THE UNIVERSITY OF CALGARY

Characterization of Wild Type and Substituted β-Glucosidases from *Aspergillus niger*

By

Heather F. Seidle

´

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN

PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

UNIVERSITY OF CALGARY

CALGARY, ALBERTA

May 2001

■✦■ National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

<div style="display:flex">
<div>

The author has granted a non-
exclusive licence allowing the
National Library of Canada to
reproduce, loan, distribute or sell
copies of this thesis in microform,
paper or electronic formats.

The author retains ownership of the
copyright in this thesis. Neither the
thesis nor substantial extracts from it
may be printed or otherwise
reproduced without the author's
permission.

</div>
<div>

L'auteur a accordé une licence non
exclusive permettant à la
Bibliothèque nationale du Canada de
reproduire, prêter, distribuer ou
vendre des copies de cette thèse sous
la forme de microfiche/film, de
reproduction sur papier ou sur format
électronique.

L'auteur conserve la propriété du
droit d'auteur qui protège cette thèse.
Ni la thèse ni des extraits substantiels
de celle-ci ne doivent être imprimés
ou autrement reproduits sans son
autorisation.

</div>
</div>

0-612-64886-9

Canadä

ABSTRACT

A 190 kDa β-glucosidase purified from commercial *Aspergillus niger* cellulase was shown to be a dimer and to be N-glycosylated (comprised mainly of mannose). It had unique pH dependent kinetics at high substrate concentrations. This was explained by the finding that the β-glucosidase is both a transglucosylase that is pH independent between pH 3 and 7, with the substrate itself being a transglucosidic acceptor at high substrate concentrations, and a glycosyl hydrolase with a pH optimum between 4 and 4.5. The main transglucosidic products of the β-glucosidase reaction at high pNPGlc and cellobiose concentrations were shown to be p-nitrophenyl-β-D-glucopyranosyl-β($1\rightarrow 6$)-glucopyranose and 4-O-gentiobiosyl-glucose, respectively. Gentiobiose was formed at high glucose concentrations. The enzyme is an exo-β-D-glucosidase and has at least 5 glucose subsites.

The gene for a β-glucosidase from *A. niger* was integrated into the genome of *Pichia pastoris*. Sequencing of a CNBr peptide and studies of the physical and kinetic properties of the recombinant enzyme showed that it was identical to the commercial β-glucosidase.

The roles of some conserved Trp and a Leu were investigated. Based on sequence alignments of glycosyl hydrolase family 3 members and on the only available family 3

structure, eight Trp (21, 49, 139, 262, 430, 551, 806, 813) and 1 Leu (426) were chosen for mutagenesis.

Analyses of the substituted β-glucosidases showed that major changes from wild type only occurred upon substituting for Trp-49 and Trp-262. Substitutions of Trp-49 caused large decreases of transglucosidic activity. Trp-49 seems to preserve the configuration at the non-reducing binding site, but adversely affects binding at other subsites, and thus the acceptor action is decreased. Substitutions of Trp-262 resulted in large decreases of the hydrolytic activity. This results from a loss of binding interactions with the non-reducing end glucose while the acceptor subsites are unaffected. Glycosidic cleavage then only takes place at very high concentrations and only transglucosylis occurs. Other more minor effects found at other positions may merit further studies. Surprisingly, Trp-430 and Leu-426, which are known to be very important at the active site of another family 3 glycosidase (*H. vulgare* β-D-glucan exohydrolase), did not seem to be important for this β-glucosidase.

ACKNOWLEDGEMENTS

To my parents

TABLE OF CONTENTS

x

List of Tables

List of Figures

Symbols & Abbreviations

Amp	ampicillin
ANTS	8-aminonaphthalene-1,3,6-trisulfonic acid
AOX	alcohol oxidase
bp	base pair
BLOSUM	Blocks substitution matrices for protein sequence comparisons
BSA	bovine serum albumin
Da	daltons
dd	double distilled
DNPC	3,4-dinitrophenyl-cellobiose
DLS	dynamic light scattering
DMF	dimethylformamide
DNA	deoxyribonucleic acid
dNTPs	deoxyribonucleotide triphosphates
DSS	2,2-dimethyl-2-silapentane-5-sulfonate
EDTA	ethylenediaminotetraacetic acid
FACE^{L}	fluorophore assisted carbohydrate electrophoresis
FPLC	fast protein liquid chromatography
g	standard gravity
GALase III	β-galactosidase
Glc	glucose
GLC	gas-liquid chromatography

HEXase III	β-N-acetylhexosaminidase
HMDS	hexamethyldisilazane
HMQC	heteronuclear multiple quantum coherence
kb	kilobase
kDa	10^3 daltons
LB	Luria-Bertani
MANase II	α-mannosidase (cleaves α1-2,3,6 linked mannose)
MANase VI	α-mannosidase (cleaves α1-6 linked mannose)
MCS	multiple cloning site
NADPH	nicotinamide adenine dinucleotide phosphate, reduced
NANaseIII	neuraminidase
NMR	nuclear magnetic resonance
NP	nitrophenol
OD	optical density
oNPGlc	o-nitrophenyl-β-D-glucopyranoside
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PNGase F	peptide N-glycosidase F
pNPGe	p-nitrophenyl-β-D-gentiobiose
pNPGlc	p-nitrophenyl β-D-glucopyranoside
psi	pounds per square inch
rpm	revolutions pre minute
SCED	sorbitol, sodium citrate, EDTA.

SDS	sodium dodecyl sulphate
TAE	TRIS acetate EDTA
TE	TRIS EDTA
TMCS	trimethylchlorosilane
TOCSY	total correlation spectroscopy
TRIS	tris(hydroxymethyl) aminomethane
TLC	thin layer chromatography
UV	ultraviolet
Vh	volt hour
Zeo^R	zeocin resistance

1. INTRODUCTION

1.1 Glycosidases

Carbohydrates are essential to all living organisms and constitute a large amount of the organic matter on earth (Kraulis *et al.*, 1989). Carbohydrates can be represented by $(C \bullet H_2O)_n$. They are constituents of DNA, RNA, many important cofactors, glycoproteins and other molecules. Monomeric sugars (monosaccharides) can be linked together via glycosidic bonds to form polymers (oligosaccharides or polysaccharides).

Due to the abundance of carbohydrates there are many enzymes that catalyze reactions involving glycosidic bonds. These enzymes transfer a glycosyl group from a donor sugar to an acceptor. If the acceptor is water, the result is hydrolysis and the enzymes that catalyze the reactions are referred to as glycosidases or glycosyl hydrolases. If the acceptor is another sugar or some other group than the enzyme is termed a glycosyl transferase (Ly & Withers, 1999).

The universal nature of oligosaccharides and polysaccharides has resulted in most organisms producing glycosidases for carbohydrate degradation. For example, most organisms contain α-amylases, cyclodextrin glucanotransferases, debranching enzymes, exo-α-amylases, glucoamylases and α-glucosidases. These are all involved in starch or glycogen degradation (Svensson *et al.*, 1989). Cells also contain enzymes that break down disaccharides and other oligosaccharides. In addition, many organisms produce glycosidases for more specific purposes. The phytopathogenic bacterium *Pseudomonas solanacearum* utilizes an endoglucanase to facilitate penetration and breakdown of host plant tissues resulting in a lethal wilting disease (Roberts *et al.* 1988). The sweet pea

2

plant (*Lathyrus adoratus* L.) uses cellulases to break down its own cell walls so that further development can proceed (Neelam and Sexton, 1995). Cytosolic β-glycosidases found in many species are thought to be involved in the metabolism of xenobiotic compounds (Hays *et al.*, 1998). In the kidney they may be involved in the absorption of sugars by the proximal convoluted tubules (Huber and Brockbank, 1988).

The above are just a few examples that show the importance of glycosidases and the varied roles that they can play. Because this type of enzyme has so many functions, deficiencies or mutations that affect the functions of these enzymes could create many problems. Heritable deficiencies in glycosyl hydrolases are among the most frequent genetically based syndromes in man (Henrissat, 1991). Lactose intolerance results from the accumulation of lactose in the lumen of the small intestine because of a lack of lactase (Stryer, 1995). Tay-Sachs disease results from a deficiency of β-hexosaminidase, an enzyme that removes β-1,4-linked N-acetylhexosamine residues from oligosaccharides (Mark *et al.*, 2001). In addition there are several glycogen storage diseases that result from glycosidase enzyme deficiencies. Pompe's disease results from a lack of α-1,4-glucosidase and Cori's disease results from a deficiency of amylo-1,6-glucosidase. (Stryer, 1995). Gaucher's disease results from the accumulation of glucosyl ceramide in monocytes or macrophages due to mutations of acid β-glucosidase (also referred to as glucocerebrosidase) (Grace *et al.*, 1997). Mutations in the α-galactosidase gene result in the accumulation of neutral glycosphingolipids in tissues and fluids, a condition known as Fabry's disease (Maranville & Zhu, 2000).

Glycosyl hydrolases are also used in industry. Approximately 12 000 tonnes of glycosidases are made per year for use in the food industry (Bains, 1998). The major enzymes used for this purpose are amylases, isoamylases, pullanases and cellulases. Amylases break starch down to smaller polymers and eventually to glucose. Both isoamylases and pullanases are starch debranching enzymes and are also used for starch degradation. Unbranched molecules feel different to our mouths than branched molecules which makes these enzymes important for controlling food texture (Bains, 1998).

Cellulases break down cellulose. Since cellulose is probably the most common biological material in the world ((70 kg of cellulose is synthesized per person per day (Coughlan, 1985)). These enzymes are very important. They are utilized for a variety of applications including: bleaching of fabrics (Klahorst *et al.*, 1994), manufacture of food, paper production (Torronen *et al.*, 1993), aroma enrichment, debittering of wines and other beverages (Shoseyov *et al.*, 1988), and bio-stoning of denim (Kumar *et al.*, 1997).

1.1.1 Glycosidase mechanisms

Glycosidases catalyze the cleavage of glycosidic linkages. There are two stereochemical possibilities that can occur at the anomeric center when glycosidic bonds are broken. If the anomeric hydroxyl group retains the same configuration as did the glycosidic bond of the substrate, the enzyme has a retaining mechanism. If the configuration of the product is inverted with respect to the substrate, the enzyme reaction occurs with inversion of configuration. The mechanism for both retaining and

4

inverting glycosidases were proposed by Koshland almost 50 years ago (Koshland,

1953). Both mechanisms are supported by a large amount of experimental evidence.

In the retaining mechanism, catalysis occurs by a two-step double displacement

mechanism. This involves the formation (glycosylation step) and hydrolysis

(deglycosylation step) of a covalent glycosyl-enzyme intermediate (Figure 1.1). Both the

glycosylation and the deglycosylation steps require an oxocarbenium ion-like transition

state for the reaction to progress. There are usually 2 carboxylic acids involved and they

are separated by approximately 5.5 Å (Ly and Withers, 1999). One of these forms a

covalent glycosyl-enzyme species by attacking the anomeric center of the sugar and is

referred to as the catalytic nucleophile. The other carboxylic acid participates in both

general-acid and general-base catalysis. It acts as an acid catalyst in glycosylation by

protonating the glycosidic oxygen. It acts as a general base in deglycosylation by

deprotonating the water. Therefore, the second carboxylic acid is referred to as the

acid/base catalyst.

The mechanism for transferolysis is very similar to the retaining glycosyl

hydrolase mechanism described above (Figure 1.1). Essentially, the glycosylation

portion of the reaction is the same. The difference occurs in deglycosylation. Instead of

hydrolysis of the glycosyl-enzyme intermediate by general base catalyzed activation of a

water molecule, the general base is thought to activate an alcohol group of a different

molecule. The activated hydroxyl attacks the anomeric carbon resulting in the formation

of a larger oligosaccharide. Since glycosyl transferolysis is so similar to glycosyl

hydrolysis, many enzymes can perform both activities. Transferolysis can occur with a

Figure 1.1. The proposed double-displacement mechanism for a retaining
glycosidase. In the first half of the reaction the enzyme is
"glycosylated". The catalytic nucleophile (a carboxyl group)
attacks the sugar anomeric center while the other carboxyl group
acts as an acid catalyst and protonates the glycosidic oxygen. This
results in the formation of a glycosyl-enzyme covalent
intermediate. The oxocarbenium ion-like transition state is
represented by (‡). In the second half of the mechanism the
enzyme is "deglycosylated". If this occurs via a water molecule it is
referred to as hydrolysis (R' = H). If R' is a different molecule the
enzyme is deglycosylated by transferolysis (if R' is a second sugar
the reaction is called transglycosylation). Regardless of the the
identity of the R' group, it is thought to be deprotonated by a
carboxyl group acting as a base catalyst. A second oxocarbenium
ion-like transition state (‡ ') occurs and the sugar is released by the
glycosidase and the anomeric hydroxyl retains the anomeric
configuration of the original glycosidic bond. It is usually thought
that the 2 oxocarbenium ion-like transition states are identical.
Adapted from Ly and Withers, 1999).
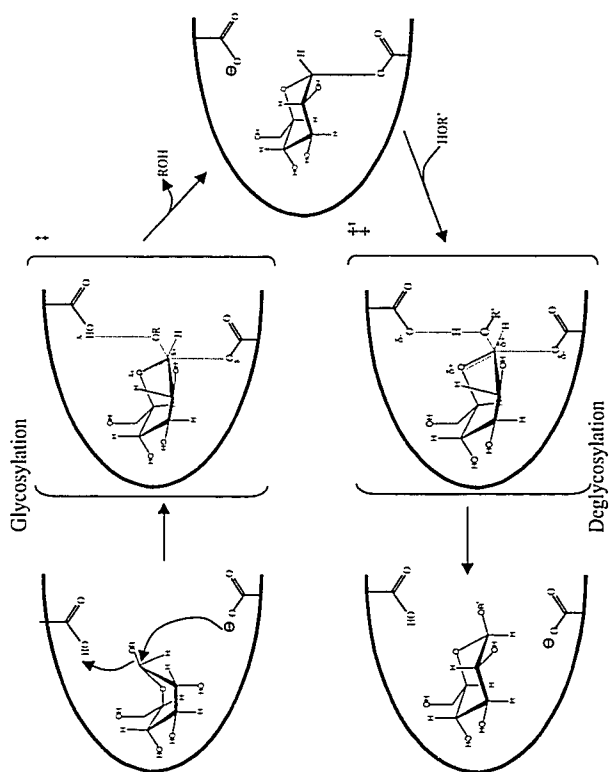
variety of different alcohols. If the alcohol is a sugar, the reaction is called transglycosylation. If the transfer reaction is specific to a sugar type (e.g. glucose) the name of the specific sugar is included in the reaction name (transglucosylation).

Inversion of the configuration at the anomeric center occurs when a glycosidase proceeds by a direct displacement reaction. In this mechanism, 2 carboxyl groups are also required and an oxocarbenium ion-like transition state occurs (Figure 1.2). However, the carboxylic acids are situated farther apart (~9.5 Å). This arrangement, provides enough room for both the substrate and water to bind together. Therefore, one of the carboxylic acids can provide general base catalytic assistance to the water while the other carboxylic acid contributes to general-acid catalysis resulting in cleavage of the glycosidic bond (Ly and Withers, 1999).

1.2 Glycosidase families

In recent years a large amount of information has been gathered on glycosidases. More than 1000 glycosidases have been sequenced and the sequences have been analyzed for similarities. Based on these analyses, glycosidases have been divided into 85 families. Generally, families contain several types of glycosidases, but since the sequences within a family are similar, the structures are likely to be similar. Table 1.1 shows some of the features of the different families that contain at least one enzyme whose structure has been determined. Proteins that belong to a particular family are likely to be functionally similar and are derived from a common ancestor (Henrissat & Bairoch, 1996). The PROSITE numbers given in the table can be used to obtain information regarding some of the glycosyl hydrolase families
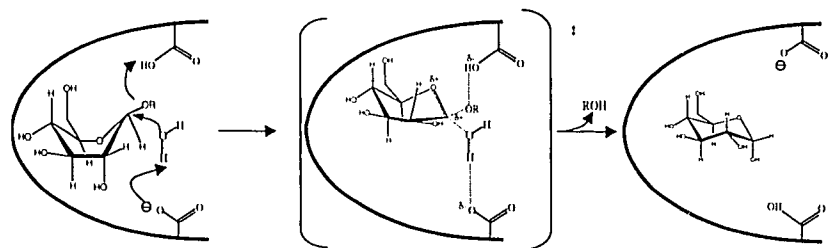
Figure 1.2. The proposed direct displacement mechanism for an inverting glycosidase. In the first step, one of the
carboxylic acids in the active site provides general base catalytic assistance to the attack of water. The
other carboxyl group provides general acid catalytic assistance for cleavage of the glycosidic bond. The
oxocarbenium ion-like transition state is indicated by (‡). Generally there is only enough room for the
substrate and water to bind at the same time. Therefore, transglycosylation is unlikely to occur by this
mechanism. Adapted from Ly and Withers (1999).

Table 1.1 Glycosyl hydrolase families that have at least one structure completed.

Family	Enzyme Types	PROSITE	Retention/Inversion
1	β-glucosidase, 6-phospho-β-galactosidase and β-glucosidase, lactase-phlorizin hydrolases, myrosinases	PDOC00495	retention
2	β-galactosidases, β-glucuronidases, β-mannosidases	PDOC00531	retention
3	mainly β-glucosidases	PDOC00621	retention
5	endoglucanases, β-mannanases, exo-1,3-glycanases	PDOC00565	retention
6	endoglucanases, cellobiohydrolases	PDOC00563	inversion
7	endoglucanases and cellobiohydrolases	none	retention
8	endoglucanases	PDOC00640	inversion
9	endoglucanases	PDOC00511	inversion
10	xylanases	PDOC00510	retention
11	xylanases	PDOC00622	retention
13	α-amylases, pullulanases, cyclomaltodextrin glucanotransferase, cyclomaltodextrinase, trehalose-6-phosphate hydrolase	none	retention
14	β-amylases	PDOC00414	inversion
15	glucoamylases	PDOC00646	inversion
16	β-glucanases (lichenases)	PDOC00794	retention
17	glucan endo-1,3-β-glucosidases, lichenases	PDOC00507	retention
18	chitinases, endo-β-N-acetylglucosaminidases, di-N-acetylchitobiases	PDOC00839	retention
19	chitinases	PDOC00620	inversion
20	β-hexosaminidases, chitobiases	none	retention
22	lysozymes type C, α-lactalbumins	PDOC00119	retention
23	lysozymes type G	none	not known
24	phage lysozymes	none	inversion
33	sialidases	none	retention
34	neuraminidases (sialidases)	none	retention
45	endoglucanases	PDOC00877	inversion
46	chitosanases	none	inversion
48	endoglucanases	none	inversion
77	amylomaltase, 4-α-glucanotransferase	none	retention
82	i-carrageenase	none	inversion
83	hemagglutinin-neuraminidase	none	retention

(http://expasy.cbr.nrc.ca/prosite/). PROSITE is a database of protein families and domains and is a way of grouping proteins that exhibit sequence similarities. Table 1.1 also indicates whether an enzyme family proceeds with retention or inversion of anomeric configuration. An updated list of the glycosidase families can be searched at http://www.expasy.ch/cgi-bin/lists?glycosid.txt or at http://afmb.cnrs-mrs.fr/~pedro/CAZY/.

Although these enzymes are all involved in hydrolysis of glycosidic bonds, they vary in size, subunit composition and carbohydrate specificities. The carbohydrate binding regions have much less diversity. The majority of glycosyl hydrolases contain a pocket or a cleft for substrate binding. The glycosyl hydrolase families have been divided into 11 clans based on common structural folds. The clans are referred to as GH-A to K (glycosyl hydrolase clan A, etc.) (Henrissat *et al.*, 1995). The structural folds that have been named and identified are the $(\beta/\alpha)_8$ barrel (there are 3 different folds of this type), β-jelly roll, β-propeller and the α+β fold (http://afmb.cnrs-mrs.fr/%7Epedro/CAZY/ghf_table.html#clans). The α/β-barrel structure, consisting of eight β-strands twisted into a cylinder surrounded by eight α-helices, is common to carbohydrate interacting enzymes. It is commonly referred to as a TIM barrel (for triose phosphate isomerase). In this structure the carbohydrate binding site is generally located in a cavity at the C-terminal end of the parallel strands (Barrett *et al.*, 1995). In Figure 1.3 the α/β-barrel structure is represented in (A) β-galactosidase, (B) β-D-glucan exohydrolase, (C) endoglucanase, and (H) amylomaltase. The TIM barrels in (A) and (C) are both representatives of clan GH-A, which is also known as the 4/7 superfamily

Figure 1.3. Several examples of glycosyl hydrolase structures illustrating
common carbohydrate binding motifs. The protein name, organism
of origin, glycosyl hydrolase family, glycosyl hydrolase clan and
primary accession number are given. All structures were obtained
from the protein data bank (PDB), and the PDB numbers are shown
after the primary accession numbers. The structures are: A) β-
galactosidase from *Escherichia coli*, family 2, clan GH-A (P00722)
(1DPO), B) β-D-glucan exohydrolase isoenzyme I from *Hordeum
vulgare*, family 3 (Q9XE13) (1EX1), C) Endoglucanase 5A from
Bacillus agaradhaerens, family 5, clan GH-A (O85465) (1A3H),
(D) Neuraminidase from Influenza virus Aternaustraliag70c75,
family 34 (P06818) (7NN9), E) κ-carrageenase from *Alteromonas
carrageenovora*, family 16, clan GH-B (P43478) (1DYP), F)
Lysozyme C from *Gallus gallus* (chicken), family 22 (P00698)
(194L), G) Chitosanase from *Streptomyces sp.* (strain N174),
family 46, clan GH-I, (P33665) (1CHK), H) 4-α-glucanotransferase
(amylomaltase) from *Thermus aquaticus*, family 77, clan GH-H,
(O87172) (1CWY).

A)   β-Galactosidase (*E. coli*)
     family 2, clan GH-A

B)   β-Glucan exohydrolase (*H. vulgare*)
     family 3

C)   Endoglucanase
        (*B. agaradherans*)
        family 5, clan GH-A

D)   Neuraminidase
        (avian influenza virus)
        family 34, clan GH-E
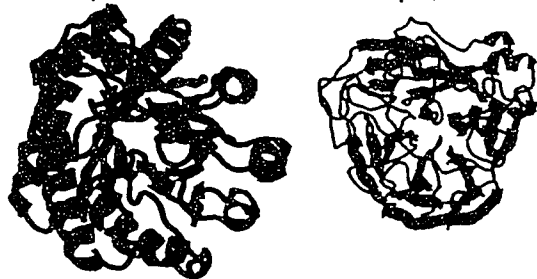
E) κ-Carrageenase   F) Lysozyme (*G. gallus*)   13
    (*A. carrageenovora*)   family 22
    family 16, clan GH-B

G) Chitosanase (*Streptomyces sp.*)   H) Amylomaltase (*T. aquaticus*)
    family 46, clan GH-I   family 77, clan GH-H
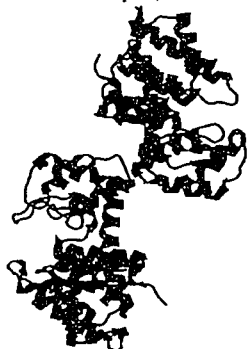
(the catalytic acid /base is on strand 4 and the catalytic nucleophile is on strand 7). The
β-D-glucan exohydrolase in (B) is not a member of a clan at this time and (H) is a TIM
barrel from clan GH-H. The β-jelly roll fold is shown in κ-carrageenase in (E), which is
a member of the GH-B clan. The neuraminidase in (D) has a β-propeller fold and is a
member of Clan GH-E. It is made up of β-sheets that are arranged in a propeller-like
formation and is referred to as α+β because enzymes belonging to this class contain α-
helixes and β-strands that do not entwine (Quiocho, 1986). In Figure 1.3 the α+β class
of glycosidase folds is represented by lysozyme (F) and chitosanase (G).

1.1.2 Glycosidase substrate binding and specificity

1.1.2.1 Carbohydrate-protein interactions

Carbohydrate binding motifs can be separated into two main structural groups.
The first group (group I) has deep clefts or cavities for carbohydrate binding. In this
group, the sugar is sequestered from the solvent and is bound with high affinity (Weis &
Drickamer, 1996). Most enzymes and bacterial sugar transport proteins are in this group.
The second group (group II) involves proteins that have shallow carbohydrate binding
sites leaving much of their ligand solvent exposed. These proteins tend to bind with a
lower affinity (Weis & Drickamer, 1996). Anti-carbohydrate antibodies, toxins, viral
proteins and galectins are members of this second group.

In the early twentieth century Emil Fischer stated, " I would like to say that the
enzyme and the glucoside have to fit to each other like a lock and a key" (Quiocho,
1986). Generally, this view of carbohydrate-protein interactions is appropriate with the
addition that conformational changes occur at the active site and result in an "induced

fit" of the substrate. Through X-ray crystallography and NMR, interactions between many proteins and their carbohydrate ligands have been determined to high resolution. Even the role of water molecules in carbohydrate binding has been visualized by these techniques. Water interactions in carbohydrate binding are important because the interactions between sugars and amino acid residues are often mediated by water (via hydrogen bonds).

Hydrogen bonds are the most prominent interactions between carbohydrates and protein (Weiss & Drickamer, 1996). The hydroxyl groups of the sugar can act as both hydrogen bond donors and acceptors and therefore can interact with several amino acids and/or water molecules. The oxygen of the sugar ring can also accept two hydrogen bonds. The glycosidic oxygen forms hydrogen bonds less often since it is a very strong binding partner and these groups tend to actually remove the hydrogen from the donating group rather than sharing it in a hydrogen bond. Thus, hydrogen bonds to glycosidic oxygens can result in glycosidic bond cleavage, although this is not always the case.

A large network of hydrogen bonds is a common feature of group I carbohydrate-protein interactions (Weiss & Drickamer, 1996). In this group the sugar is mostly buried in the binding site. There is high affinity binding because each hydroxyl group has at least one bonding partner. The bonds to the transition state are similar to the ground state condition but they are more numerous and stronger. Van der Waals interactions are also very important in carbohydrate-protein binding and there are three types that occur; polar-polar, polar-nonpolar and nonpolar-nonpolar (Quiocho, 1986). Polar-polar interactions

16

occur between the sugar hydroxyls and polar groups of the protein. Thus, they resemble hydrogen bonds except that the distance between the interacting groups is greater than the maximal 3.4 Å distance allowed for hydrogen bonding. The hydroxyl groups may also interact with nonpolar carbon atoms of amino acid side chains forming polar-nonpolar van der Waals interactions. The majority of nonpolar-nonpolar associations are made with aromatic side chains, a phenomenon referred to as stacking. Stacking occurs because most sugars have one face that is less polar than the other and is therefore, able to pack against the faces of Trp, Phe and/or Tyr. The interaction is not only due to complementary non-polar surfaces but is also thought to result from the proximity of the aliphatic protons of the sugar ring which carry small net positive partial charges and the π-electron cloud of the aromatic ring (Weis & Drickamer, 1996). This stacking interaction is known to occur in many carbohydrate binding sites and is significant in both sugar binding and recognition. Trp and Tyr are common in sugar binding sites since they can interact with sugars via stacking in addition to forming hydrogen bonds.

Specificity for different sugars is usually a result of the placement of the amino acid side chains and through steric exclusion. The placement of amino acids indicates whether axial or equatorial hydroxyl groups are favored in the ligand. For example, the specificity of the mannose-binding protein towards mannose results from hydrogen bonds between the sugars' equatorial 4-hydroxyl group and an Asn. If this Asn is substituted with an Asp, the binding becomes specific for an axial 4-hydroxyl group. The protein then binds galactose but not mannose (Drickamer, 1997). Other interactions in the binding site may increase the stability of sugar binding but are not as important in selectivity.

Selectivity for oligosaccharides arises from a primary monosaccharide binding specificity that is enhanced through multiple binding of additional sugar molecules in subsites. This is due to the cooperative effects of each bond, which in turn is partly a result of more favorable entropic effects (Quiocho, 1986). These subsites may also show hydrogen bond specificities similar to the main binding site or their interactions may result from correct spacing. Proteins tend to show more selectivity for their ligands when the oligosaccharides are buried or mostly buried in their binding sites. Generally, this results from steric exclusion and additional hydrogen bond and van der Waals interactions. Furthermore, sugar stacking prohibits the binding of sugar epimers because of steric interference (Quiocho, 1993). An aromatic amino acid would not be positioned correctly to allow the sugar to stack if a hydrogen is replaced with a hydroxyl group.

1.1.2.2 Analysis of carbohydrate binding sites by site specific mutagenesis of Trp

Probing the role of individual residues involved in substrate binding by site directed mutagenesis is a very useful method for explaining structure-function relationships of enzymes. As described above, Trp is important in carbohydrate-protein interactions. The importance of these interactions has been demonstrated by structural data that show that Trp at the active sites of proteins are often in a position to interact with the sugar substrates (Koivula et al., 1998). Additionally, Trp is important for catalysis of some glycosidases because it maintains the non-polar environment that is required at the active sites of these glycosidases (Mark et al., 2001). Since Trp can be involved in carbohydrate associations, it is a good candidate for investigation via mutagenesis. Examples of these

investigations illustrate some of the information that can be gained from studies in which conserved Trp have been substituted by other residues.

The cellulose binding domains (CBDs) of various cellulases have been examined for their ability to bind amorphous or crystalline cellulose. The CBD of *Pseudomonas fluorescens* subsp. *cellulosa* xylanase A contains a Type IIa CBD. These CBDs have 3 strictly conserved Trp on the surface (Nagy *et al.*, 1998). These Trp form a hydrophobic strip on the CBD that allows it to interact with cellulose (Nagy et al, 1998). When the xylanase CBD conserved Trp were individually substituted with Ala, the CBD showed at least 20 times less affinity for crystalline and amorphous cellulose and for cellohexaose. The structure for this CBD has been determined and the Trp side chains are co-planar. They are thought to form strong stacking interactions with the glucose pyranose rings (Nagy et al, 1998).

Trp-62 of hen egg white lysozyme has been shown to be important. The non-polar surface of the sugar residue in subsite B makes substantial contact with the indole ring of Trp-62 (Maenaka *et al.*, 1994). The importance of this residue has been investigated by site-directed mutagenesis to other aromatic residues (Kumagai, 1989) and to non-aromatic residues (Maenaka *et al.*, 1994). The results of these substitutions suggested that this residue has a role in defining substrate specificity, through steric hindrance. Mutations to non-aromatic residues resulted in a greater proportion of non-productive substrate binding (the K_m is about 6 times higher). Additionally, substitutions for Trp-62 resulted in decreased hydrolytic activity with non-charged substrates (glycol chitin and (PNP-

GlcNAc)₃). These results suggested that Trp-62 is important for carbohydrate recognition and for maintaining a non-polar environment for catalysis (Maenaka *et al.*, 1994).

The maltodextrin-binding protein has 4 aromatic residues that interact with maltose (Quiocho, 1993). Through structural studies it is thought that one of these, Trp-230, has van der Waal's contacts with the reducing-end glucose. A 12-fold decrease in the affinity for maltose was observed when Trp-230 was substituted with Ala. This was equivalent to a decrease of 1.5 kcal mol⁻¹ in binding energy (Quiocho, 1993). The decrease in binding energy was interpreted as the loss of the van der Waal's contacts (0.5 kcal mol⁻¹ per contact, 3 contacts) due to replacement of Trp-230 (Quiocho, 1993).

α-Galactosidase is an exoglycosidase that cleaves α-linked galactose monomers from glycomolecules (Maranville & Zhu, 2000). Trp-16 has been shown to be important in coffee bean α-galactosidase (Zhu *et al.*, 1996). Trp-16 was substituted by every other amino acid at that position. Only conservative mutations (Tyr and Phe) allowed activity. They had about 10 times less affinity for the substrate (4-methylumbelliferyl-α-galactose) than wild type (Maranville & Zhu, 2000). Substitutions with Met or Asn showed much less activity than wild type and substitutions with Lys or Cys showed only faint activity. Substitutions by all other amino acids abolished enzyme activity. The data indicated that the indole ring is not essential for catalysis since the aromatic substituted enzymes still showed significant amounts of activity. Therefore, it is likely that Trp-16 provides a hydrophobic environment that alters the pK$_a$ of a nearby carboxyl group and can be replaced by other aromatic amino acids (Maranville & Zhu, 2000).

1.2 Cellulose

Cellulose is the major polysaccharide in plants, where it is part of the cell walls (Rouvinen *et al.*, 1990). Termites, a few species of cockroaches, and ruminants (cows, sheep, goats, camels, etc.) can all use cellulose as a food source, as they either contain cellulose degrading enzymes or microorganisms containing these enzymes (Kotz and Purcell, 1987). The enzymatic conversion of cellulose to glucose is far from simple due to the complex nature of the substrate. Substrates contain both crystalline and amorphous regions (Penttila *et al.*, 1986). Native cellulose is an insoluble polymer composed of long, linear chains of β-1,4-linked glucose units. Hydrogen bonding and van der Waals forces create an interchain binding network associating 60 to 70 single chains into elementary fibrils. These fibrils aggregate to larger bundles that form ordered crystals (Reinikainen *et al.*, 1992). Furthermore, the fibers are embedded in a matrix of hemicellulose and lignin which further complicates the conversion as it reduces the availability of the cellulose fibers to cellulose degrading enzymes (Beguin, 1990). Cellulose degrading enzymes have been termed "cellulases". In most cellulolytic microorganisms, several enzyme components are involved in converting cellulosic substrates to glucose by their cooperative and synergistic action (Okada, 1985). This group of enzymes is often called a "cellulase system".

1.2.1 Cellulases

The cellulolytic enzymes are usually secreted into the medium by the organism. Enzymes such as glucohydrolases, cellobiose oxidases and cellobiose dehydrogenases play a role in cellulolysis, but the main enzymes involved are cellobiohydrolases (1,4 β - D-glucan cellobiohydrolase), endoglucanases (1,4 β -D-glucan 4-glucanohydrolase) and

β-glucosidases (β -D-glucoside glucohydrolase) (Wood, 1992). Classification of enzymes

found in cellulase systems is based on their activity with different substrates. Some

cellobiohydrolases (CBH) and endoglucanases appear to be relatively nonspecific.

Regardless, cellobiohydrolases are described as exo-hydrolases and remove cellobiose

from the non-reducing ends of cellulose chains. Endoglucanases are endo-hydrolases and

cleave cellulose chains randomly. β-Glucosidases are perhaps the easiest to classify as

they hydrolyze cellobiose and some soluble cello-oligosaccharides to glucose (Wood,

1992). It is common for microorganisms to have several of the same types of cellulase

enzymes. *Humicola insolens* produces 7 different cellulases (5 endoglucanases and 2

cellobiohydrolases) (Schülein, 1998). It is interesting to note that these 7 cellulases belong

to 5 different glycosyl hydrolase families and that they are not just isozymes (Schülein,

1998).

Because these enzymes all have similar substrates (made up of glucose) they all

contain a specific number of glucose subsites where each glucose of the substrate binds to

the protein. Previously, numbering systems were confusing since different investigators

chose to number the glucose subsites differently (e.g. lysozyme subsite B (Maenaka *et al.*,

1994) does not provide the reader with much information about the subsite unless one

studies lysozyme). Recently, this problem was noted and a subsite nomenclature system

was designed (Davies *et al.*, 1997). The glucose subsites are labeled left and right from the

scissile glycosidic bond. The glucose subsites on the left (towards the non-reducing end)

are numbered from the subsite closest to the scissile bond, which is −1. The next subsite

would be −2 etc., until a glucose on the non-reducing side of the sugar polymer was bound

at the last subsite available (-n). The glucose subsites following the scissile glycosidic bond (on the right side) are numbered starting with +1, which is followed by +2 etc. until the last subsite available for the reducing side is +n. Thus, the glycosidic bond to be cleaved is always located between the –1 and +1 subsites (Davies *et al.*, 1997).

The mechanisms of these cellulolytic enzymes are similar to those of other glycosidases (Gilkes *et al.*, 1991). Cleavage of the glycosidic bond can proceed with either inversion or retention of configuration. Due to the increasing interest in cellulase enzymes, mostly due to interest from the biotechnology sector, a large number of these enzymes have been investigated. The majority of the research, however, has been on the cellulase systems of *Trichoderma reesei*, *Cellulomonas fimi* and *Clostridium thermocellum*. Six cellulase enzymes have been identified from *Trichoderma reesei*, 4 endoglucanases and 2 cellobiohydrolases (Divne *et al.*, 1998). Of the six enzymes the most intensely studied has been cellobiohydrolase I (CBH I), because it is very proficient at degrading microcrystalline cellulose (Abuja *et al.*, 1988). Like all CBHs, CBH I is made up of 2 domains, a catalytic domain and a cellulose binding domain. The enzyme cleaves cellobiose from the reducing end of cellulose chains, with retention of the anomeric configuration. The structure has been determined and is a large β-sandwich. The cellulose binding site is a long tunnel (~50Å) formed by loops that extend out from the β-sandwich. The cellulose chain threads through the tunnel (Divne *et al.*, 1998). There are 11 glucose subsites that have been identified and they are arranged from –7 to +4 (thus there are 7 glucose bound before the glycosidic bond cleavage site and 4 bound after the cleavage site). Interestingly, this enzyme cleaves from the reducing end rather than the non-

reducing end (Divne *et al.*, 1998). However, the subsites are still numbered according to the nomenclature system discussed above. The probable active site residues have been identified as Glu-212 and Glu-217, where Glu-212 is the active site nucleophile and Glu-217 is the catalytic acid/base (Divne *et al.*, 1994). Cellobiohydrolase II is structurally different from CBH I. Although it is still a tunnel (20 Å), it is formed by 2 loops and a TIM barrel (Teeri *et al.*, 1998) rather than a β-sandwich. Additionally, it only contains 6 glucose subsites, 2 at the non-reducing end of the glucose chain and 4 at the reducing end (Teeri *et al.*, 1998). Cleavage occurs from the non-reducing end of the glucose chain. This enzyme cleaves the glycosidic bond with inversion of configuration at the anomeric carbon. The proposed proton donor is Asp-221 (Rouvinen *et al.*, 1990) but it is unclear if Asp-263 or Asp-401 is the catalytic base. The *Trichoderma reesei* endoglucanases have not been investigated to the same extent as the cellobiohydrolases and won't be discussed here.

The cellulase enzymes from *C. thermocellum* have also been investigated by a number of researchers. One reason for this is that there are a huge number of cellulolytic enzymes produced by these organisms. Many of these proteins form a high molecular weight complex, which has been called a cellulosome (Lamed *et al.*, 1983). The enzymes that have been found in the cellulosome are: 21 endoglucanases, 3 exo-glucanases (cellobiohydrolases), 2 β-glucosidases, 6 xylanases, 2 lichenases and 1 xylan esterase (Beguin, 1998). Each of these components contains a catalytic domain and a region of 65 conserved amino acids referred to as a dockerin domain (Beguin *et al.*, 1996). The dockerin domain anchors the catalytic components to a noncatalytic scaffolding subunit-

CipA (Beguin *et al.*, 1996). This assists in the attachment of the cellulases to the cellulose. The Cip A protein interacts with the dockerin domains via a cohesion domain (see Bayer *et al.*, (1994) for a good review of cellulosomes). Several of the endoglucanases from the cellulosome have been studied in detail. Cel A (a family 8 glycosyl hydrolase) has an $(\alpha/\alpha)_6$ barrel catalytic domain that has 6 glucose subsites, 3 on either side of the scissile bond (Alzari *et al.*, 1996). The glucose subsites are not positioned in a straight line, and this forces the substrate to bend at the glycosidic bond between the -1 and $+1$ subsites. It was also found that the glucose on either side of the bend were in the same orientation (Alzari *et al.*, 1996). This can only occur if there is $180°$ rotation at the bent glycosidic bond. The bending of the substrate and the rotation at the scissile bond together result in a lower activation energy for cleavage of the bond (Alzari *et al.*, 1996). However, since the enzyme evolved this structure these interactions are now required for bond cleavage to occur. Thus, a trisaccharide is the smallest sugar that can be hydrolyzed by this enzyme because 3 glucose are necessary to cause the bending at the scissile bond (Alzari *et al.*, 1996). Hydrolysis occurs via inversion of configuration. Glu-95 is known to be the proton donor, but the catalytic base has not been determined. Cel D is also an endoglucanase (glycosyl hydrolase family 9) that is part of the cellulosome. It is similar to Cel A since it also has an $(\alpha/\alpha)_6$ barrel catalytic domain that has 6 glucose subsites (Beguin, 1998). Unlike Cel A, the glucose binding subsites in Cel D are positioned in a straight line in a cleft. The enzyme has no trouble hydrolyzing substrates with only 3 glucose units since the bending required for the Cel A reaction is not needed (Beguin, 1998). Cel D reacts with inversion of configuration, and Glu-555 is known to be the proton donor (Tomme *et*

al., 1992) and Asp-201 is the general base. Another endoglucanase from *C. thermocellum*
is Cel C (family 5 glycosyl hydrolase). This endoglucanase is not part of the cellulosome
and is quite different from Cel A and Cel D discussed above. It has a TIM barrel structure
and is a member of the 4/7 superfamily (Clan GH-A) (Dominguez *et al.*, 1995).
Hydrolysis occurs with retention of the anomeric configuration. Glu-280 is the catalytic
nucleophile (Wang *et al.*, 1993) and Glu-140 is the proton donor (Navas and Beguin,
1992).

A number of cellulase enzymes have also been identified from *Cellulomonas fimi*.
Probably the most studied enzyme is Cex, which is a family 10 glycosyl hydrolase. Cex
has both xylanase and glucanase activities (Birsan *et al.*, 1998). This broad substrate
specificity has been observed for other members of this family
(http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-e+[BRENDA-ECNumber:'3.2.1.8']), which is
also a member of the GH-A clan (4/7 superfamily). The enzyme carries out hydrolysis
with retention of configuration. A mechanism based inactivator was used to trap the
glycosyl-intermediate and identify the catalytic nucleophile (Glu-233) (Tull *et al.*, 1991).
The catalytic acid was identified as Glu-127 (MacLeod *et al.*, 1994). It has also been
determined that the glycosyl-enzyme intermediate is stabilized by a covalent bond
(Notenboom *et al.*, 1998).

1.2.2 β-Glucosidases

β-Glucosidases are often referred to as cellobiases (Kovar *et al.*, 1987). They
hydrolyze the β1-4 glycosidic bond of cellobiose. They are not usually active with
cellulose itself and are not, technically, cellulase enzymes. However, β-glucosidases play

a pivotal role in cellulose degradation since they are involved in the final step that forms glucose. In addition, since they hydrolyze cellobiose, they remove the inhibitory effect that cellobiose usually has on the other cellulase enzymes (Kuriyama *et al.*,1995).

β-Glucosidases are found in many prokaryotes and eukaryotes. They have a variety of functions depending on their location and substrate specificity. In humans, acid β-glucosidase is responsible for the catabolism of glucocerebroside (glycosphingolipid glucosylceramide) (Mikhaylova *et al.*,1996). Some β-glucosidases in plants are utilized in the chemical defense against pathogens and herbivores. Toxic compounds are released from cyanogenic glucosides, or arbutin when they are hydrolyzed. Others are utilized in the regulation of phytohormones such as cytokinin, gibberellin, and auxin, by causing the release of the active forms from the inactive hormone-glucoside conjugates (Leah *et al.*,1995). In many fungi and bacteria, β-glucosidases are utilized in the conversion of cellobiose to glucose. Many of these enzymes are also capable of utilizing other sugars as acceptors so that transglycosylation occurs rather than hydrolysis. This additional activity has been noted in mammal (Vanderjagt *et al.*, 1994), plant (Kuriyama *et al.*, 1995), fungal (Christakopoulos *et al.*, 1994), and bacterial (Watt *et al.*, 1998) β-glucosidases. There has been some interest in using transglycosylation for production of oligosaccharides. This could provide a more economical method of production since it occurs in one step and the sugar product can be altered depending on the initial substrate and the acceptor (Yazaki *et al.*,1997). Additionally, β-glucosidases can be used in ethanol production by including an ethanol producing organism in the saccharification process (Hoh *et al.*,1992). Studies of β-glucosidases that function in cellulose breakdown are

important because the degradation of cellulose to glucose may provide a solution to resource problems involving fuel, feedstock and food materials that may occur in the future (Skory *et al.*, 1996).

1.3 *Aspergillus*: a prolific β-glucosidase producing fungi:

Aspergillus is a fungal genus that secretes a wide range of carbohydrate-hydrolyzing enzymes. It is of the subdivision Ascomycotina, referring to its ability to produce a structure called the "ascus" that holds spores during sexual reproduction (Deacon, 1984). *Aspergillus* is a genus of fungi that is associated with food spoilage and plant diseases (Stewart & Parry, 1981). *Aspergillus* species lack chlorophyll and must, therefore, be capable of obtaining nutrients from their environment (Boyd, 1988). They live primarily on dead organic matter and use it as a source of carbon for cellular synthesis and energy. Due to their rigid cell wall they cannot engulf food. Instead they absorb simple soluble nutrients that are obtained from complex polymers by releasing extracellular enzymes into the environment (Deacon, 1984).

The *Aspergillus* genus is known to be an excellent producer of β-glucosidase(s) and the β-glucosidase(s) from *Aspergillus niger* have been found to be highly active with cellobiose (Woodward and Wiseman, 1982). This property has promoted interest in the *Aspergillus* β-glucosidase(s) as a supplement to the saccharification processes to increase the reaction efficiency (Gupte and Madamwar, 1997).

1.3.1. β-Glucosidases isolated from *Aspergillus*

β-Glucosidases have been isolated from *Aspergillus phoenicis* (Duff, 1985), *A. terreus* (Pushalkars, 1995), *A. heteromorphus* (Solovyeva *et al.*, 1997), *A. nidulans* (Lee

et al., 1996), *A. wentii* (Srivastava *et al.*, 1984, Bause and Legler, 1980), *A. fumigatus*

(Rudick and Elbein, 1973, 1975, Ximenes *et al.*, 1996), *A. ellipticus* (Gupte and

Madamwar, 1997), *A. japonicus* (Sanyal *et al.*, 1988, Solovyeva *et al.*, 1997), *A. oryzae*

(Roiu *et al.*, 1998), *A. sojae* (Kimura and Tajima, 1999, Kimura *et al.*, 1999), and *A. niger*.

The enzymes from *A. phoenicis*, *A. terreus*, *A. heteromorphus* and *A. nidulans* were not

kinetically characterized. Only studies of the best growth conditions for their production

were done. Generally, the production of β-glucosidase could be induced by the inclusion

of cellulosic materials in the media. If readily metabolized carbon compounds were

available (such as glucose) the β-glucosidase production was repressed. In some cases it

has been shown that the presence of specific disaccharides and cellulose resulted in

induction of β-glucosidase enzymes (Solovyeva *et al.*, 1997).

The β-glucosidases that were derived from *Aspergillus* have molecular weights

between 40 and 340 kDa. Most of the enzymes seem to be dimers and many are

glycosylated. The pH optima are usually between 4 and 5.5, and the temperature optima

vary between 55 and 65°C. The enzymes' activities were not affected by EDTA, but in

some specific cases the presence of Mn^{2+} resulted in activation. The substrate specificities

of the enzymes varied.

The enzyme isolated from *A. oryzae* showed the most activity with laminaribiose,

then gentiobiose and finally cellobiose. This enzyme could also hydrolyze maltose,

isomaltose and nigerose. This is unusual since these sugars have α glycosidic linkages

rather than β. Compared to other β-glucosidases this enzyme was extremely glucose

tolerant (meaning that the enzyme was still active in the presence of large amounts of

glucose). Usually β-glucosidases are inhibited by glucose at low concentrations (K_i

between 1-5 mM). The K_i obtained for glucose was 1.36 M, and the K_i obtained with the

transition state analog inhibitor glucono-δ-lactone was 12.5 mM which is also an

unusually high value for a transition state analog inhibitor (Riou *et al.*, 1998). The β-

glucosidase characterized from *A. wentii* showed more expected specificities (being the

most active with cellobiose). However the analysis was not done on purified enzyme. This

enzyme was also more glucose tolerant than was expected and had a K_i for glucose of 14

mM (Sriyastava *et al.*, 1984). The enzyme isolated from *A. japonicus* preferred cellobiose

as a substrate over oNPGlc. This enzyme was inhibited by glucono-δ-lactone (K_i=0.04

mM) and especially by nojirimycin (K_i=1.8 μM) (Sanyal *et al.*, 1988). The β-glucosidase

investigated in *A. sojae* was quite different. It had the greatest activity with pNPGlc, then

pNPX (X=xylose), followed in order by sophorose, laminaribiose, cellobiose, gentiobiose,

and xylobiose (Kimura *et al.*, 1999). Additionally, this enzyme was found to have

transglycolytic activity (the substrate was the acceptor) if the substrate concentrations

were over 20 mM. The *A. sojae* enzyme was also reacted with xylooligosaccharides

between 2 and 6 xylose units in length. The K_m was found to decrease with increased

length and indicated that there are several additional sugar subsites on the enzyme

(Kimura and Tajima, 1999). Finally, several β-glucosidases were isolated from *A.*

fumigatus. The first was observed in 1973, and preferred substrates with β1-2 or β1-3

glycosidic linkages (Rudick and Elbein, 1973). The second enzyme that was isolated by

the same group preferred pNPGlc, followed by cellobiose and finally gentiobiose (Rudick

and Elbein. 1975). A third β-glucosidase had low activity with both pNPGlc and cellobiose (Ximenes *et al.*, 1996).

1.3.1.1. β-Glucosidases isolated from *Aspergillus niger*

β-Glucosidases from *Aspergillus niger* have been investigated because these organisms are known as the most efficient producers of β-glucosidase among micro-organisms of interest to industry (Birk *et al.*, 1997). The enzymes are usually not very specific for substrate. They are stable at higher temperatures, and active at pHs lower than those from other fungal sources. Another advantage is that *Aspergillus niger* is considered nontoxic and can, therefore, be used in the food industry (Birk *et al.*, 1997). The fungus secretes the enzyme into the media and the enzyme can then be isolated. In some cases the media, containing all secreted enzymes, is lyophilized and that mixture is sold to researchers interested in cellulolytic enzymes. In other cases, the enzymes are purified from the media and sold as a liquid preparation of β-glucosidase (although other enzymes are also present). There are several of these mixtures available from *A. niger* strains, although the exact strains used are proprietary. Table 1.2 summarizes the *Aspergillus niger* β-glucosidases that have been investigated.

Ueda Kagakukogyo sells Cellulosin (a cellulase mixture produced by *Aspergillus niger*) in Japan. A 240 kDa protein was isolated from this mixture by Watanabe *et al.* (Watanabe *et al.*, 1992). The enzyme is a dimer with a pH optimum of 4.5 and an optimum temperature of 50°C (Watanabe *et al.*, 1992). The enzyme was found to tolerate a wide variety of aglycones in glycosidic bonds with glucose, provided that the glucose has a β-D-glucopyranose configuration. The K_m values reported for pNPGlc, cellobiose,

cellotriose, cellotetraose and cellopentaose were 0.9, 2.3, 1.1, 1.1, and 1 mM,
respectively. This enzyme could also catalyze transglucosylation to another cellobiose to
form a trisaccharide at very high concentrations of cellobiose (400 mM was used). The
principal linkages observed in this transglucosylation reaction were β1-6 but β1-2, β1-3,
and β1-4 were also observed. The main product of the transglucosylation reaction was 4-
O-β-gentiobiosylglucose (Watanabe *et al.*, 1992).

Table 1.2 A summary of the β-glucosidases isolated from *Aspergillus niger.*

A. niger strain (if known)	Native molecular weight (kDa)	pH optima	Substrate specificity	Temp. optima	Reference
Cellulosin powder	240-dimer	4.5	β-D-glucopyran		Watanabe *et al.*, 1992
Novozyme 188	118-140 dimer??	4-4.5	Cellooligo-saccharides		Grous *et al.*,1985, Dekker, 1986, McCleary & Harrington 1988, Himmel *et al.*,1993, Unno *et al.*,1993
USDB 0827	230	4.6	Lamina-ribiose		Hoh *et al.*, 1992
USDB 0828	230	4.6	Lamina-ribiose		Hoh *et al.*, 1992
IBT-90	200	4.8	Cellooligo-saccharides		Galas & Romanowska, 1997
322	64 (SDS)	5.5	Cellobiose	50	Peshin & Mathur, 1999
NIAB 280	330-homotrimer	4.6-5.3	N/A		Rashid &Siddiqui, 1999
CCRC 31494	100-homodimer	5.0	pNPGlc	55	Yan & Lin, 1997
CCRC 31494	360-homotrimer	4.5	cellobiose	60	Yan *et al.*, 1998

The most popular liquid preparation of β-glucosidase from *Aspergillus niger* is

Novozyme 188. It is a product of Novo Nordisk Bioindustry Ltd. Denmark. The β-

glucosidase(s) in the Novozyme mixture were inhibited by glucose (K_i = 2.87 mM) (Grous

et al., 1985). Dekker provided more information on the kinetics of the Novozyme mixture.

The observed pH optimum was 4.5, and the temperature optimum was between 60 and

70°C. The K_m values of pNPGlc and cellobiose were 1.03 and 5.63 mM, respectively and

the V_{max} values were 3.76 and 33.74 µmol/mg/min, respectively. Glucose and glucono-δ-

lactone were again found to be inhibitory. Xylose was also an inhibitor. Additionally,

Dekker noted that there was substrate inhibition at pNPGlc concentrations greater than 4

mM while cellobiose concentrations up to 10 mM were not inhibitory (Dekker, 1986). The

Novozyme 188 β-glucosidase preparation was analyzed again by McCleary and

Harrington (1988). The β-glucosidase was purified and found to have a subunit molecular

weight of 118 kDa. The enzyme had a pH optimum of 4 and a temperature optimum

between 70 and 75°C. The enzyme was reacted with a variety of substrates and was most

reactive with cellotriose, then cellobiose, cellotetraose, cellopentaose, gentiobiose,

pNPGlc, sophorose and oNPGlc. The enzyme activity with gentiobiose was approximately

half of the activity observed with cellobiose. The activity reported with oNPGlc was only

about 1% of the total activity that was observed with cellobiose. The K_m values that were

determined were different than the values that had been previously reported. Cellobiose

had a K_m of 1.89 mM and pNPGlc a K_m of 0.8 mM (McCleary, and Harrington, 1988). The

Novozyme 188 β-glucosidase preparation was investigated again by Himmel *et al.* (1993).

The experiments showed that there were β-glucosidase isoforms present in the Novozyme

33

mixture. The enzymes showed similar molecular weights, approximately 120 kDa, but differed in the amounts of glycosylation. This resulted in differences in the K_m values obtained for pNPGlc. One of the isoforms had a K_m value of 0.47 mM while the other had a K_m value of 0.36 mM. The N-terminal sequence was obtained for both enzyme forms and found to be the same, and when all carbohydrate was removed the enzymes had the same molecular weights (Himmel et al., 1993). Due to the data discrepancies, the Novozyme 188 β-glucosidase preparation was again studied by Unno et al (1993). They found that the enzyme had a molecular weight of 137 kDa and consisted of about 12.5% covalently attached carbohydrate. They also determined that the enzyme hydrolyzed its substrates with a retaining mechanism, and was specific for β-glycosidic linkages. Cellobiose was the best substrate, and the relative rate of hydrolysis decreased with increasing cellooligosaccharide length. Additionally, laminaribiose was shown to be a good substrate (92% relative rate of cellobiose), while pNPGlc was a relatively poor substrate (58% relative rate). The enzyme showed no activity with lactose, pNPGalactoside, pNPXyloside, or disaccharides with α linkages (Unno et al., 1993). Inhibition studies showed that glucono-δ-1,5-lactone was a good inhibitor (K_i=0.01 mM) and it is considered to be a transition state analogue that would bind at subsite -1 of the β-glucosidase. Other sugars were also assayed for their ability to inhibit the enzyme. Galactose and lactose were tested and both caused some inhibition to occur (Yazaki et al., 1996). The same group determined that the enzyme had 6 glucose binding subsites (Yazaki et al., 1997).

β-Glucosidases have also been purified from different strains of *Aspergillus niger* and characterized. Two extracellular enzymes were isolated from 2 strains, USDB 0827 and USDB 0828 (Hoh *et al.*, 1992). Both enzymes had molecular weights of approximately 230 kDa, had a pH optimum of 4.6 and a temperature optimum of 65°C. The enzymes had similar activities with cellooligosaccharide substrates and pNPGlc. In both cases the enzymes were more active with cellobiose than with cellotriose. The greatest amount of activity was seen with laminaribiose. Glucose was a competitive inhibitor. The trends were the same for the two enzymes studied, but the actual values obtained differed. The USDB 0828 enzyme was always more active than the USDB 0827 enzyme. A β-glucosidase from *Aspergillus niger* IBT-90 has also been isolated and characterized (Galas and Romanowska, 1997). The enzyme was found to be a dimer with a molecular weight of 200 kDa, and contained about 33% carbohydrates. It has a pH optimum of 4.8 and a temperature optimum of 65°C. The enzyme was unaffected by incubation with EDTA and divalent metals, indicating that they are not required for enzyme activity. N-Bromosuccinimide had a strong inhibitory effect suggesting that Trp residues are necessary for enzyme activity. Glucono-δ-lactone and glucose were competitive inhibitors, with K_i values of 0.4 and 22 mM, respectively. The enzyme had the highest V_{max} with cellobiose, followed in order by gentiobiose, cellotriose, amygdalin, cellotetraose, cellopentaose, and pNPGlc. K_m values, 1.5 and 0.71 mM, were reported for cellobiose and pNPGlc, respectively. Another β-glucosidase was isolated from *Aspergillus niger* strain 322. This time the enzyme was smaller, 64 kDa, had an optimal pH of 5.5 and an optimal temperature of 50°C. This enzyme was quite different from those previously

mentioned. This enzyme showed very high affinity for cellobiose (K_m=0.1 mM) and was inhibited by the presence of divalent metal ions (Peshin, and Mathur, 1999). Additional β-glucosidases have been isolated by other groups. Rashid and Siddiqui (1997) isolated a 300 kDa trimeric protein from *A. niger* strain NIAB280, and Yan and Lin isolated an extremely glucose-tolerant 100 kDa dimer from *A. niger* CCRC 31494 (Yan and Lin, 1997). This latter enzyme is quite different from previously reported enzymes since the K_i for glucose is 543 mM. This group also isolated a second β-glucosidase from the same *Aspergillus niger* strain that was a 360 kDa trimer, and that was competitively inhibited by glucose (K_i=5.7 mM) (Yan *et al.*, 1998).

1.3.1.2 An *Aspergillus niger* β-glucosidase from glycosyl hydrolase family 3

A β-glucosidase was identified in *Aspergillus niger* and cloned (Dan *et al.*, 2000). The molecular weight was obtained from a native gel and was approximately 160 kDa. Upon denaturation, SDS-PAGE indicated a molecular weight of 120 kDa. The enzyme was deglycosylated using N-glycosidase-F. SDS-PAGE of the deglycosylated protein indicated that approximately 20 kDa of the enzyme's mass was due to N-glycosylation. The quaternary structure of the enzyme was not discussed despite the contradictory molecular weights reported.

The gene was cloned and sequenced (GenBank accession # AJ132386). The genomic DNA encoding the β-glucosidase consisted of 7 exons that were interrupted by 6 introns. The upstream sequence coded for a leader sequence that was not part of the

mature protein (19 amino acids). The open reading frame encoded for a polypeptide that was predicted to be about 92 kDa. The sequence showed similarity to family 3 glycosyl hydrolases. All of the family 3 enzymes proceed with retention of configuration and the *A. niger* β-glucosidase was not an exception. Studies with pNPGlc using ^{1}H NMR provided evidence that the β-anomer of glucose was formed initially. The catalytic nucleophile was identified as Asp-261 (Dan *et al.*, 2000). This enzyme is the subject of this thesis.

1.4 Structure of a family 3 glycosyl hydrolase

To my knowledge only one family 3 glycosyl hydrolase structure has been solved, although the family now boasts over 100 members. The enzyme is a β-D-glucan exohydrolase isolated from *Hordeum vulgare* (barley) (Varghese *et al.*, 1999). Its preferred substrate for hydrolysis is laminarin, a β(1-3) glucan. It also hydrolyzes β(1→3,1→4) (this nomenclature indicates that the non-reducing end glucose is joined by a β1-3 glycosidic bond to a second glucose, which is joined to a third by a β1-4 glycosidic bond) and β(1→3,1→6) oligosaccharides and aryl-glucosides. Regardless of the substrate, a glucose is released from the non-reducing end of the sugar. The cleavage of the glycosidic bond occurs with retention of configuration (Varghese *et al.*, 1999).

The structure of β-D-glucan exohydrolase was determined by x-ray crystallography to a resolution of 2.2Å. The binding site of the enzyme was located by the presence of a bound glucose in a pocket of the structure (Varghese *et al.*, 1999). The protein is an elongated structure that contains 2 domains, which are connected by a 16 residue helix. An $(\alpha/\beta)_8$ barrel (TIM barrel) makes up one of the domains and the other domain is a $(\alpha/\beta)_6$ sheet. The C-terminus forms a long antiparallel loop. The bound glucose was in a

37

pocket that lies between the TIM barrel domain and the $(\alpha/\beta)_6$ sheet domain, just above

the pseudo-eightfold axis of the barrel. The pocket is 13Å deep (Varghese *et al.*, 1999).

The amino acids located in the pocket are situated as follows and are shown in Figure 1.4.

Trp-286 is at the top of the pocket, and Trp-434 is at the bottom. The left side of the

pocket is lined by a triple glycine loop (Gly-56,-57,-58) and Arg-291 is on the right side.

There is a hydrophobic patch at the top of the pocket (M250, F144, M316, and L54) and at

the bottom there is a basic patch (K206, H207, R158). Lys-206 and His-207 are joined by

a *cis* peptide bond. The catalytic nucleophile, identified by labeling with conduritol B

epoxide, is Asp-285, which is found at the top surface of the pocket and is thought to be in

close contact to the anomeric carbon (<3Å). It was suggested that the catalytic acid/base is

Glu-491. It is located at the bottom of the pocket, approximately 3.5Å from the anomeric

carbon. However, it is also possible that Glu-220 is the catalytic acid/base group since it is

also close by, and is highly conserved in the family 3 enzymes. Glu-491 is not conserved

in some of the family 3 enzymes. The positioning of these residues indicates that the β-D-

glucan exohydrolase is not a member of the "4/7 superfamily" that have their catalytic

nucleophile on strand 7, and their catalytic acid/base on strand 4.

The hydroxyl groups of the bound glucose are all involved in multiple hydrogen

bonding and ionic interactions. The anomeric hydroxyl forms H-bonds with Glu-491 and

Tyr-253. The hydroxyl of carbon-2 interacts with Arg-158 and Asp-285 and the C3

hydroxyl has interactions with Arg-158, Lys-206, and His-207. The hydroxyl group at C4

interacts with both Asp-95 and Lys-206 while the C6 hydroxyl interacts with Asp-95 and

a water molecule. Hydrophobic interactions also occur. The hydrophobic face of the

38

Figure 1.4. The important groups in the active site of *Hordeum vulgare* β-D-glucan exohydrolase. The catalytic nucleophile (D285) and the putative catalytic acid/base (E491 & E220) are shown in red. The glycine loop (Gly-56, -57, -58) on the left side is shown in green. The hydrophobic patch (L54, F144, M250, M316) is indicated in orange, and Trp 286, 430 and 434 are shown in yellow. The basic residues lining the bottom of the pocket (K206, H207 and R158) are shown in magenta. Tyr-253 and Asp-95, which are discussed in the text are shown in cyan. The structure was obtained from the Protein Data Base (PDB#1EX1) and RasMac version 2.6 was used to produce this Figure.

39

glucose ring (C3H, C5H and C6H) interacts with Leu-54, Met-316, Phe-144, and Met-250. The C2H and C4H on the other face interacts with Trp-430 (Varghese *et al.*, 1999).

1.4.1 Comparison of *Aspergillus niger* β-glucosidase with the structure of β-D-glucan exohydrolase from barley.

The *H. vulgare* β-D-glucan exohydrolase sequence shows 37% identity to the *A. niger* β-glucosidase sequence (Figure 1.5). Many of the residues discussed above are conserved in *A. niger* β-glucosidase. The catalytic nucleophile Asp-285 is conserved in the *A. niger* β-glucosidase but Glu-491 is not. Instead a Gln is found at that position. However, Glu-220, which is also a candidate for the catalytic acid/base is conserved in the *A. niger* β-glucosidase. Both Trp-286 and Trp-434 are conserved, while Trp-430 is replaced with a Leu. From the placement of Trp-286 and Trp-434 it does not appear that they have stacking interactions in the barley enzyme. The Trp-434 has some hydrophobic interactions with the corner of the glucose ring, as mentioned above. Other residues that appear important in the barley active site that are conserved in the β-glucosidase are D95, R158, H207, M250, M316, Y253, and G56 and G58. In addition to the above residues there are several other patches of conservation: from 309-314 (*H. vulgare*), LAGLDM, from 390-394 (*H. vulgare*), VLLKN, and from 594-597 (*H. vulgare*), GYGL. These regions seem to be common in many of the family 3 proteins and are highly conserved. Additionally 2 Trp, other than those discussed above are also conserved in both the barley and the *A. niger* proteins. Trp-70 and Trp-544 (*H. vulgare*).

Figure 1.5 Alignment of the *Aspergillus niger* β-glucosidase sequence with the *Hordeum vulgare* (barley) β-D-glucan exohydrolase. Both proteins belong to the family 3 glycosyl hydrolases and there was 37% identity. The sequences were aligned using Gene Inspector™ version 1.5 by Textco. The scoring matrix used was Blosum 55. The alignment parameters used were a k-tuple (word) size of 1, a maximum gap length of 5, a gap penalty of 3, the number of top diagonals to use was 5, a gap creation of 12 and a gap extension of 10. Red indicates the most conserved residues and pink indicates sequence gaps. (Unpublished results Seidle and Huber, 2001).

1.5. OBJECTIVES AND SEQUENCE OF STUDY

The original goal of this study was to thoroughly characterize the β-glucosidase from *Aspergillus niger* in a commercial "cellulase", to clone the gene that encodes for this β-glucosidase and to study the residues at the active site by site specific mutagenesis. As discussed, the properties of quite a large number of β-glucosidases from *A. niger* have been described in the literature and there is confusing evidence about their properties. It was felt that at least one of the β-glucosidases should be studied in detail, and the gene producing it should be isolated. In addition, preliminary work in this laboratory had shown that the β-glucosidase in a commercial (Sigma) cellulase had interesting kinetic properties. To this end, I worked with the commercial cellulase from *Aspergillus niger* from Sigma, and identified the β-glucosidase in the enzyme mixture. Only one β-glucosidase was apparent in the cellulase mixture. A suitable method of isolation was developed. Studies were done to explain the unique pH kinetic effects with various substrates that this enzyme has. The temperature optimum and stability, substrate specificities, inhibition kinetics, glucose subsite affinities, analyses of substrate breakdown and product formation, identification of unique sugar products were all carried out. The enzyme was found to be a glycoprotein, and I analyzed the carbohydrate moiety attached to the glycoprotein.

When I had finished most of these studies and before I had started to isolate the gene I learned that Dr. Shoseyov's group at the Hebrew University in Jerusalem was working with a β-glucosidase from *A. niger* strain CMI CC 324262. This group had already isolated and sequenced the gene. The work has now been published (Dan *et al.* 2000). A collaboration was undertaken. The gene and the enzyme sequence were obtained.

I incorporated the gene for the β-glucosidase into an appropriate plasmid and expressed the recombinant β-glucosidase in *Pichia pastoris*. I fully characterized the recombinant enzyme to compare it with the β-glucosidase isolated from the Sigma mixture. Sequencing of a peptide of the Sigma enzyme indicated that the two enzymes were identical. Other physical studies confirmed this. Detailed studies with the recombinant β-glucosidase showed that it had the same unique kinetic character, the same pH and temperature optima, and the same substrate and inhibition specificities. Substrate breakdown and product formation studies provided further evidence of the identity of the two enzymes.

After this was done, I carried out substitution studies. Because the study was a collaboration it was determined that I would study some of the conserved Trp residues (and one Leu) to define their roles in binding and activity. Trp are often found at sugar binding sites and there were a number of conserved Trp. The characterization of the β-glucosidase from the Sigma source suggested the presence of at least 5 glucose subsites. The sequence of the β-glucosidase showed that the β-glucosidase is a member of the family 3 glycosyl hydrolase family 3. Comparison of the sequences of family 3 members suggested which conserved residues should be investigated. Site specific mutagenesis was completed for Trp 21, 49, 262, 430, 551, 806, and 813 as well as Leu 426. Of the 21 substituted enzymes that were produced, 17 were characterized in detail. Attempts were also made to explain the properties of the substituted *A. niger* β-glucosidases on the basis of the only known family 3 structure (*H. vulgare* β-glucan exohydrolase).

2. MATERIALS AND EQUIPMENT

2.1 Biochemical Reagents

2-Nitrophenyl β-D-glucose (oNPGlc) and 4-nitrophenyl β-D-glucose (pNPGlc) were obtained from Fluka or Sigma. D-Arabitol, D-glucose, D-gluconic acid lactone, maltose, cellobiose, (β,β)-trehalose (iso-trehalose), sophorose, laminaribose, gentiobiose, D-allose, D-sorbitol, 4-nitrophenyl-β-D-gentiobiose, cellopentaose, cyanogen bromide, ATP, NADP⁺, EDTA, L-histidine, L-alanine, biotin, hexamethyldisilazane, dimethylformamide (HPLC grade), trimethylchlorosilazane, magnesium chloride, Coomassie Blue G-250, triethanolamine, iodoacetic acid, bromophenol blue and lyticase were all purchased from Sigma as was the crude cellulase (*Aspergillus niger*) powder. Cellotriose and cellotetraose were purchased from either Sigma or Calbiochem. The 3,4-dinitrophenyl-cellobiose (DNPC) and glucopyranosyl-D-(1-6)-glucopyranosyl-β-D-(1-4)-glucose (4-O-gentiobiosyl glucose) were prepared in the laboratory. Yeast extract, tryptone, peptone and yeast nitrogen base were purchased from Difco. The hexokinase/glucose-6-phosphate dehydrogenase enzyme mixture (hexokinase, 340 units/mL; and glucose-6-phosphate dehydrogenase, 170 units/mL) used in the coupled assay was from Boehringer Mannheim. Hydrochloric acid, glacial acetic acid, methanol, ethanol, butanol, ethyl acetate, 2-propanol, propanol, sulfuric acid, formic acid, sodium chloride, sodium hydroxide, sodium phosphate, acetic anhydride, magnesium sulfate, potassium phosphate, phosphoric acid and glycerol were all obtained from BDH. Manganese sulfate and ethylene glycol were purchased from J.T. Baker Chemical Co. Sodium acetate, sodium carbonate, calcium sulfate, zinc sulfate, sea sand and sucrose

were purchased from Fisher. Ultra pure TRIS was obtained from ICN, and Tricine was from Aldrich. Ethidium bromide, Ultra Pure SDS, Ultra Pure agarose, and Ultra Pure low melting point agarose were from Gibco BRL. The restriction enzymes and the lambda molecular weight marker were obtained from Amersham/Pharmacia Biotech, New England Biolabs or Gibco BRL. Zeocin was provided by Invitrogen. The D_2O used for NMR was purchased from Cambridge Isotope Laboratories. Helium, air, hydrogen and nitrogen gases were from Praxair.

2.2 Purification and Chromatography Equipment

Gel filtration chromatography was used for protein purification. A Sephacryl S-300 (Pharmacia) column (26 mm x 91 cm) was used for the first step, followed by Superose 6 and 12 columns joined in series (bed volume of 24 mL each) (Pharmacia). Protein was concentrated using Filtron macro- and micro-centrifugal devices. Sugars were purified using Sephadex G10 and G25 (with 10 mm x 152 cm, and 20 mm x 60 cm columns, respectively) from Pharmacia Biotech. Silica from ICN and sea sand from Fisher was used for flash chromatography of sugars. The column was 10 mm x 20 cm. A small cotton plug was placed in the bottom and a small amount of sea sand covered the cotton (the sand was about 3 mm thick). Silica made up the rest of the column. The column was utilized for additional purification of sugars for NMR identification.

Thin layer chromatography was done on Whatman K6F Silica Gel 60A plates. A flat chamber was used to expose the plates to solvent and the plates were developed horizontally. The chambers used were either 5 cm x 5 cm or 10 cm x 10 cm.

Gas liquid chromatography (GLC) was done with a 5710A Hewlett Packard Gas

Chromatograph with a Hewlett Packard 7123A recorder using an OV-101 packed column

from Chromatographic Specialties Inc. The stainless steel column had a 0.6 cm diameter

and was 200 cm long. Detection was by flame ionization. The gas pressures utilized were

50 psi for helium, 60 psi for compressed air and 40 psi for hydrogen. The injection

temperature was 350°C. The detector temperature was 300°C. The same temperature

program was used for the majority of the samples. The temperature was kept at 200°C for

4 min. The temperature was then increased at a rate 8°C/min to 310°C. It was held at

310°C for 32 min. A second temperature program was used for some samples. The

temperature was kept at 200°C for 4 min. The temperature was then increased at a rate of

8°C/min to 320°C. It was held at 320°C for 22 min. Attenuation was adjusted depending

on peak size.

2.3 NMR Equipment

NMR spectra were recorded using a Bruker AMX-500 spectrophotometer

operating at 500.139Mhz for ^{1}H and at 125.71MHz for ^{13}C.

2.4 Fluorescence Studies

All fluorescence studies were done with a Jobin Yvon Spex: Fluorolog-3, using

Datamax for Windows software. The slit width for excitation was 5 nm and was 1 nm for

emission. Emission was recorded in counts per second. Emission was measured at 1 nm/s

from 300 to 450 nm. The excitation wavelengths used were 275, 285 and 295 nm.

Excitation spectra were collected at an emission wavelength of 339 nm (the highest point

of the emission spectra usually occurred at this wavelength). The slit width for excitation

was 1 nm and was 5 nm for emission. Excitation was recorded as millivolts (mv). The

spectra were measured at 1 nm/s from 200 to 300 nm.

2.5 Light Scattering

Light scattering experiments were completed on both the wild type β-glucosidase

from Sigma and on the unsubstituted recombinant enzyme. A Dyna Pro MSTC Dynamic

Light Scattering Device with the Dynals software from Protein Solutions, Inc. was used.

2.6 Enzyme Assays

All spectrophotometric assays were carried out using a Shimadzu UV2101 PC

spectrophotometer with a CPS260 temperature controlled multi-cell changer set at 25°C.

The spectrophotometer was interfaced to a Packard Bell 386SX-II computer. The

software used was UV-2101/3101PC software (version 3) and UV-2101/3101 PC

optional kinetics software (version 2).

2.7 *Escherichia coli* strains

Plasmids obtained from Dr. Shoseyov (Hebrew University of Jerusalem in Israel)

containing the β-glucosidase gene were transformed into *E. coli* XL2 blue ultracompetent

cells. These cells were purchased from Stratagene. The genotype of XL2 Blue cells is

recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacI^qz_M15 Tn10 (Tet^r)

Amy Cam^r]. Mutants of the β-glucosidase gene were transformed into E. coli XL1 Blue

Supercompetent cells. The genotype is *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1*

lac[F' proAB lacI^qzΔM15 Tn10 (Tet^r)].

2.8 *Pichia pastoris* strains

P. pastoris is a methylotrophic yeast. This means that it can use methanol as a

carbon source. There are two genes that allow the metabolism of methanol, *AOX1* and

AOX2. These genes code for alcohol oxidases that are required for the initial step of methanol metabolism (in addition to molecular oxygen), the oxidation of methanol to formaldehyde. Since the AOX enzymes have a low affinity for oxygen, a large amount of these enzymes is produced. Therefore, the promoter used to regulate the expression of the AOX enzymes has been utilized to express the heterologous gene expression in *P. pastoris*. The *AOX1* gene is responsible for the majority of the alcohol oxidase activity. Therefore strains that express *AOX1* grow very well in the presence of methanol and are referred to as Mut⁺ (methanol utilization plus). Strains where only the *AOX2* gene is expressed grow more slowly and are referred to as Mut⁺ (methanol utilization suppressed) (Invitrogen, 1997). The *P. pastoris* strains used for expression of the β-glucosidase gene were X-33, KM71 and GS115. The X-33 stain was usually used. It is a wild type strain of *P. pastoris* and therefore is Mut⁺. In some cases KM71 was used. Its genotype is *arg4 his4 aox1::ARG4*. This indicates that it has a mutation in the histidinol dehydrogenase gene that prevents it from synthesizing His, and a mutation in the argininosuccinate lyase gene that prevents it from synthesizing Arg. However, it also contains a wild type *Arg* gene inserted into the *AOX1* gene so that KM71 is a Mut⁺, Arg⁺, His⁻ strain. The strain is Mut⁺ because the insertion of the *Arg* gene into *AOX1* causes disruption of the gene so that only the *AOX2* gene is functional. The GS115 strain also has a mutation in the histidinol dehydrogenase gene, and so cannot synthesize His. However, it contains no mutations to *AOX1*, and is, therefore, a Mut⁺ strain. The GS115 strain was utilized initially, but was found to have lower transformation efficiencies than the X-33 or KM71 strains and thus was not used for expression of β-glucosidase mutants. All *P. pastoris*

strains were obtained from Invitrogen. The strain GS115/His⁻Mut⁺ Albumin was used as a control for secreted protein expression.

2.9 Plasmids

The β-glucosidase gene was provided by Dr Shoseyov and his group at The Kennedy Leigh Center for Horticultural Research and The Warburg Center for Agricultural Biotechnology. The Faculty of Agriculture. The Hebrew University of Jerusalem, Israel. The gene was obtained in three different plasmids: pET3d, pBS2SK, and pPic9K. The gene was removed from the pPic9K plasmid and placed into pPICZAα (Figure 2.1). The pPICZAα plasmid contained several features that made over expression in *P. pastoris* easier. The plasmid contained an α-factor signal sequence, AOX1 promoter region, a polyhistidine tag, and the Zeocin resistance gene (*Sh ble* ORF). The signal sequence caused the protein to be secreted to the media, the AOX1 promoter allowed the gene expression to be induced by the presence of methanol, and the polyhistidine tag could be used for purification. Zeocin is an antibiotic that is structurally related to bleomycin/phleomycin. It shows strong toxicity against both bacteria and fungi (including yeast). The yeast will only grow in the presence of Zeocin if they contain the resistance gene.

2.10 Primers

2.10.1 Sequencing Primers

Primers used for sequencing the wild type and mutant β-glucosidase genes are shown in Table 2.1. All primers had phosphorylated 5' ends. The sequencing primers from Invitrogen were designed from the pPICZAα plasmid sequence surrounding the
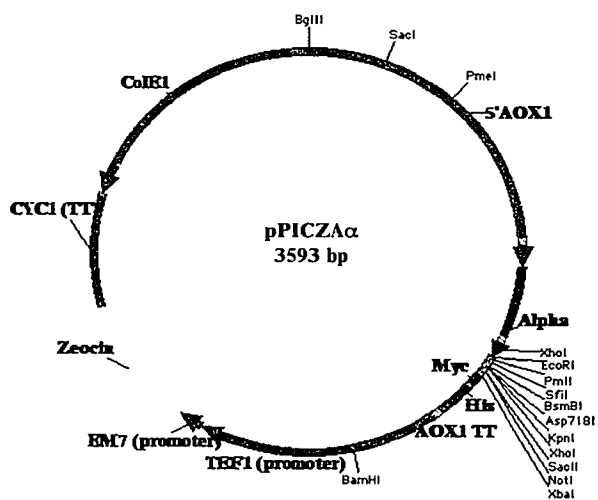
Figure 2.1. Plasmid map of pPICZAα. The multiple cloning site restriction enzyme cut sites are indicated in addition to several other important restriction sites. Other important features are also indicated. AOX refers to the alcohol oxidase gene, Alpha refers to the α–factor signal sequence, Myc refers to the myc epitope tag, His refers to the polyhistidine tag made up of 6 His residues, TT refers to transcription termination region, and ColE1 is a pUC-derived replication origin. Primer sites are indicated by stripes.

multiple cloning site. Both the 5'AOX1 and the α-factor primers were designed to

provide sequence starting at the 5' end of the β-glucosidase gene. The 3'AOX primer was

designed to provide sequence starting at the 3' end of the gene. Due to the size of the β-

glucosidase gene, additional primers were required for complete sequencing. The 5'-BG-1

primer was designed for sequencing from the interior of the β-glucosidase gene in the 5'-

3' direction. The 3'-BG-1 primer was designed for sequencing from the interior of the β-

glucosidase gene in the 3'-5' direction. The primers were designed to allow maximal

overlap so that sequence errors could be easily detected and confirmed.

Table 2.1: Primers used for sequencing the β-glucosidase genes. The primers were either supplied by Invitrogen or designed from the β-glucosidase sequence and produced by UCDNA Services. All of the sequencing primers had a Tm of 62°C.

Primer	Sequence
5'AOX1ᵃ	5'-GACTGGTTCCAATTGACAAGC-3'
3'AOX1ᵃ	5'-GCAAATGGCATTCTGACATCC-3'
β-Factorᵃ	5'-TACTATTGCCAGCATTGCTGC-3'
3'-BG-1ᵇ	5'-CTCGTTTGAGATGGCCTGCT-3'
5'-BG-1ᵇ	5'-TGACCGTGGATGCGACAATG-3'

ᵃsupplied by Invitrogen
ᵇproduced by UCDNA services

2.10.2 Mutagenesis Primers

Mutations to the β-glucosidase gene were produced using both specific and

degenerate primers. Substituted enzymes were named by placing the letter abbreviation

for the amino acid to be substituted followed by its place in the amino acid sequence

followed by the letter abbreviation for the replaced amino acid. Therefore, W49F-β-glucosidase indicates that the Trp at position 49 in the β-glucosidase amino acid sequence was replaced with a Phe. Specific primers were used to produce W49F/A, W262F/A, W430F/A, and W551F/A (See Table 2.2). Degenerate primers were used to produce W21L/H/Y, W262C/P, L426G/V/A, W430L, W806F/A and W813C (See Table 2.3). The primers were designed to have optimal T_m values and GC content according to the Stratagene Quik Change Site-Directed Mutagenesis kit manual (See Section 3.13). Primers were named according to the mutation desired and the direction of the primer (e.g. W49F-f is for a substitution of the Trp at position 49 by Phe and the primer is in the forward direction).

2.11 Growth Media

2.11.1 Growth Media for *Escherichia coli*

E. coli strains containing the pPicZAα-plasmid were grown in low salt LB (Luria-Bertani media) in order to allow selection using the antibiotic Zeocin. Zeocin is a broad spectrum antibiotic whose exact mechanism is unknown. It is thought to have a mechanism similar to bleomycin and phleomycin because of structural similarities and because its effects are prevented by the *Sh ble* resistance protein, which also prevents the other drugs from binding DNA. Zeocin is isolated as a copper-chelated glycopeptide. The copper-chelated form is inactive. However, when Zeocin enters the cell the copper is removed and this activates the Zeocin. The activated form of the antiobiotic then binds and cleaves DNA resulting in cell death (for more information refer to Berdy (1980)).

Table 2.2: Primers designed to produce substitutions to Phe or Ala at specific amino acid positions in the β-glucosidase sequence. The altered codons are indicated in bold and underlined. The primer melting temperature (T_m) is in °C. Primer names indicate the amino acid identity, position, new amino acid identity and direction (f-forward, r-reverse).

Primer	Sequence	T_m
W49F-f	5'-CCAACACATAGTTCCAATTC**GAA**TCCAGTTCCTGTGGTC-3'	76
W49F-r	5'-GACCACAGGAACTGGA**TTC**GAATTGGAACTATGTGT-3'	76
W49A-f	5'-CCAACACATAGTTCCAATTCC**GGC**TCCAGTTCCTGTGGTC-3'	74.4
W49A-r	5'-GACCACAGGAACTGGA**GCC**GAATTGGAACTATGTGT-3'	74.4
W262F-f	5'-GCATGGTGAGCAGC**GAA**ATCACTCATGACAAAGCCC-3'	78.5
W262F-r	5'-GGGCTTTGTCATGAGTGAT**TTC**GCTGCTCACCATGC-3'	78.5
W262A-f	5'-GCATGGTGAGCAGC**GGC**ATCACTCATGACAAAGCCC-3'	77.2
W262A-r	5'-GGGCTTTGTCATGAGTGAT**GCC**GCTGCTCACCATGC-3'	77.2
W430F-f	5'-GCACTACCACTTCC**GAA**GCCCATCGCCAATGTTC-3'	77.4
W430F-r	5'-GAACATTGGCGATGGGC**TTC**GGAAGTGGTACTGC-3'	77.4
W430A-f	5'-GCACTACCACTTCC**GGC**GCCCATCGCCAATGTTC-3'	75.8
W430A-r	5'-GAACATTGGCGATGGGC**GCC**GGAAGTGGTACTGC-3'	75.8
W551F-f	5'-CCGGGCAAACCACC**GAA**GAGGATAGCGGTAAC-3'	78.5
W551F-r	5'-GTTACCGCTATCCTC**TTC**GGTGGTTTGCCCGG-3'	78.5
W551A-f	5'-CCGGGCAAACCACC**GGC**GAGGATAGCGGTAAC-3'	78.1
W551A-r	5'-GTTACCGCTATCCTC**GCC**GGTGGTTTGCCCGG-3'	78.1

Table 2.3: Degenerate primers designed to produce substitutions for specific amino acid positions in the β-glucosidase sequence[a]. The mutated codons are indicated in bold and underlined. The primer melting temperature (T_m) is in °C. Primer names indicate the amino acid identity, position, and direction (f-forward, r-reverse). Degenerate nucleotides were indicated by K, M, R, Y, H, and N (see [b] for symbol description).

Primer	Sequence	Tm
W21-f	5'-CAATGGCCAGGGGCAC**YNC**GGGCAGGCATAC-3'	79.7
W21-r	5'-GTATGCCTGCGC**GNR**GTCGCCCTGGCCATTG-3'	79.7
W262-f	5'-GGCTTTGTCATGAGTGAT**YNC**GCTGCTCACCATGCTGG-3'	78.8
W262-r	5'-CCAGCATGGTGAGCAGC**GNR**ATCACTCATGACAAAGCC-3'	78.8
L426-f	5'-GGATGCGACAATGGAACA**KKK**GCGATGGGCTGGG-3'	76.8
L426-r	5'-CCCAGCCCATCGC**MMM**TGTTCCATTGTCGCATCC-3'	76.8
W430-f	5'-GCACTACCACTTCC**NHA**GCCCATCGCCAATGTTC-3'	78.1
W430-r	5'-GAACATTGGCGATGGGC**THN**GGAAGTGGTACTGC-3'	78.1
W806-f	5'-CGCCGTGACCTTGCAAAC**YNC**AATGTTGACAAGCAGG-3'	79
W806-r	5'-CCTGCTTGTCAACATT**GNR**GTTTGCAAGGTCACGGCG-3'	79
W813-f	5'-GGAATGTTGAGAAGCAGGAC**YNC**GAGATTAGCTCGTATCC-3'	79.1
W813-r	5'-GGATACGAGCTAATCTC**GNR**GTCCTGCTTCTCAACATTCC-3'	79.1

[a]Primers were produced by Sigma Genosys.
[b]Symbols used in degenerate primers: K=G/T, M=A/C, R=A/G, Y=C/T, H=A/T, N=A/C/T/G.

Zeocin activation is inhibited by high salt concentrations. Therefore, it is necessary to reduce the amount of salt that is usually used to make LB media. Low salt LB contains: 10 g Tryptone, 5 g NaCl, 5 g Yeast extract in 1 L ddH$_2$O. The pH was adjusted to 7.5 with NaOH and then autoclaved on liquid cycle (30 min.). For plate preparation, 15 g of

agar was added prior to autoclaving. After the media was cooled. Zeocin was added to 25

μg/mL final concentration. Plates were stored in the dark at 4°C. For transformation into

E. coli XL2-Blue or XL1-Blue cells, NZY⁺ broth was used. NZY⁺ broth contains: 10 g

NZ amine (casein hydrolysate enzymatic (Sambrook and Russell, 2001)), 5 g yeast

extract. 5 g NaCl. This solution was brought to pH 7.5 with NaOH and then autoclaved

(liquid cycle, 30 min.). After autoclaving, 12.5 mL of filter sterilized $MgCl_2$ (1 M), and

$MgSO_4$ (1 M), and 10 mL of filter sterilized 2 M glucose were added. The total volume

was 1 L.

2.11.2 Growth Media for *Pichia pastoris*

P. pastoris strains were grown in several different media. YPD (yeast extract

peptone dextrose medium) was used for plates and growth of strains for glycerol stocks.

YPD contains 1% (w/v) yeast extract, 2% (w/v) peptone and 2% (w/v) dextrose

(glucose). For plates, 20 g/L agar was added prior to autoclaving. Filter-sterilized

dextrose (20% w/v) was added to the autoclaved media. Zeocin was added in some cases

to give a final concentration of 100 μg/mL. Liquid media that does not contain Zeocin

was stored at room temperature while plates and media containing Zeocin were stored at

4°C. Media and plates containing Zeocin were kept in the dark and had a maximum shelf

life of 2 weeks.

After *P. pastoris* strains were transformed (see section 3.14.2 for details) with the

β-glucosidase gene plasmid, they were grown on YPDS (yeast extract peptone dextrose

medium with sorbitol) plates with Zeocin. 182.2 g sorbitol, 20 g peptone and 20 g of agar

in 900 mL ddH_2O. This medium contained 10 g yeast extract, 182.2 g sorbitol, 20 g

peptone and 20 g of agar in 900 mL ddH$_2$O. This was autoclaved and then 100 mL of

20% (w/v) dextrose was added. After cooling, Zeocin was added to a concentration of

100 μg/mL. Once the plates were poured they were stored at 4°C in the dark.

To confirm the Mut phenotype of yeast transformants (X-33 and GS115 strains

should be Mut$^+$, and KM71 should be Mut^s) they were grown on MDH (minimal dextrose

medium with histidine) and MMH (minimal methanol medium with histidine) plates.

MDH was made by adding 100 mL of filter sterilized 13.4% (w/v) YNB (yeast nitrogen

base with ammonium sulfate without amino acids, 2 mL of filter sterilized 0.02% (w/v)

biotin, 100 mL of filter sterilized 20% dextrose (w/v) and 10 mL of filter sterilized 0.4%

(w/v) His to 800 mL autoclaved water containing 15 g of agar. MMH was produced in

the same manner but 100 mL of 5% (v/v) methanol was added instead of dextrose. Both

types of plates were stored at 4°C.

The media used for expression of the β-glucosidase gene were BMGY (buffered

glycerol complex medium with yeast nitrogen base) and BMMY (buffered methanol

complex medium with yeast nitrogen base). These media were made (1 L) by dissolving

10 g yeast extract and 20 g of peptone in 700 mL of water and autoclaving. When this

mixture was cooled, 100 mL of 1 M potassium phosphate buffer (pH 6), 100 mL 13.4%

(w/v) YNB, 2 mL 0.02% (w/v) biotin and 100 mL of 10% (v/v) glycerol (BMGY) or

5%(v/v) methanol (BMMY) were added. Both of these media were stored at 4°C.

The media used for growing up yeast strains for the purpose of isolating genomic

DNA was MGYH (Minimal glycerol medium +/- histidine with yeast nitrogen base).

This media was made up of 1.34% (w/v) YNB, 1% (v/v) glycerol, 4x10-5% (w/v) biotin

and +/- 0.004% histidine.

2.12 Electrophoresis

2.12.1 SDS/native PAGE

The Pharmacia Phastgel™ system was used for both native and SDS-PAGE. Phast gels with a 10-15% acrylamide gradient (G10-15) were used for native PAGE and gels with an 8-25% acrylamide gradient (G8-25) were used for SDS-PAGE. Both types of gels had 2% crosslinking. Buffer strips were made for both types of gels. SDS-PAGE buffer strips contained 0.2 M Tricine, 0.2 M Tris, and 0.55% (w/v) SDS (analytical grade) at a pH of 8.1. Native buffer strips contained 0.88 M L-alanine and 0.25 M Tris (pH 8.8). For both types of buffer strips, 2% (w/v) agarose was added prior to use. Sample Buffer was made up of 10 mM Tris/HCl, and 1 mM EDTA (pH 8). SDS was added to a final concentration of 2.5% (w/v), mercaptoethanol to 5% (w/v) and bromophenol blue was added to 0.01% (w/v). Coomassie blue tablets and the Phast gel silver stain kit were purchased from Amersham/Pharmacia Biotech. Protein size was determined by comparison to "high molecular weight markers" for native or SDS PAGE. These markers were also purchased from Amersham/Pharmacia Biotech.

2.12.2 Agarose Gel Electrophoresis

Agarose gels contained 0.8% (w/v) agarose which was melted in sterile 1X TAE buffer (50X TAE (1 L) contains 242 g Tris, 16.8 g EDTA, and 57.1 mL acetic acid (pH 8)). After cooling, ethidium bromide (0.0025 µg/mL, final concentration) was added to the gel. Samples were prepared by mixing with TE buffer (10 mM Tris-HCl (pH 7.5) and 1 mM EDTA) and loading buffer (0.25% (w/v) bromophenol blue and 40% (w/v)

sucrose). The sample was generally 10% of the total volume. A Hind III digested lambda ladder was used to determine DNA size.

2.13 Analysis of Carbohydrate Chains

The sugars removed from the β-glucosidase protein were analyzed using equipment supplied by Glyko. All reagents were supplied in the FACE® N-Linked Oligosaccharide Profiling Kit or in the FACE™ N-Linked Oligosaccharide Sequencing Kit. The FACE™ N-Linked Oligosaccharide Profiling Kit contained all the gels, buffers, labeling agents, dyes, standards and controls necessary to perform the profiling reactions. The FACE™ N-Linked Oligosaccharide Sequencing Kit contained all of the enzymes, gels, buffers, standards and controls necessary for the oligosaccharide sequencing.

59

3. METHODS

3.1 Purification of the β-glucosidase from the Sigma cellulase mixture

3.1.1 Ammonium sulphate precipitation

The cellulase powder obtained from Sigma (catalogue number C-1184) was

dissolved in 150 mM acetate buffer (pH 4.5) at a concentration of 0.1 g/mL. Sigma ultra

pure $(NH_4)_2SO_4$ was added to the cellulase solution to a concentration of 1.7 M (40%

saturation) at 4°C. The resulting suspension was stirred for 30 min. at 4°C and then

centrifuged at 25 000 x g for 30 min. The supernatant was then brought to 2.7 M (60%

saturation) $(NH_4)_2SO_4$ at 4°C and again stirred for 30 min. Centrifugation was repeated.

The pellet was redissolved in a small volume of 150 mM sodium acetate buffer. The

protein solution was kept at 4°C until it could be applied to a Sephacryl S-300 column.

3.1.2 Size exclusion chromatography

3.1.2.1 Sephacryl S-300

The protein solution was applied to a Sephacryl S-300 column (see Section 2.2).

Sodium acetate buffer (pH 4.5, 150 mM) was the eluent (2 mL/min). Fractions (4.5 mL)

were assayed for β-glucosidase activity (see Section 3.6.2.1). Active fractions were

collected and concentrated to 500 μL. SDS-PAGE was used to assess the purity of the

concentrated protein samples.

3.1.2.2 FPLC

If necessary, enzyme solutions obtained from the Sephacryl column were

concentrated and loaded (500 μL) onto an FPLC with a Superose 6 and a Superose 12

column connected in series. Elution was with 150 mM sodium acetate buffer (pH 4.5) at a

speed of 0.25 mL/min. Fractions (1 mL) were assayed for activity. The active fractions were analyzed by SDS-PAGE and those containing only the β−glucosidase band were pooled and concentrated.

3.2 Polyacrylamide gel electrophoresis

Samples were prepared for SDS-PAGE by mixing equal parts of 2X sample buffer and protein and boiling for 3 to 5 min. These mixtures were applied to the gel in 1 μL aliquots. Protein samples for native PAGE were applied directly to the gel in 1 μL aliquots. SDS-PAGE gels were run for 105 volt hours and native PAGE gels for 300 volt hours. Gels were stained with Coomassie Blue or the Phast™ gel silver stain kit from Pharmacia.

3.3 Determination of β-glucosidase size

3.3.1 SDS and native PAGE

The molecular weight of the β-glucosidase was estimated by several methods. Initially SDS PAGE and native PAGE were utilized and the results indicated that it was likely that the protein was a dimer. However, it was found that the β-glucosidase is a glycoprotein and so other methods were also utilized.

3.3.2 Size exclusion chromatography

Blue Dextran (2 000 000 Da), thyroglobulin (669 000 Da), ferritin (440 000 Da), catalase (232 000 Da) and adolase (158 000Da) were applied to Superose 6 and 12 columns connected in series and eluted on the FPLC. ß-Glucosidase was applied separately and eluted. The absorbance was monitored at 280 nm. Elution of the standards

was repeated several times to ensure that the FPLC column results were reproducible. The logarithms of the molecular weights vs. the elution volumes were plotted.

3.3.3 Dynamic light scattering

Dynamic light scattering (DLS) is a non-invasive approach to protein characterization. The sample is irradiated with a mono-chromatic laser and the scattered light is collected at a photon detector placed at a known angle relative to the incident beam. The strength of the scattering signal is dependent on the molecular weight, density, shape, concentration, viscosity and refractive index properties of the medium, as well as the magnitude of the particle-particle interactions and the solvent-particle interactions (http://protein-solutions.com).

The DLS data produces an intensity correlation curve, which is an indirect measure of the particle's diffusion coefficient and this information is used to calculate the radius of a molecule and then the molecular weight. The radius is determined by:

$$D = \frac{kT}{f} = \frac{kT}{6\pi\eta r} \qquad (1)$$

D is the diffusion coefficient, k is Boltzmann's constant, T is the temperature, and f is the frictional coefficient for a compact sphere in a viscous medium. The frictional coefficient is determined from $6\pi\eta r$, where η is the solvent viscosity and r is the radius. The radius that is determined from the diffusional properties indicates the apparent size of a dynamic, hydrated particle. Hence the term hydrodynamic radius. The parameters utilized by the Dynals program to calculate the hydrodynamic radii for the β-glucosidase are given in Appendix I.

3.4 Determination of protein concentration

The concentration of enzyme was determined using the modified Bradford method (Kruger, 1994). The Bradford assay is dependent on the binding of Coomassie Blue dye to protein which allows spectrophotometric measurements to be taken at 595 nm. The assay is somewhat dependent on amino acid composition since the dye binds most readily to Arg residues (Kruger, 1994). Bovine γ-globulin was used as a standard since the dye binding capacity of this protein is closest to the mean of a variety of proteins that have been assayed (Kruger, 1994). Additionally, the percentage of Arg in γ-globulin (4.3%) is similar to the percentage of Arg in β-glucosidase (4.1%). The protein assay reagent was composed of 10 mg of Coomassie Brilliant Blue, 2.5 mL of 95% ethanol, 5 mL of 85% phosphoric acid and 50 mL of water. This was added (4 mL) to 6 samples of 1 mg/mL bovine γ-globulin (0, 0.04, 0.08, 0.12, 0.16, 0.20 mL) that were made up to 0.2 mL with the addition of water. Tubes were incubated for 5 min. at 20°C. Absorbance at 595 nm was recorded for each tube and plotted to produce the standard curve. The enzyme was treated in the same manner. The concentration of protein in the unknown sample was determined by comparing the A_{595} values on the standard curve. All samples were measured in duplicate.

3.5 Determination of extinction coefficient

Once the protein concentration was determined the extinction coefficient (280 nm) was calculated. The extinction coefficient was determined from Beer's Law. The calculated extinction coefficient was used to determine protein concentration from the absorbance (280 nm) in subsequent work.

3.6 Enzyme assays

3.6.1 3,4-Dinitrophenyl cellobiose (DNPC)

Assays with DNPC were carried out in 150 mM sodium acetate buffer at pH 5 and 25°C. The absorbance of dinitrophenol at 420 nm could be followed directly as a function of time since the pK_a value of the dinitrophenol is relatively low (5.4).

3.6.2 pNPGlc/oNPGlc

Assays with oNPGlc and pNPGlc were carried out in sodium acetate (150 mM) buffer at 25°C. The assays were started by the addition enzyme. The final volume of the assay was 0.33 mL. The reactions (done in duplicate) were allowed to incubate for 3 to 15 minutes. Zero time blanks were also obtained. The reactions were stopped by the addition of 0.67 mL of 1 M sodium bicarbonate (Na_2CO_3). This raised the pH to 10 and inactivated the enzyme. Nitrophenols are fully ionized at pH 10 and absorb at 400 nm. Absorbance versus time of incubation was plotted in order to ensure that the increase in product concentration was linear. Extinction coefficients (pH 10) of 4.65 $cm^{-1}mM^{-1}$ for oNP and 13.01 $cm^{-1}mM^{-1}$ for pNP were used. All assays were completed at least twice and in many cases assays were done in triplicate. A small amount of error was noted (approximately 5-15%).

3.6.3 Determination of V_{max}, K_m and V_{max} (trans)

The kinetics of the reactions were studied using Eadie-Hofstee plots. Unusual Eadie-Hofstee plots were obtained with the ß-glucosidase. Figure 3.1 illustrates a typical Eadie-Hofstee plot obtained upon reaction of β-glucosidase with pNPGlc. The results
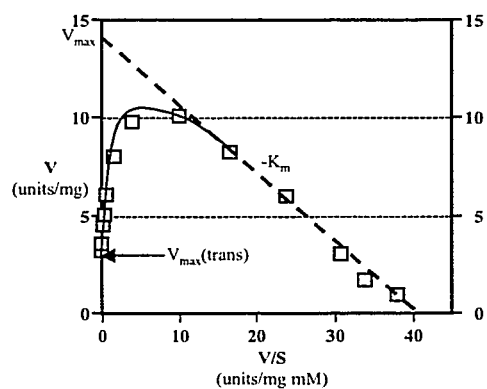
Figure 3.1. An example of an Eadie-Hofstee plot obtained with pNPGlc
and the β-glucosidase purified from the Sigma cellulase mixture.
The dashed line represents the hydrolytic reaction. The V_{max} of the
hydrolytic reaction is the y-intercept of the extrapolated dashed
line, while the K_m is the negative slope of the dashed line. The
transglycosylation reaction is illustrated by the solid line. The V_{max}
of the transglycosylation reaction (V_{max}(trans)) is the y-intercept
indicated by the arrow.

section will show that the unique profiles were obtained because both hydrolytic and

transglucosidic reactions occur. The kinetic constants were determined as follows. Straight

lines were fitted to the data obtained at low substrate concentrations as shown in Figure

3.1. These lines correspond to the hydrolytic portion of the reaction. The lines were

extrapolated to the Y-axis. The Y-intercepts of these straight lines are equal to the V_{max}

values while the negative slopes of these straight lines are the K_m values of the hydrolytic

part of the reaction. The V_{max} values for the transglycolysis reaction (called V_{max} (trans)

values) were determined by statistically extrapolating the curves to infinite substrate

concentration.

   3.6.3.1 Statistical analysis of the V_{max} (trans) values using a SAS

(statistical analysis system) non-linear regression program

   The V_{max} (trans) values were calculated using a non-linear least squares statistical

analysis program, SAS (www.sas.com). A model equation was derived from the

mechanism (to be presented later – Figure 4.20) using King-Altman patterns and the

Cleland method.

$$V_o = \frac{a[S]+b[S]^2}{c + d[S] + e[S]^2} \qquad (2)$$

The equation has this form (with the substrate concentration squared) because the

substrate is also a transglucosidic acceptor and thus affects the rates in two ways. The

substrate usually binds with less affinity to the acceptor site and thus the effects of the

terms with substrate concentration squared are only seen at high substrate concentration.

The SAS program solved for the 5 unknowns (a, b, c, d, and e) using the Marquardt iterative process. When the equation is solved for the five unknowns, the V_{max}, K_m and V_{max} (trans) can be obtained by calculating a/d, c/d and b/e, respectively. This program was, however, only used to obtain the V_{max} (trans) value. The V_{max} and K_m values were more accurately obtained from the straight line (see Figure 3.1).

3.6.4 Rapid qualitative assays for β-glucosidase activity

Fractions obtained from purification columns were qualitatively assayed to determine the presence of β-glucosidase. Aliquots (15 μL) of the fractions were added to 15 μL of 100 mM pNPGlc in sodium acetate buffer (150 mM, pH 4.5). This was incubated at room temperature for 10 to 30 min. A small amount (30 μL) of 1 M Na_2CO_3 was then added. Mixtures with enzyme became yellow.

3.6.5. pH profiles with nitrophenyl glucosides as substrates

Profiles of the enzyme activity at various pH values were done with pNPGlc and oNPGlc. The assays were similar to those in section 3.6.2 but were done in a buffer made up of sodium acetate (150 mM) and sodium phosphate (150 mM). The effects of pH from 3 to 7 were investigated.

3.6.6. Temperature effects

The thermal stability was investigated by heating the enzyme to 65°C, removing aliquots at timed intervals and assaying them for activity with 50 mM pNPGlc. Assays were done as described in Section 3.6.2. V_{max} values were also obtained at 5°C temperature intervals from 25° to 85°C.

The effects of freezing the enzyme were also investigated. An aliquot of enzyme was placed in liquid nitrogen. The frozen enzyme was then lyophilized and later was suspended in buffer (sodium acetate buffer, 150mM, pH 4.5). The protein was assayed for activity as detailed in Section 3.6.2.

3.6.7 Non-aryl glucosyl substrates

Assays of glucose production with cellobiose, cellotriose, cellotetraose, cellopentaose, isotrehalose, sophorose, laminaribiose, gentiobiose and 4-O-gentiobiosyl-glucose (see Appendix II for structures) were done at 25°C and pH 4.5 (150 mM sodium acetate buffer). Because of the expense of the substrates and reagents, the assays were done very carefully in a small volume (50 μL). The mixtures were allowed to react for 5 min. at 25°C at which time 40 μL aliquots of the assays were removed and transferred to tubes containing 50 μL of a mixture of hexokinase (4.2 units) and glucose-6-phosphate dehydrogenase (2.1 units) in a 1.54 mL total volume. Magnesium chloride (13.3 mM), ATP (1.26 mM) and NADP⁺ (1.87 mM) were also present. The high concentrations of ATP and NADP⁺ were used to ensure that they would not limit the extent of the reaction. The buffer was 33.3 mM triethanolamine (pH 8.3). At this pH the β−glucosidase is inactive and therefore the assay measured only the glucose that was produced before the coupled enzymes were added. The absorbance was measured at 340 nm for a total of 5 min. This allowed all of the glucose produced by the β-glucosidase to react with the hexokinase/glucose-6-phosphate dehydrogenase mixture and the absorbance plateaued (Figure 3.2). The total amount of glucose produced by the enzyme was calculated from the plateau region of the absorbance using an extinction coefficient of NADPH (6220
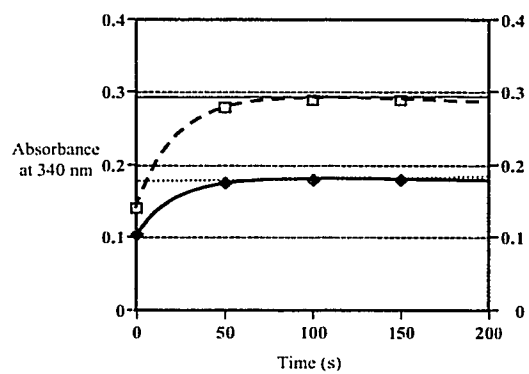
Figure 3.2. Typical spectrophotometric results from the hexokinase/glucose-6-phosphate dehydrogenase assay coupled to the β-glucosidase assay. These results were obtained from the β-glucosidase purified from the Sigma cellulase mixture reacted with cellobiose at concentrations of 10 mM (□) and 4 mM (◆). The absorbance readings obtained from the graph plateaus are illustrated.

$M^{-1}cm^{-1}$). Zero time blanks, obtained by the same method but without the addition of enzyme, were subtracted.

3.6.7.1 pH studies with cellobiose as a substrate

The effects of pH (pH 3-7) on the enzyme were studied with cellobiose. Assays were completed as in Section 3.6.7 except that the buffer was made up of acetate (150 mM) and sodium phosphate (150 mM).

3.6.8 Inhibitor studies

The inhibitory effects (K_i values) of a variety of sugars on the enzyme reaction with pNPGlc were determined in 150 mM sodium acetate buffer at pH 4.5 and 25°C. The sugar concentrations studied were varied depending on the inhibitory effect found in preliminary studies. The sugars tested were glucose, mannose, allose, galactose, xylose, gluconic acid δ-lactone (also referred to as gluconolactone), cellobiose, gentiobiose, maltose, and cellotriose (see Appendix II for structures). Since pNPGlc was the substrate used, the product that was measured was nitrophenol. This allowed the inhibition effects of cellobiose, gentiobiose and cellotriose (all good substrates) to be determined because they have different products than pNPGlc. Therefore, if hydrolysis of these sugars occurred it would not interfere with the inhibition assay. Additionally, the reaction was done quickly so that only a very small amount of the inhibitor could be reacted. Assays were carried out as in Section 3.6.2 but with the addition of the inhibitor. Only low pNPGlc concentrations were investigated so that the K_i values could be obtained with the straight line section of the plots. The reaction mechanism for β-glucosidase in the presence of a competitive

inhibitor is shown in Figure 3.3. The Figure shows that inhibitors also act as acceptors but the K_i values obtained by equation (3) below are for when inhibitors bind to free enzyme.

$$K_i = \frac{[I]}{\left\{ \dfrac{{}^{App}K_m}{{}^{App}V_{max}} \Big/ \dfrac{K_m}{V_{max}} \right\} - 1} \qquad (3)$$

The K_i values are competitive inhibition constants. The equation accounts for changes in the reaction rate that occur in the presence of a competitive inhibitor that also acts as an acceptor (Deschavanne et al., 1978; Huber et al., 1983). The V_{max} and K_m values (Equation (3)) are obtained in the absence of an inhibitor while the apparent V_{max} and K_m values (${}^{App}V_{max}$ and ${}^{App}K_m$) are obtained in the presence of an inhibitor (I) that also acts as an acceptor (Figure 3.3).

3.6.9 Acceptor studies

As already stated, most sugars and alcohols were found to act as acceptors and bind to the glucosyl form of the enzyme. Different alcohols were tested for their ability to affect the rate when acting as acceptors. The assays were done at low pNPGlc concentrations to minimize the effects of the substrate itself acting as an acceptor, but were otherwise carried out as above. Assays were done with ethanol (1 M), glycerol (1 M), propanol (1 M) and methanol (0.1 M, 0.5 M, 1 M). These are rather high concentrations but the K_i" values are large and so effects are only seen at such high values. Also, this was only done for a qualitative indication of the acceptor action. The high concentrations of the alcohols did not appear to cause denaturation.

E•Glc•I

Glc-I

k_4

K_i''

I

E•I

K_i

E

k_1

k_{-1}

E• Glc-NP

k_2

E-Glc

I

NP

k_3

Glc

H$_2$O
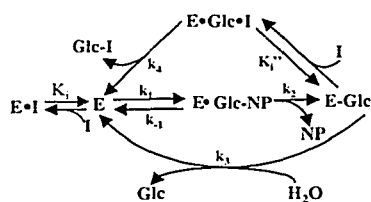
Figure 3.3. Inhibition of the β-glucosidase reaction. E represents the free
enzyme, I is an inhibitor that can also act as an acceptor, Glc-
NP is pNPGlc, NP is the released nitrophenol, and Glc is
glucose. Rate constants are shown in addition to the
dissociation constants for the inhibitor/acceptor (K_i and K_i'').
The K_i value of a particular inhibitor is determined by its
binding affinity for the free enzyme and is a measurement of
competitive inhibtion . If the inhibitor binds to the glucosyl
form of the enzyme and acts as an acceptor, this is measured
by the K_i'' constant. The K_i'' value was not determined. The
solid dots represent an interaction of some kind with the
enzyme.

Acceptor activity was also followed by TLC. Reactions were done in 150 mM sodium acetate buffer (pH 4.5). The assay solutions contained 20 mM pNPGlc. Enzyme was added at 0 time. Aliquots (1 μL) were spotted onto TLC plates at appropriate time intervals. The TLC plates were then processed according to Section 3.7.1. Various sugars and alcohols were tested. The appearance of novel products on the TLC plate suggested acceptor activity.

3.6.10 Metals and EDTA

The effect of metals (Mg^{2+}, Ca^{2+}, Mn^{2+} and Zn^{2+} (10 mM)) and EDTA (10 mM) was investigated. Reactions were completed in the same manner as normal (Section 3.6.2) except that metals or EDTA were added. In all cases, the effects of these additions were determined from changes in the resulting Eadie-Hofstee plots.

3.7 Analysis of substrate breakdown

3.7.1 Thin layer chromatography

Reactions were usually carried out in 150 mM sodium acetate buffer at pH 4.5 (25°C) but in some cases the reactions were done at pH 7.0. Substrates and enzyme were mixed at zero time and 1 μL aliquots were applied to TLC plates at timed intervals. The reaction was stopped by this action. Plates were placed into a developing chamber and the solvent was allowed to travel to 0.5 cm from the top. The developing solvent was a mixture of n-butanol, ethyl acetate, 2-propanol, acetic acid and water in the ratio of 1:3:2:1:1. Sometimes the plates were dried and run a second time. Products containing nitrophenyl groups were detected by illuminating the dried plate with UV light. Products

containing sugars were detected by dipping the dried plates into 2% (v/v) concentrated

sulfuric acid in methanol and then charring the plates at 450°C. Except for the cases where

unknown sugars were produced, the identities of the sugars were determined by using

known standards.

3.7.1.1 Densitometer analysis

In some cases, the amounts of the various sugars present on TLC plates were

quantitated by using LKB Broma Ultroscan XL Enhanced Laser Densitometer scans of

reproductions of the charred plates on transparencies. The densitometer was programmed

for an x-width of 4, a y-step of 4, and a peak width of 2. The baseline was the average of

the 16 lowest data points.

3.7.2 NMR analysis

NMR was used to analyze the enzyme reaction over time and to determine the

identity of the products resulting from the enzyme reaction. The reactions were done in

150 mM acetate buffer at pH 4.5 and 25°C using a 10 mm probe.

3.7.2.1 Cellobiose

Assays were done to continuously follow product formation with 80 mM cellobiose

in the presence of 2 µg of enzyme. The final volume was 2.5 mL. The reaction was done

in 150 mM sodium acetate buffer (pH 4.5) at 25°C. Spectra were collected every 4 hrs, and

analyzed for the presence of cellobiose, gentiobiose, gentiobiosyl-glucose and glucose.

3.7.2.2 pNPGlc

Assays utilizing NMR were done to continuously follow product formation with 75

mM pNPGlc in the presence of 4 µg of enzyme. The final volume was 2.5 mL. Spectra

were collected every 4 hrs and analyzed for the presence of pNPGlc, pNPGe and glucose. The reaction was done in sodium acetate buffer (150 mM) at pH 4.5 and 25°C, using a 10 mm probe.

3.8.1 Gas-liquid chromatography

Gas-liquid chromatography was utilized to follow the reactions of the β-glucosidase. Standards were run to determine where the expected substrates and products (cellobiose, glucose, gentiobiose and gentiobiosyl- glucose) would elute from the column and to construct standard plots. In some cases these experiments were done with pNPGlc and gentiobiose as the substrates. All of the β-glucosidase reactions were carried out in 150 mM sodium acetate buffer at pH 4.5, 25°C.

The GLC analyses were done as follows. Aliquots of (100 μL) of the reaction were removed after timed intervals that depended on the concentration of the substrate. The zero time aliquot was taken before enzyme was added. The aliquots were added to tubes that were immersed in liquid nitrogen and so the aliquots froze instantly. An internal standard (100 μL) of 2 mM inositol was also added to the reaction tube and it also froze instantly. The tubes were then lyophilized overnight. The samples were then silylated by adding dimethylformamide (500 μL) as a solvent and then hexamethyldisilazane (HMDS) (250 μL) and trimethylchlorosilane (TMCS) (100 μL). After vortexing, the tubes were incubated for 30 min. at 70°C. The tubes were vortexed again and the incubation was repeated. The silylated sugars became concentrated in a small layer of liquid at the top of the microfuge tube. A small sample was injected into the injection port of the gas liquid chromatograph instrument. The peaks that were obtained were displayed on a chart

recorder. The areas under the peaks of the products were measured by weighing. The ratios of the weights of the product peaks to the weight of the internal standard peaks were determined. Weighing was done with a Mettler Micro-Gramatic balance that was accurately calibrated for very small weights. The ratios were then compared to standard ratio curves to determine the quantities of the sugars produced.

3.9 Identification of unknown products

3.9.1 Production of large amounts of unknowns

The enzyme produced unknown transglucosidic reaction products. These unknown sugars were produced and purified in large amounts for NMR analysis. The unknown compounds produced by the reaction of ß-glucosidase on cellobiose were obtained by incubating β-glucosidase (1 mL, 0.5 mg/mL) with 24 mL of 100 mM cellobiose for 210 min. at 25 °C. The enzyme was reacted in a similar manner with DNPC but the reaction was done with a smaller amount of substrate (5 mL of 5 mM) and therefore, less enzyme was used (0.6 mL of 0.43 mg/mL). Also the reaction was incubated at 25°C for only 5 min. A reaction with pNPGlc was also done. In this case it was done with 25 mL of 100 mM pNPGlc and 3 mL of enzyme (0.5 mg/mL) and the reaction was allowed to proceed for 120 min. at 25°C. These times and concentrations were chosen because maximum levels of the unknowns were present at those times and concentrations for each substrate. The reactions were stopped by freezing the reacting vessels in liquid N_2. The solutions were lyophilized.

3.9.2 Purification of unknown sugars

The concentrated mixtures were applied to a Sephadex G10 column (1 m x 3 cm diameter) and eluted with deionized H_2O. Sephadex G-10 was used for sugar separation since it is highly cross-linked and separates molecules below a molecular weight of 1000 g/mol. The matrix also interacts weakly with the carbohydrates at low salt concentrations and this allowed for the separation of molecules of the same size but with different structures. Fractions were assayed for sugars by TLC and fractions containing pure individual sugars were pooled. The pooled fractions were lyophilized. TLC was used to estimate the purity of the unknowns.

3.9.3 NMR

To identify the unknown sugars using NMR, the purified lyophilized unknowns were dissolved in D_2O (500 µL). Internal standard, DSS (2 uL), was then added. The spectra were acquired at 25°C using a 5 mm probe. HMQC (heteronuclear multiple quantum coherence) and TOCSY(Total Correlation Spectroscopy) spectra were utilized to determine the chemical shift values shown in Appendix III.

3.9.4 TLC

TLC was also used to identify some of the unknowns. This was done when a standard was available for comparison. Generally, the more common sugars, glucose, cellobiose, and gentiobiose were identified by comparison to R_f values of standards. However, NMR was used to verify the identities of glucose, cellobiose and gentiobiose as described in Section 3.9.3.

3.10 Identification of *Aspergillus niger* β–glucosidase

After the initial studies on the β–glucosidase obtained from Sigma had been accomplished I was informed that another group, Dr. Shoseyov's lab in Israel, was working on an *Aspergillus niger* β–glucosidase and had already cloned the gene. Therefore, it became important to determine if the two β–glucosidases were the same. Dr. Shoseyov generously sent me a lyophilized sample of his protein and the gene sequence. Once I was sure that the two proteins were the same, Dr. Shoseyov also provided the cloned gene in three plasmids (pPet3d, pPic9K, and pBS₂SK) as well as two strains of *Pichia pastoris* containing plasmids for the genes of two substituted enzymes, W262L and W139L-β-glucosidase.

3.10.1 Cyanogen bromide cleavage

Cyanogen bromide cleavage was used to cleave the protein into small peptides. The sulfhydryl and disulfide groups of the protein were initially carboxymethylated. Solutions of the enzyme were dialyzed overnight in 0.1 M Tris-HCl (pH 8.6) with β-mercaptoethanol (at a concentration equivalent to 2-3x that of the total cysteine concentration). After dialysis, urea (8 M final concentration), 2 M Tris-HCl (pH 8.5) (0.4 M final concentration), and 5% (w/v) EDTA (0.2% final concentration) were added. The solution was then degassed with nitrogen and incubated at room temperature for 4 hours. Following the incubation the solution was wrapped in aluminum foil to protect it from light and a 20-fold molar excess of iodoacetic acid was added (20X the number of sulfhydryl groups in the solution including those from β-mercaptoethanol). After sitting in the dark for 15 minutes, 1 mL of 14 M β-mercaptoethanol was added. This was done in

order to react the excess iodoacetic acid. This solution was dialyzed against 0.2 M

ammonium bicarbonate (2 changes). The carboxymethylated protein was frozen in liquid

nitrogen and lyophilized. The lyophilized protein was dissolved in 70% formic acid and 5

crystals of CNBr were added. This mixture was incubated overnight at room temperature.

A rotovap was used to remove the formic acid and cyanogen bromide. The protein was

resuspended in 30% acetic acid and dried. This step was repeated 3X.

3.10.2 Peptide sequencing

The peptides were analyzed by Dr. D. McKay, University of Calgary. He separated

the peptides by gel electrophoresis. One of the peptides was chosen and sequencing was

done for 12 degradation cycles on a PE-Applied Biosystems 491 Protein Sequencer.

3.10.3 Other methods for determining the identity of the two β–glucosidases

The lyophilized protein provided by Dr. Shoseyov was also tested kinetically to determine

if it was similar to the β-glucosidase purified from the Sigma cellulase mixture. Both

enzymes were assayed with pNPGlc at pH 4.5 at concentrations from 0.1 to 50 mM.

Assay conditions were the same as those described in section 3.6.2. The two enzymes

were also analyzed by SDS- and native PAGE.

3.11 Transformation into *E. coli*

The 3 plasmids containing the β-glucosidase gene were transformed into *E. coli*

Ultracompetent XL2-Blue cells according to the protocol provided by Stratagene. Initially

the cells were thawed and 75 μL were added to sterile tubes with 1.42 M β-

mercaptoethanol (1.55 μL). After incubating on ice for 10 min., 1-2 μL of plasmid DNA

was added and the mixture was kept on ice for an additional 30 min. The mixture of cells,

plasmid DNA and β-mercaptoethanol was then heat shocked at 42°C for 30 s. Following heat shock, the mixture was returned to ice for 2 min. Preheated NZY⁺ broth was then added (0.9 mL, 42°C) and the cells were incubated with shaking at 37°C for 1 hr. After this, the cells were plated on LB agar plates containing a suitable antibiotic (if the plasmid contained the gene for Zeocin resistance, the cells were plated on low salt LB with 25 µg/mL Zeocin; if the plasmid contained the gene for ampicillin resistance, the cells were plated on LB plates with 100 µg/mL ampicillin).

3.11.1 Purification of plasmid DNA

All plasmid DNA was purified using the Qiagen Mini-, Midi-, or Maxi-plasmid purification kits (depending upon the amount of DNA required). The Qiagen plasmid purification protocols are based on a modified alkaline lysis procedure, followed by binding of DNA to Qiagen anion-exchange resin. RNA, proteins, dyes and low-molecular weight impurities were removed by a medium-salt wash and plasmid DNA was eluted in a high-salt buffer. The DNA was then concentrated and desalted by isopropanol precipitation. The expected yield for high-copy plasmids for Mini, Midi, and Maxi kits were 10-20 µg, 75-100 µg, and 300-500 µg, respectively.

3.11.2 Introduction of the β-glucosidase gene into the pPicZAα plasmid

The β-glucosidase gene was transferred from the pPic9K plasmid to the pPicZAα plasmid because the latter had several beneficial features. A description of these features is given in Table 3.1. The sequence of the pPic9k plasmid containing the β-glucosidase gene was analyzed and found to have two unique restriction sites on either end of the gene that were also unique to the multiple cloning site of the pPicZAα plasmid (Figure 3.4).

Table 3.1. Features of the pPicZAα plasmid.

Feature	Benefit
5' AOX1	A fragment of the alcohol oxidase gene containing the AOX1 promoter that allows methanol-inducible, high-level expression in Pichia. It also targets the plasmid for integration into the AOX1 locus in P. pastoris
Native Saccharomyces cerevisiae α-factor secretion signal	Allows for secretion of proteins to the media by P. pastoris.
Multiple cloning site (10 unique sites)	Allows insertion of gene of interest.
C-terminal myc epitope tag	Permits detection of the fusion protein by the Anti-myc antibody.
C-terminal polyhistidine tag	Encodes 6 His that form a metal-binding site for affinity purification.
AOX1 Transcription termination (TT)	Native transcription termination and polyadenylation signal from AOX1 gene that permits efficient 3' mRNA processing-including polyadenylation for increased stability.
TEF1 promoter	Transcription elongation factor 1 gene promoter from S. cerevisiae that drives expression of the Sh ble gene in Pichia.
Sh ble gene (Streptoalloteichus hindustanus bleomycin (ble) gene	Zeocin resistance gene.
EM7 promoter	Synthetic prokaryotic constitutive promoter that drives Sh ble gene expression in E. coli.
CYC1 transcription termination region	3' end of the S. cerevisiae CYC1 gene that allows efficient 3' mRNA processing of the Sh ble gene for increased stability.
CoLE1 origin	Allows replication and maintenance of the plasmid in E. coli.

Figure 3.4. Construction of pPICZAα:BGL plasmid from pPic9k and
pPICZAα. The β-glucosidase gene is represented by BGL. The
pPic9K plasmid and pPICZAα plasmid share some common
features: ColE1-a pUC-derived origin of replication, the alpha
factor signal sequence for secretion (α), and the 3'*AOX1* TT-the
AOX1 transcription termination region. Both also contained a
multiple cloning site (MCS). pPic 9k also contains a fragment of
the 3'*AOX1* gene, the 5'*AOX1* promoter region, Ampicillin
(Amp^R) and Kanamycin (Kan^R) resistance and the histidinal lyase
gene (*His4*). The pPICZAα plasmid also has some additional
features; a region of the 5'*AOX1* gene (5'*AOX1*), a myc epitope
tag (m), a polyhistidine tag (H), a TEF1 promoter, an EM7
promoter, the *Sh ble* ORF (Zeocin resistance, Zeo^R), and the
CYC1 transcription termination region (CYC1 TT). The benefits
of the above features are described in Table 3.1.
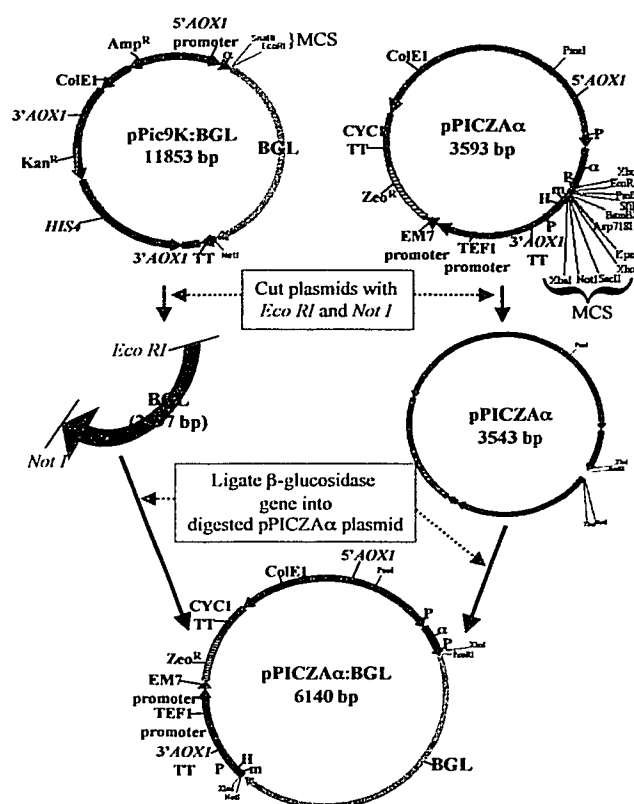
Therefore, both plasmids were cut at these unique sites (*Eco RI* and *Not I*) in a double digest. The digestions were carried out according to the manufacturer's instructions (Pharmacia Biotech). The digestion of the pPic9K plasmid produced 2 DNA fragments (a 2.5 kb fragment containing the β-glucosidase gene and a 9.3 kb fragment representing the rest of the plasmid) (Figure 3.5). The sites in the pPicZAα plasmid were in the MCS and only a small piece of DNA was removed by the digestion process. The restriction enzymes produced a large 3.54 kb plasmid fragment and a smaller 0.05 kb fragment was removed from the cloning site. The DNA fragments resulting from the digestions were visualized on a 1% low-melting point agarose gel. The β-glucosidase gene band and the pPicZAα band were excised from the gel and purified using the Wizard DNA Purification System (Promega). The gene was then ligated into the plasmid using the Rapid DNA Ligation kit (Roche). The new plasmid was loaded onto a 0.8% agarose gel to determine if the ligation was effective. This is shown in Figure 3.6.

3.11.3 Transformation into *Escherichia coli*

The new plasmid, pPicZAα:BGL (pPicZAα containing the β-glucosidase gene) was transformed into *E.coli* XL2 blue cell following the Stratagene protocol (see section 3.11). Following transformation, the cells were plated on low salt LB plates containing 25 μg/mL Zeocin and grown overnight at 37°C. Zeocin resistant colonies were then selected and grown in 2 mL of low salt LB with 25 μg/mL Zeocin overnight at 37°C with shaking (225 rpm). Plasmid was purified using the Qiagen Mini-prep plasmid purification kit and sequenced (UCDNA Services).
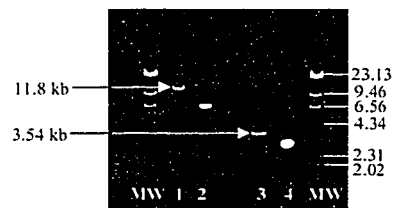
Figure 3.5 Agarose gel illustrating the migration of the parent plasmid
containing the β-glucosidase gene (pPic9k:BGL) and the
expression plasmid, pPicZAα, prior to β-glucosidase gene
insertion. Gel lanes are indicated in white. MW is the molecular
weight markers and sizes for the bands are shown
on the right. Both plasmids were digested with *Eco RI* to linearize
the plasmid and illustrate their sizes. Lanes 1 and 2 contain
pPic9k:BGL, lane 1 is the linear form of the plasmid while lane 2
shows the circular DNA. Lanes 3 and 4 contain
pPicZAα plasmid DNA in the linear and circular forms,
respectively.

85
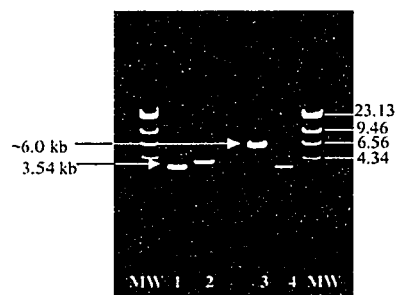
Figure 3.6. Agarose gel illustrating the migration of the pPicZAα plasmid
before, and after, the incorporation of the β-glucosidase gene.
The lanes are indicated on the bottom of the gel. MW is the
molecular weight marker. Sizes of the marker bands are shown on
the right side of the gel. Lanes 1 and 2 show the digested, and
undigested, forms of pPicZAα, respectively. Lanes 3 and 4
contained the pPicZAα plasmid with the β-glucosidase insert
(~2.5 kb) in digested, and undigested forms respectively. Both
plasmids were digested with the *Eco RI* restriction enzyme.

3.11.4 Restriction digestions

Restriction enzyme digestions of plasmid DNA were done to determine plasmid size, quantitate DNA and isolate DNA fragments. The digestions were usually done in 10-20 μL volumes using the compatible buffers provided by the manufacturers. The restriction enzyme solution usually made up 10% of the total volume (i.e. 1 μL in 10 μL total volume) and the reactions were incubated from 1 hr to 24 hrs at 37°C.

3.11.5 Agarose gel electrophoresis

Agarose gel electrophoresis was utilized to analyze all purified plasmids prior to sequencing. The agarose was melted in TAE buffer to form a 0.8% gel. A *Hind* III digested lambda ladder was used to determine plasmid size. The lambda ladder was either produced in the lab or purchased from Amersham/ Pharmacia Biotech. The advantage of the purchased marker is that it can be used to estimate the concentration of the DNA. The gels were run in TAE buffer at an appropriate voltage for the gel size. Plasmids that were not digested (1 μL) were mixed with 1 μL tracking dye and 8 μL TE buffer to a total volume of 10 μL. Plasmids that were digested (2 μL) were mixed with 1.5 μL loading buffer and 11.5 μL TE buffer to a total volume of 15 μL. Agarose gels were visualized using an Eagle III gel scanner.

3.12 Identification of important tryptophan residues

A major portion of the studies I did concerned β-glucosidases with substitutions for Trp residues. It was determined that the β-glucosidase belonged to the family 3 glycosyl hydrolases. The presence of conserved residues in this family of proteins provided information as to which Trp could be important in β-glucosidase function. Sequence

alignments were done using the Gene Inspector program by Textco Inc. The matrix used was BLOSUM 62. Additionally, analysis of the one known structure (Varghese *et al.*, 1999) of family 3 glycosyl hydrolases gave further clues as to which Trp could be important.

3.13 PCR mutagenesis

Mutagenesis of the β-glucosidase gene was accomplished using the Quik-Change™ Mutagenesis kit from Stratagene. The kit provides a rapid 4 step procedure for producing mutations (Figure 3.7). Initially, 2 synthetic oligonucleotides containing the desired mutation were synthesized (UCDNA Services). The primers are complementary to opposite strands of the plasmid and are extended during temperature cycling by *Pfu* Turbo DNA polymerase. The polymerase replicates both plasmid strands with high fidelity and without displacing the mutant primers. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. After temperature cycling the mixture is treated with *Dpn* I endonuclease which is specific for methylated and hemi-methylated DNA. This causes the digestion of the parental DNA template and results in selection for synthesized DNA containing the mutation. The nicked plasmid DNA was transformed into *E. coli* XL1-Blue supercompetent cells where the nicks in the mutated plasmid were repaired.

Primers designed to replace specific Trp were designed according to the Stratagene protocol for the Quik-Change mutagenesis kit. Both forward and reverse primers must contain the desired mutation and anneal to the same sequence on opposite strands of the

Figure 3.7. Diagram of the steps involved in the Quik Change Site
Specific Mutagenesis Kit™ obtained from Stratagene. Step 1
shows the target site for mutagenesis. Step 2 shows how the
mutagenic primers anneal and illustrates the production of
plasmid complementary to the original with the exception of the
site of mutagenesis. Step 3 shows how non-mutated plasmid
DNA is removed by the nuclease *Dpn I* which is specific for
methylated DNA (all plasmid produced in *E. coli* will be
methylated). The final step (step 4) shows how the nicks are
repaired after transforming the new DNA (containing the mutant
sequence) into *E. coli*. Adapted from the Stratagene Catalog,
1999.

Temperature
Cycling

Digestion

Transformation

Step 1: The gene of interest is present in a plasmid and the target site (■) for the mutation has been determined.

Step 2: Oligonucleotide primers (▬) containing the desired mutation (★) anneal to the denatured plasmid.

The mutagenic primers are extended by *Pfu Turbo* DNA polymerase resulting in nicked circular strands.

Step 3: The methylated, nonmutated parental DNA template is digested with *Dpn I*.

Step 4: The circular, nicked double-stranded DNA is transformed into XL1-Blue supercompetent cells. After transformation, the cells repair the nicks in the mutated plasmid.

plasmid. The primers should be 25-45 bps in length and should have a $T_m \geq 78^\circ C$. The T_m for each primer was calculated using the formula:

$$T_m = 81.5 + 0.41(\%GC) - 675/N - \% \text{ mismatch} \qquad (4)$$

where N is equal to the primer length. In addition, the mutation should be in the middle of the primer. The optimal minimum GC content is 40% and the primer should terminate in one or more C/G bases. Finally, the primers must be purified. The PCR reactions contained: 5 µL 10X reaction buffer, 25 ng double-stranded DNA template, 125 ng primer #1 and 125 ng primer #2, 1 µL dNTP mix (10 mM), ddH$_2$O to a final volume of 50 µL and 1 µL Pfu Turbo™ DNA polymerase (2.5 U/µL). The reaction was overlaid with sterile mineral oil (30 µL). PCR was done according to the protocol shown in Table 3.2 (PCR #1) except for when substitutions for Trp 813 and Trp 262 were made. Those primers contained a high GC content which caused them to preferentially form primer dimers and therefore they did not bind to the DNA template. Initially, the annealing temperature was adjusted to try to decrease the chance of primers annealing to each other. Several different temperatures were chosen (Table 3.2, PCR #2). This did not alleviate the problem and so the amount of template DNA was adjusted (50, 100, and 150 ng were added). The PCR #2 program was still used. This resulted in the production of Trp 262 mutants. However, mutations for replacement of Trp 813 remained a problem so the PCR program was modified a third time (Table 3.2, PCR #3) and the amount of template was increased to 200 ng. This time a product was visible.

3.13.1 Digestion of parental DNA

Upon completion of the PCR reactions, the tubes were subjected to *Dpn* I digestion.

This digested the parental DNA (because it is methylated) and left the newly synthesized

DNA which contained the mutant codon. *Dpn* I (1 μL (10 U/μL)) was added directly to

the PCR reaction tube. The tube was incubated at 37°C for 1 hr.

3.13.2 Transformation into *Escherichia coli*

The *Dpn* I digested DNA was transformed into *E. coli* XL1 Blue competent cells

according to the Stratagene protocol (see section 3.11). After transformation the reactions

Table 3.2: PCR Cycling Parameters

	Segment	Cycles	Temperature (°C)	Time
PCR#1	1	1	95	30s
	2	16	95	30s
			55	1 min.
			68	12 min.
PCR#2	1	1	95	30s
	2	4	95	30s
			75	1 min.
			68	12 min.
	3	4	95	30s
			70	1 min.
			68	12 min.
	4	4	95	30s
			65	1 min.
			68	12 min.
	5	4	95	30s
			60	1 min.
			68	12 min.
PCR#3	1	1	98	1 min.
	2	8	98	30s
			74	1 min.
			68	12 min.
	3	8	98	30s
			70	1 min.
			68	12 min.

were plated on low salt LB plates containing 25 µg/mL Zeocin. These plates were incubated overnight at 37°C.

3.13.3 Plasmid preparation for sequencing

Colonies that could grow in the presence of Zeocin contain the plasmid that was subjected to the mutation protocol. These were grown overnight at 37°C with shaking (225 rpm) in low salt LB (100 mL) with Zeocin (25 µg/mL). The plasmids were then purified from the *E. coli* suspension using the Qiagen protocols. Plasmids were cut with *Eco* RI and analyzed by electrophoresis to determine size and concentration for sequencing.

3.13.4 Sequencing

All sequencing was done at UCDNA Services. Primers were supplied in 5 uL aliquots at a concentration of 15 pmol. Template DNA was supplied in 10 uL aliquots with approximately 5 µg DNA.

3.14 Transformation into *Pichia pastoris*

Once the plasmids were sequenced and the mutations were verified, the plasmids were transformed into *P. pastoris*. Stable transformants can be generated in *P. pastoris* using linear DNA, which undergoes homologous recombination between the transforming DNA and regions of homology within the genome. Single cross-over events (insertions) are usually more likely than double cross-over events (replacements). With the Invitrogen Easy Express system, the gene insertion occurs at the *AOX1* (for X-33 and GS115) or *aox1::ARG4* (for KM71) loci from a single cross-over event between the loci and either of the two *AOX1* regions contained in the pPicZAα vector. This results in insertion of the vector upstream or downstream of the *AOX1* or *aox1::ARG4* genes in the Mut⁺ (X-33,

GS115) or Mut⁺ (KM71) strains (see Figure 3.8).

3.14.1 Preparation of competent *Pichia pastoris* cells

P. pastoris competent cells were prepared using the guidelines supplied for the *Pichia* EasyComp™ kit. Competent cells of the X-33, GS115 and KM71 *Pichia* strains were produced. The method was the same for all strains. The *Pichia pastoris* strains were grown in YPD media at 28-30°C. After 16 hours the cells were diluted with YPD to an $O.D._{600}$ of 0.1-0.2 in 10 mL YPD media. The diluted cells were allowed to continue growing until the cultures reached an $O.D._{600}$ between 0.6 and 1. The cells were then pelleted by centrifugation (500 x g) and were made competent chemically. The actual identities of the solutions in the kit were not available, but the process involved a variety of buffered salts ($CaCl_2$, KCl, KOH, $MnCl_2$) and DMSO. The competent cells were stored at -70°C.

3.14.2 Transformation into *Pichia pastoris*

The Invitrogen transformation protocol was utilized for transformation into *Pichia pastoris*. If the plasmid DNA is linearized the frequency of insertion into the genome is increased. Linearization was done by incubating 10 μL of plasmid DNA with 4.4 μL (10 Units/μL) of *Pme* I and 1.6 μL of NEB4 reaction buffer. *Pme* I was chosen since it cuts uniquely in the pPicZAα plasmid and does not cut within the β–glucosidase gene. The digestion was allowed to proceed overnight at 37°C in order for all of the plasmid DNA to be cut (Figure 3.9). Each transformation required 50 μL of cells and 3 μg of linearized plasmid DNA. The transformation reactions (linear plasmid & yeast cells) were plated on YPDS containing 100 μg/mL Zeocin. The plates were incubated for 2-4 days at 30°C.
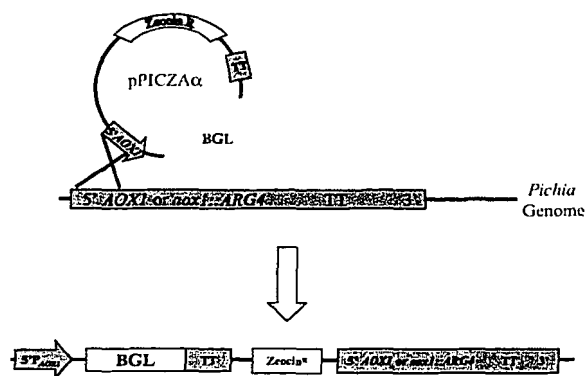
Figure 3.8. Gene insertion at *AOX1* (X-33 or GS115 *Pichia* strains) or
aox1::ARG4 (KM71 *Pichia* strain) loci. These events arise from a
single crossover event between the loci and either of the two
AOX1 regions (the *AOX1* promoter or the *AOX1* transcription
terminator (TT) region) on the pPICZAα vector. This event
results in at least one copy of the vector downstream or upstream
of the *AOX1* or the *aox1::ARG4* genes. For X-33 and GS115, the
resulting phenotype of the yeast strain is Mut⁺. For KM71 strains,
the resulting phenotype is Mutˢ. Recombination is facilitated by
initially linearizing the plasmid. BGL represents the β-
glucosidase gene. Adapted from the Invitrogen Pichia Easy Select
(Version C) Manual, 1999.
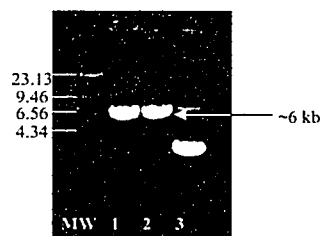
Figure 3.9. Agarose gel of restriction enzyme digestion of pPicZAα:β-
glucosidase with *Pme I*. *Pme I* was used for linearization prior to
transformation into *P. pastoris* since it cuts only the plasmid
DNA, and only cuts once. MW indicates the molecular weight
markers, and sizes for the observable bands are provided to the left
of the gel. Lanes 1 and 2 contain the digested pPicZAα:β-
glucosidase plasmid and lane 3 contains the undigested plasmid.

Transformation controls were also done. The first control showed that the *Pichia* strain cannot grow without the presence of the plasmid. The second control showed that the plasmid must be transformed into an organism to show any growth.

3.15 Confirmation of β-glucosidase mutations in the *Pichia pastoris* genome

The *Pichia pastoris* were tested to determine if the insertion of the β-glucosidase gene occurred. Genomic DNA was isolated (Section 3.15.1) from the mutant and wild type strains and subjected to PCR (Section 3.15.2) using primers specific for the alcohol oxidase gene. PCR products of the appropriate size were isolated from an agarose gel. The PCR product still contained the unique restriction sites used to place the β-glucosidase gene in the MCS. Therefore the PCR product was digested with these enzymes (*Not I* and *Eco RI*) to allow the β-glucosidase gene to be ligated into the pPicZAα plasmid that had been digested by the same restriction enzymes (Section 3.11.2). The plasmid was then transformed into *E. coli* XL2 Blue cells (Section 3.11.3). The plasmids were purified from *E. coli* using the Qiagen mini-prep protocol and sent for sequencing.

3.15.1 Total DNA isolation from *Pichia*

The *Pichia* strains were grown at 30°C to an $O.D._{600}$ of 5-10 in 10 mL of MGYH. The cells were collected by centrifugation at 1500 x g for 10 min. at room temperature. The cells were then washed with sterile water (10 mL) and centrifuged again. Spheroplasting of the cells was done by adding 2 mL of SCED buffer, pH 7.5, and 0.1-0.3 mg of Lyticase and incubating at 37°C for 50 minutes. The cells were then lysed by the

addition of 2 mL of 1% SDS (incubated on ice for 5 min.) and 1.5 mL of 5 M potassium

acetate, pH 8.9. After mixing, the lysed cells were centrifuged at 10 000 x g for 10 min. at

4°C. The DNA could then be precipitated from the supernatant by adding two volumes of

ethanol. After incubating for 15 min. at room temperature, the mixture was centrifuged at

10 000 x g for 20 min at 4°C. The supernatant was carefully removed and the pellet

resuspended in 0.7 mL of TE buffer, pH 7.4. The DNA was then extracted by adding an

equal volume of phenol:chloroform (1:1 v/v) followed by an equal volume of

chloroform:isoamyl alcohol (24:1). The aqueous layer was removed and a half volume of

7.5 M ammonium acetate (pH 7.5) and 2 volumes of ethanol were added. The tubes were

then incubated at –20°C for 60 min. After the incubation, the mixture was centrifuged at

10 000 x g for 20 min at 4°C and the pellets were washed with 1 mL of 70% ethanol. The

pellets were then air dried and resuspended in TE buffer, pH 7.4 (50 μL). The DNA

concentration was estimated from an agarose gel and stored at 4°C.

3.15.2 PCR analysis of *Pichia* integrants

Table 3.3 shows the PCR reaction protocol for the analyses of the *Pichia* genomic

DNA for integration of the β-glucosidase gene. Controls of pPicZAα plasmid and

pPicZAα plasmid containing the insert were also done (~100 ng of each). The reactions

were layered with mineral oil (30 μL) prior to PCR. The PCR instrument was set up to run

the program as indicated in Table 3.4 .

Table 3.3. PCR Reaction Set Up.

PCR Reaction Contents	Volume (µL)
10X PCR Buffer	5
Genomic DNA (~1µg)	10
25mM dNTPs	4
5'AOX primer (0.1µg/µL)	2
3'AOX primer (0.1µg/µL)	2
Sterile ddH₂O	26
Pfu Turbo™ DNA Polymerase (2.5U/µL)	1

Table 3.4. PCR Program

Step	Temperature (°C)	Time (min.)	Cycle
Heat Soak	94	2	1
Denaturation	94	1	25
Annealing	55	1	
Extension	68	5	
Final Extension	68	7	1

The results were analyzed on an agarose gel. In the case of Mut⁻ integrants (see Figure 3.10) two bands should be visible, one should correspond to the size of the gene of interest and the other to the *AOX1* gene (~2.2 kb). In Mut⁺ strains (KM71), the *AOX1* gene
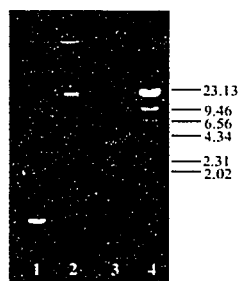
Figure 3.10. PCR analysis of *Pichia pastoris* integrants. The agarose gel illustrates the presence of the β-glucosidase gene in the genomic DNA of *P. pastoris*. DNA was isolated according to Section 3.15.1. 5'*AOX1* and 3'*AOX1* primers were used for PCR. The gel lanes were: 1) A positive control with the pPICZAα plasmid only (no gene insert, and not transformed into *P. pastoris*) -expected size is 588 bp (the distance from 5' end of the 5'*AOX1* primer to the 3' end of the 3'*AOX1* primer), 2) genomic DNA isolated from *P.pastoris* that was transformed with the wild type β-glucosidase enzyme, 3) products observed from PCR reaction with the genomic DNA of lane 2 (the wild type β-glucosidase gene (2597 bp)) plus 538 bp of *AOX1* DNA (588 bp - 50 bp lost during integration of the β-glucosidase gene into the pPICZAα plasmid) gives a product of ~ 3.1 kb). The second band is from the genomic copy of the *AOX1* gene (~ 2.2 kb). 4) molecular weight markers, sizes are indicated in kb on the right.

band will be larger due to the ARG4 insert (3.6 kb). The control with a pPicZAα plasmid

without an insert should give a 588 bp band, while the control with the plasmid containing

the insert should give a band of ~3.1 kb (size of insert (2.56 kb) & some plasmid DNA

(0.54 kb)).

3.16 Expression in *Pichia pastoris*

3.16.1 Determining Mut phenotype

In order to verify that the yeast strains were behaving as expected the Mut

phenotype was determined. The X-33 and GS115 strains should be Mut⁺ since the

linearized plasmids favor single crossover recombination at the *AOX1* locus. However,

because of the presence of the *AOX1* sequences in the plasmid, there is a chance that

recombination will occur in the 3'*AOX1* region and disrupt the wild-type *AOX1* gene,

creating Mut⁺ transformants. The Mut phenotype of the KM71 strains do not need to be

examined since they will all be Mut⁺ because the wild-type *AOX1* gene is already

disrupted in these strains. To determine the phenotype of X-33 and GS115 strains,

individual colonies were streaked onto MMH and MDH plates that had been divided into

small sections. The MMH plates were streaked before the MDH plates. Since growth is

more difficult for the organism when utilizing methanol as a carbon source, it is important

that the MMH patches contain a lot of cells (especially for the growth of the Mut⁺ strains).

Control strains of Mut⁺ and Mut⁺ strains (supplied by Invitrogen) were streaked onto the

plates for comparison. The plates were then incubated at 30°C for 2 days. Cells that grew

normally on the MDH plates but had little growth on the MMH plates were Mut⁺ strains.

Cells that grew normally on both types of plates were Mut⁺.

3.16.2 Yeast growth

A repression/derepression mechanism and an induction mechanism regulate the expression of the *AOX1* gene. Growth on glucose represses the transcription of the *AOX1* gene even if methanol is present. Therefore, it is important to use a glycerol containing medium to have optimal induction by methanol (BMGY media was used). However, methanol must be present for any level of *AOX1* expression. Growth on glycerol alone (derepression) is not sufficient to cause expression of the *AOX1* gene.

3.16.3 Methanol induction

To express Mut⁺ strains, 25 mL of BMGY was inoculated and incubated overnight at 28°C with shaking (225 rpm). The cells were harvested by centrifuging at 3000 x g for 5 min. The supernatant was removed and the cells resuspended in 100 mL BMMY to induce expression of the *AOX1* gene and therefore the gene of interest. Methanol (100%) was added every 24 hrs. to give a final concentration of 0.5% (v/v). This maintained the induction. In order to determine the optimal length of induction, aliquots of the media (1 mL) were taken to determine expression levels. Aliquots were also taken from an X-33 strain without the plasmid insert, a KM71 strain containing the plasmid but no β-glucosidase gene, and from the control for secreted expression. The aliquots were centrifuged for 2-3 min at room temperature and the supernatant was frozen in liquid nitrogen until it could be assayed. For the Mut⁺ strains, aliquots were taken at 0, 6, 12, 24, 36, 48, 60, 72, 84, and 96 hrs. For the Mut⁺ strains, aliquots were analyzed after 1, 24, 48, 72, 96, 120, and 144 hrs. Since the plasmid contained the α secretion factor, the β-glucosidase was present in the media. Therefore, after centrifugation to remove the

cells, the supernatant was examined for activity. The media aliquots were analyzed for β-glucosidase activity by reacting with 100 mM pNPGlc for 10 min. and adding Na_2CO_3. The intensity of the yellow pNP was noted (Figure 3.11). The aliquots were also analyzed by SDS-PAGE (Figure 3.12). After the initial time trial, X-33 strains were induced for 4-5 days before the media was harvested and KM71 strains were induced for 5-6 days. Two control strains were also expressed. The GS115 Mut⁺ albumin strain is a control for secreted expression, and KM71/pPicZAα (no insert) and X-33 strains were used to determine any background expression of β-glucosidase by *P. pastoris*. To express Mut⁺ strains, 100 mL of BMGY was inoculated and grown overnight at 28°C with shaking. The cells were treated the same as the Mut⁺ strains with the exception that they were resuspended in 25 mL of BMMY. The expression for both strains could be scaled up so that more protein is obtained. However, the small scale expression produced about 0.7 mg of protein, which allowed a large number of assays to be done with fresh enzyme.

3.16.4 Purification of enzymes from *Pichia pastoris*

The major advantage of secretion of the expressed β–glucosidase is that *Pichia pastoris* secretes very low levels of endogenous proteins. Additionally, under the conditions of methanol induction, the alcohol oxidase 1 (AOX1) protein is expressed at very high levels. Since the *AOX1* promoter is linked to the β-glucosidase gene, the gene will also be highly expressed. Furthermore, since very few other proteins are secreted, the protein found in the media will be mostly the heterologously expressed protein. To purify the β–glucosidase from the media, the cell suspension were centrifuged at 3300 x g for

103

Figure 3.11. Analysis of β-glucosidase expression in *Pichis pastoris*. Results
are based on enzyme reaction with pNPGlc (20 μL, 100 mM), for
a specific time period. A small amount of media containing the
secreted enzyme was added (10 μL). Reactions were halted by
the addition of 1M Na_2CO_3 (40 μL). The degree of yellow color is
indicative of enzyme activity. This analysis shows that the β-
glucosidase was expressed in both: A) Mut[+] (methanol utilization
plus), and B) Mut[S] (methanol utilization suppressed) strains and
that the amount increased with time. Controls of the Mut[+] and
Mut[S] strains were also assayed for β-glucosidase activity which
indicated that no β-glucosidase was induced in these strains. A
control of pNPGlc alone (buffer was added (10 μL)) is also
shown to illustrate the amount of color observed in the absence of
enzyme.

A)

Time (hrs)	X-33-3	X-33-4	X-33-5	X-33-6	GS115-1	Mut⁺ (control)	pNPG (control)

0
6
24
33
48
72

B)

Time (hrs)	KM71-1	KM71-2	Mut⁻ (control)

0
24
48
72
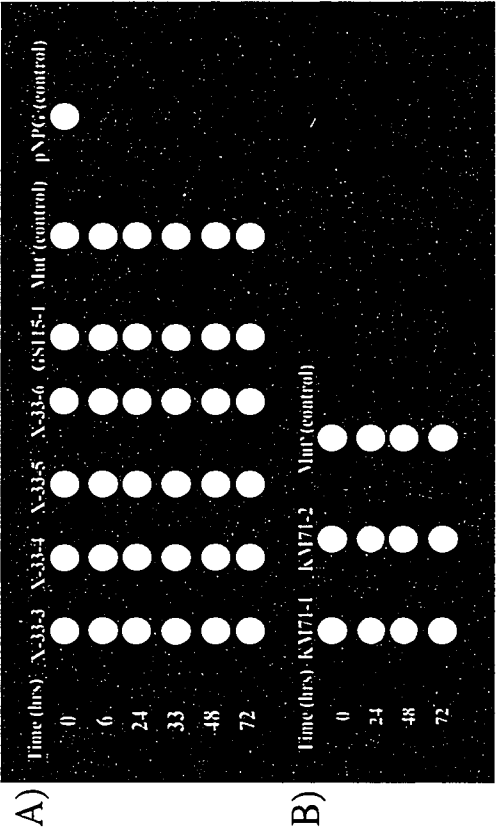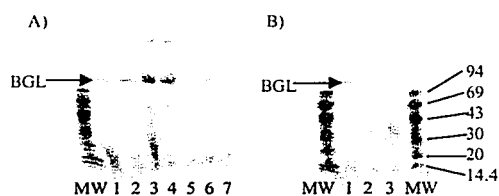
Figure 3.12. SDS-PAGE analysis of β-glucosidase expression by different
strains of *Pichia pastoris* prior to gel exclusion chromatography,
but following ammonium sulfate precipitation. The smallest
volume possible was used to solubilize the salted out proteins, but
the volumes were not equal. The molecular weight markers are the
same for both gels and the sizes are indicated at the right of gel B.
The position of β-glucosidase (BGL) is indicated with arrows. In
gel (A) lanes 1 to 4 illustrate the β-glucosidase expression in X-33
strains of *P. pastoris* (colonies 3-6). The volumes of the samples
were 500, 500, 100, 250 μL, respectively. Lane 5 shows the
expression of the enzyme in a GS115 strain (250 μL). Lane 6
shows the expression of the W262L strain obtained from Dr.
Shoseyov, a GS115 strain of *P. pastoris* that was transformed with
the mutant enzyme(250 μL). Lane 7 is a *P.pastoris* Mut⁺ control
strain that does not contain/secrete the β-glucosidase gene (200
μL). In gel (B), lanes 1 and 2 display the β-glucosidase expression
from KM71 *P. pastoris* strains (500 μL), and lane 3 shows the *P.
pastoris* Mutˢ control strain expression (200 μL).

20 min. at 4°C. This removed the yeast cells from the media. Ammonium sulfate (60%) was added to the media containing the β-glucosidases and the mixture was stirred at 4°C for 30 min. These were then centrifuged at 15 500 x g for 30 min. and the pellet was resuspended in a small amount of water (500-1000 μL) for SDS-PAGE analysis. Water was utilized because protein samples did not separate properly on a gel in the high salt environment present after ammonium sulfate precipitation. After an aliquot was removed for SDS-PAGE analysis, sodium acetate (150 mM, pH 4.5) was added to a volume of 10 mL. The suspension was then applied to a Sephacryl S-300 column and eluted with sodium acetate buffer (150 mM, pH 4.5) at a rate of 2.5 mL/min. Fraction collection was begun at 1 hr. and fractions were assayed for enzyme activity using pNPGlc (Section 3.6.4). SDS-PAGE gels of the concentrated fractions containing β-glucosidase were completed to determine purity. The protein was usually pure after elution from the Sephacryl column. However, in some cases the concentrated protein required further purification and was applied to Superose 6 and 12 columns connected in series with an FPLC. The FPLC was run at 0.25 mL/min. and 1 mL fractions were collected. The purity of these fractions was also assessed by SDS-PAGE.

3.17 Storage of *Escherichia coli* and *Pichia pastoris* strains

All *E. coli* and *P. pastoris* strains were stored as glycerol stocks. In the case of *E. coli*, 2 mL of low salt LB was inoculated with the strain of interest and grown overnight at 37°C in the presence of Zeocin. Sterile glycerol was added to 15% (v/v), and the mixture was frozen in liquid nitrogen and stored at -70°C. *P. pastoris* cells were grown overnight

at 30°C in YPD containing Zeocin. Sterile glycerol was added to 15% and the mixture was frozen using liquid nitrogen and stored at -70°C.

3.18 Fluorescence spectroscopy

The fluorescence of the enzymes was measured by the rate of photon emission. Spectra were recorded at 1 nm/s from 300 to 450 nm and one reading was taken per nm. The excitation wavelengths used were 275, 285 and 295 nm. The slit widths were 5 nm excitation and 1 nm emission. The emission spectrum obtained from a blank (150 mM acetate buffer, pH 4.5) at the same excitation wavelengths was subtracted from the protein spectra. The absorbance of the protein samples at 280 nm was determined prior to the fluorescence experiment and the concentration was determined. Protein samples were prepared by diluting with acetate buffer (150 mM, pH 4.5) to a concentration of 0.017 mg/mL immediately before running the experiment. Samples (2.75 mL) were placed in a four-sided fluorometer cuvette.

Excitation spectra were also obtained. The spectra were recorded at an emission wavelength of 339 nm. The slit widths were 1 nm excitation and 5 nm emission. An excitation spectrum was also obtained for the blank and it was subtracted from the protein excitation spectra. The excitation was measured from 200 to 300 nm at a speed of 1 nm/s and one reading was taken per nm. Protein samples were prepared in exactly the same way as they were for the emission spectra. In most cases the diluted enzyme utilized for the emission spectra was also used for the excitation spectrum.

3.19 Analysis of the oligosaccharide component of β-glucosidase

This was done by two methods. 1) NMR of the β-glucosidase enzyme was done using a 5 mm probe at an enzyme concentration of 0.5 mg/mL. 2) The amount and the identity of the sugars present was done with two FACE™ (Fluorophore-Assisted-Carbohydrate-Electorphoresis) kits purchased from Glyko.

3.19.1 N-linked oligosaccharide profiling

The β-glucosidase was purified and the concentration was estimated. The protein was concentrated to a small volume (~100 µL) and diluted with water to lower the salt concentration of the buffer. Following dilution, the enzyme was again concentrated and diluted. Following the second dilution the protein was lyophilized and resuspended in 100 µL of water. Approximately 200 µg of protein was present. The protein was treated with SDS and β-mercaptoethanol and this was boiled to ensure complete denaturation. Following boiling, a detergent, NP-40, was added. Peptide N-glycosidase F (PNGase F) was then added and the mixture was incubated overnight at room temperature. The protein was then precipitated with 100% cold ethanol and pelleted out by centrifugation (5 min. at 10 000 x g). The oligosaccharides remained in the supernatant. The supernatant was lyophilized. After lyophilization the samples were prepared for the profiling gel.

The oligosaccharides were labeled with a fluorophore, 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) by reductive amination (Figure 3.13). The reaction was allowed to occur over 16 hours at 37°C. The fluorescently labeled oligosaccharides were then separated on an OLIGO Profiling Gel. The resulting bands were visualized with UV light (~312 nm). An oligo ladder standard was also supplied by the manufacturer. The ladder
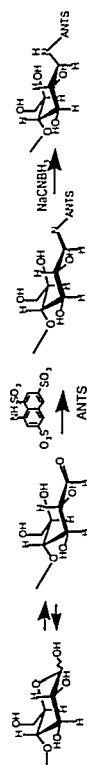
Figure 3.13 Fluorophore labeling of oligosaccharides by reductive amination. Released oligosaccharides are labeled with a fluorophore (ANTS) at the reducing termini by reductive amination. Only one molecule of fluorophore is attached to each molecule of oligosaccharide. Adapted from the Glyko Profiling manual, 1994.

was composed of glucose polymers of 1 to 20 glucoses. The glucose units (1 to 20) can be identified on the gel since the oligosaccharide with 4 glucoses was at a lesser concentration and was therefore, identifiable (the only faint band). A control that consisted of trypsin inhibitor was also provided. The profile for this enzyme is known and could therefore be used to determine if the enzyme digestion and fluorophore labeling occurred properly. A quantitation control is also provided to determine the amount of sugar present in the carbohydrate bands. However, in order to quantitate the control, the FACE imaging system must be used for visualization of the carbohydrate gels. The prepared samples, controls and glucose ladder were loaded (4 μL) onto the profiling gel which was run at a constant current of 15 mA per gel. Profiling gels were run until the dye front reached the bottom of the gel (~1hr). Once the gel was finished, a picture was obtained using an Eagle III gel scanner.

3.19.2 N-Linked Oligosaccharide Sequencing

The unique oligosaccharide chains observed on the profiling gel were sequenced by utilizing specific endo and/or exo-glycosidase enzymes. This was accomplished by the purification of the unique bands observed in the profiling gel and exposing them to different mixtures of glycosidase enzymes. The individual bands were removed from the gel with a razor blade, and the bands were eluted from the gel slice with water. Once the oligosaccharides were eluted from the gel they were treated with a variety of exoglycosidases. The result of this digestion was visualized on an oligosaccharide sequencing gel, and the pattern of bands indicated which sugars were present in the oligosaccharide chain. The enzymes utilized were neuraminidase (NANaseIII) which is

111

specific for α2-(3,6,8) linked N-acetylneuraminic acid, β-galactosidase (GALaseIII) which is specific for β1-4 linked galactose, β-N-acetylhexosaminidase (HEXaseIII) which is specific for β1-(2,3,4,6) N-acetylglucosamine, α-mannosidase (MANaseII) which is specific for α1-(2,3,6) linked mannose and a second α-mannosidase (MANaseVI) which is specific for α 1-6 linked mannose. For each oligosaccharide that was sequenced, 5 separate enzyme reaction digests were required. Therefore, the first tube contained only the oligosaccharide and no enzymes, the second tube contained only NANaseIII which releases sialic acid. The third tube contained NANaseIII and GALaseIII, which releases both sialic acid and galactose. The fourth tube contained NANaseIII, GALaseIII and HEXaseIII that releases sialic acid, galactose and N-acetylglucosamine. The fifth tube contained all of the above enzymes in addition to the α-mannosidases (MANaseII) which degrade the oligosaccharide down to the Man(β1-4)GlcNAc(β(1-4)GlcNAc or Man(β1-4)GlcNAc(β1-4)[Fuc(α1-6)]GlcNAc core structures. MANase VI was added to the fifth tube following digestion by the other enzymes to complete the digestion of Man(α1-6).

Oligosaccharide sequencing was performed by carrying out the digestions discussed above. The digestions were loaded onto a sequencing gel (profiling and sequencing gels are the same) and electrophoresed in the same manner as the profiling gels. The gels were run until the dye front reached the bottom of the gel. The gels were visualized with a UV light (~312 nm). Images were evaluated according to the migration values provided by Glyko (1994) and by comparison with the glucose polymer ladder.

4. RESULTS

Part I:

Characterization of the β-Glucosidase from *Aspergillus niger* Cellulase Powder

4.1 Purification, size and quaternary structure

The β-glucosidase from the Sigma *A. niger* cellulase powder (catalogue number C-1184) was purified. Two gel filtration steps usually proved sufficient since the majority of the proteins in the cellulase mixture were of smaller size. The enzyme was shown to be > 98% pure (determined at high protein concentration) by SDS (Figure 4.1(A)) and native PAGE (Figure 4.2(A)).

Comparisons to markers on SDS (Figure 4.1(B)) and native PAGE (Figure 4.2(B)) indicated that the monomeric molecular weight was approximately 125 kDa and that the native β-glucosidase is a dimer with an approximate molecular weight of 230 kDa. Native PAGE gels are not a very accurate method of determining molecular weight. However, the values obtained were more or less similar to the molecular weights from gel filtration and dynamic light scattering, which are discussed below. Subjection of the purified enzyme to NMR (to be described later) gave spectral bands characteristic of carbohydrate. That study suggested that the enzyme is a glycoprotein. Other analyses, to be shown, confirmed the glycoprotein nature of the enzyme. Glycoproteins migrate anomalously on SDS and native gels and that also indicates that some caution must be exercised in assigning definite molecular weights.

A)

β-glucosidase

B)

log
Molecular
Weight

Relative Mobility
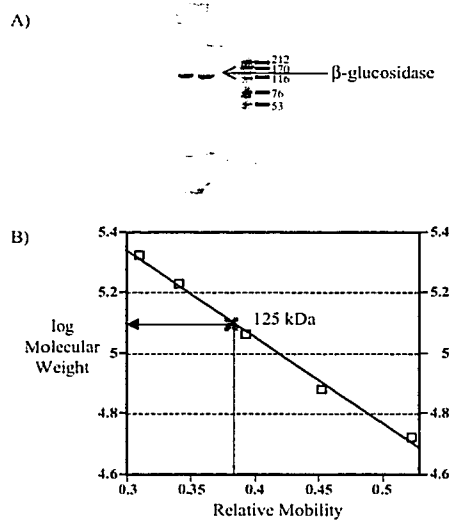
Figure 4.1. SDS-PAGE (A) of the β-glucosidase after purification.
Molecular weight markers were myosin (212 kDa), α2-
macroglobulin (170 kDa), β-galactosidase (116 kDa),
transferrin (76 kDa) and glutamic dehydrogenase (53 kDa). B)
Plot of the log molecular weight of the markers and their
respective relative mobilities that were used to determine the
monomeric molecular weight of the β-glucosidase.

114

A)

β-glucosidase

669
440
232
140
67

B)

log
Molecular
Weight

Relative Mobility
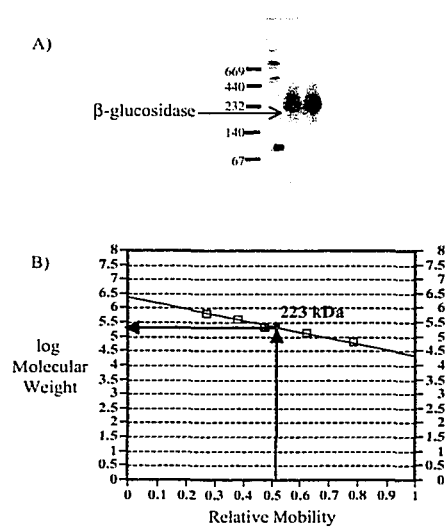
Figure 4.2. Native PAGE (A) of the β-glucosidase after purification. The
native gel was overloaded to indicate enzyme purity. Molecular
weight markers were: thyroglobulin (669 kDa), ferritin (440 kDa),
catalase (232 kDa), lactate dehydrogenase (140 kDa) and albumin
(66 kDa). B) Plot of the log molecular weight of the markers in (A)
and their respective relative mobilities that were used to determine
the native molecular weight of the β-glucosidase.

An analysis of the native molecular weights was also done by size exclusion chromatography. Comparison with standards suggested a molecular weight of approximately 210 kDa (Figure 4.3). In addition, dynamic light scattering (DLS) suggested that the molecular weight of the native protein is 188 kDa. This molecular weight was calculated from the hydrodynamic radii illustrated in the histogram in Figure 4.4. Finally, the values of molecular weight obtained from gel filtration and from DLS are consistent with the estimated molecular weight (assuming the protein is a dimer) obtained from the protein sequence (see later), which showed that the monomeric molecular weight is 91.4 kDa. This molecular weight was determined only from the protein sequence and does not take into account the oligosaccharide component of this enzyme.

4.2 Extinction coefficient

The extinction coefficient of the enzyme was determined using bovine γ-globulin as the standard. γ-Globulin has been shown to provide accurate protein concentration data when the Bradford assay method is used (Kruger, 1994). The extinction coefficient of the β-glucosidase calculated from Beer's Law was found to be $1.8\ cm^2 mg^{-1}$.

4.3 Temperature stability

The V_{max} with pNPGlc increased with temperature up to 65°C. The activity decreased after 65°C (data not shown) and the drop was especially sharp between 75°C and 85°C. Additional data indicated that the enzyme was stable at 65°C since only a small decrease in the V_{max} was observed over a 30 min. time period. The enzyme was sensitive to freezing and thawing. The β-glucosidase in 150 mM sodium acetate buffer, pH 4.5, at a
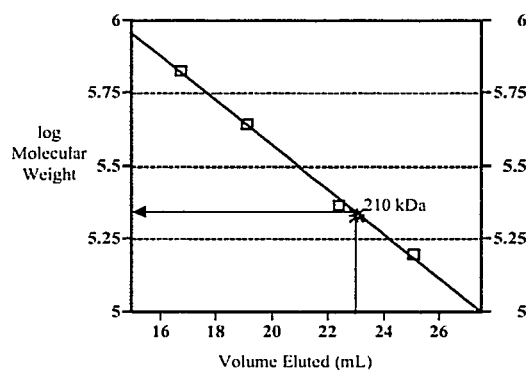
Figure 4.3. Plot of the relationship between the elution volume from the
Superose 6 and 12 size exclusion columns joined in series on the
FPLC and the log of the molecular weight of specific proteins. The
proteins that were used were thyroglobulin (669 kDa), ferritan (440
kDa), catalase (232 kDa), and adolase (158 kDa). The FPLC was
run at a rate of 0.2 mL/min. The β-glucosidase eluted at 23.1 mL
(determined for 3 runs) which was equivalent to a molecular weight
of 210 kDa.
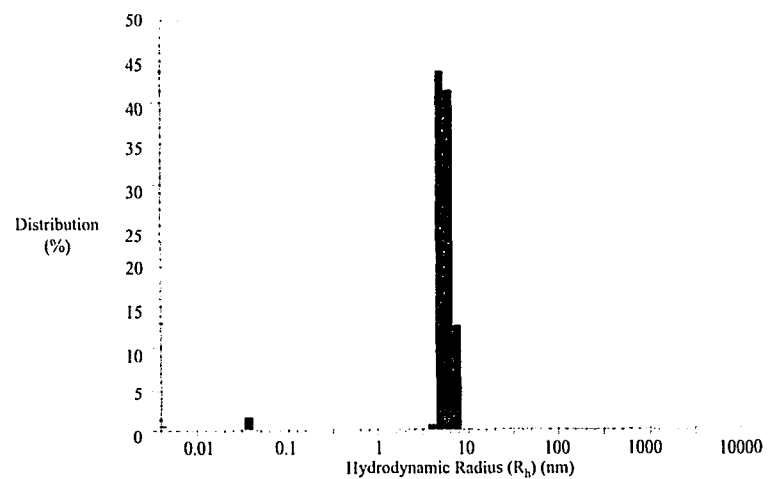
Figure 4.4 Distribution of macromolecule size in a 5.05 mg/mL solution of β-glucosidase (isolated from the Sigma cellulase powder mixture). Macromolecule size is shown in terms of the hydrodynamic radii. The small peak that occurs at an approximate radius of 0.015 nm results from the sodium acetate buffer. The large peak centered at about 8 nm resulted from the β-glucosidase.

concentration of 0.5 mg/mL was frozen in liquid nitrogen for 10 min. and thawed at room

temperature. It lost approximately half of its activity when it was assayed at 25°C (data not

shown).

4.4 Fluorescence characterization

Emission spectra (Figure 4.5) were obtained for the β-glucosidase at excitation

wavelengths of 275, 285, and 295 nm. The enzyme showed the greatest fluorescence

intensity (photon counts per second (cps)) when excited at 275 nm. The fluorescence

peaks occurred at wavelengths between 337-339 nm for the 3 emission spectra.

An excitation spectrum was also obtained. The emission wavelength used for the

excitation spectrum was 339 nm. The spectrum showed a main peak at 283 nm, and a

shoulder peak at 288 nm (Figure 4.6).

4.5 Unusual kinetics of the Sigma enzyme

4.5.1 Eadie-Hofstee plots for analysis of enzyme kinetics

The central expression for steady state enzyme kinetics (initial velocity) is the

Michaelis-Menten equation:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

The Eadie-Hofstee equation is derived from the above equation by multiplication of both

sides by Km +[S], division by [S], and a rearrangement to produce:

$$v = V_{max} - K_m\left(\frac{v}{[S]}\right)$$

A plot of v (initial velocity) as a function of v/[S] would be expected to produce a straight-
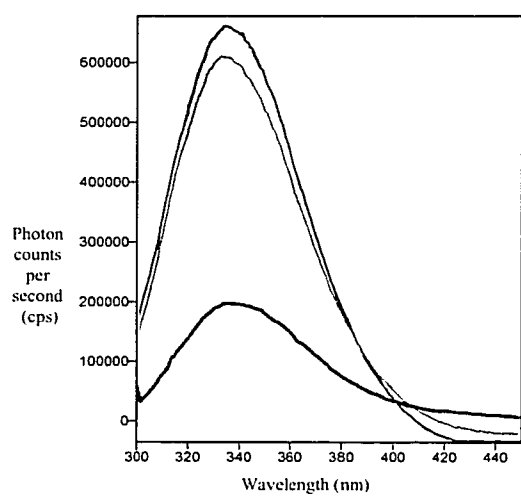
Figure 4.5. Emission spectra of β-glucosidase excited at 275 nm (blue) 285
nm (pink), and 295 nm (green). Spectra were recorded on a
Jobin Yvin Spex: Fluorolog-3 from 300 to 450 nm at a rate of 1
nm/s. The β-glucosidase was diluted to a concentration of 17
μg/mL in sodium acetate buffer, 150 mM, pH 4.5. A spectrum
of the buffer was also obtained at each of the excitation
wavelengths and this was subtracted to produce the spectra
(shown above). The slit width for excitation was set at 5 nm and
the slit width for emission was set at 1 nm.
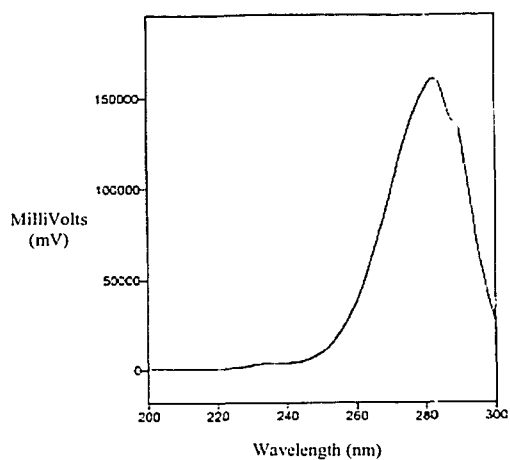
Figure 4.6. Excitation spectrum of β-glucosidase. The spectrum was
acquired on a Jobin Yvon Spex: Fluorolog-3 from 200 to 300
nm. The excitation spectrum was collected at an emission
wavelength of 339 nm, with a slit width of 1 nm for excitation
and a slit width of 5 nm for emission. The protein was diluted to
a concentration of 17 µg/mL in 150 mM sodium acetate buffer,
pH 4.5. The spectrum of the sodium acetate buffer under the
same conditions as the protein was subtracted to give the
spectrum shown.

line relationship with a slope of $-K_m$ and a y intercept of V_{max}. An example of a normal

hypothetical Eadie-Hofstee plot is shown in Figure 4.7.

4.5.2 Kinetic analyses of the reactions with pNPGlc, oNPGlc and cellobiose

Initial velocity experiments with the β-glucosidase with pNPGlc at pH 4.5 resulted

in atypical Eadie-Hofstee plots. The points only followed a straight line at low substrate

concentrations. When the substrate concentration was over 5 mM the initial velocity

values no longer followed a straight line. The line curved downward as it approached the y

axis. An example of such an Eadie-Hofstee plot at pH 4.5 is shown in Figure 4.8(A). A

different Eadie-Hofstee plot was obtained when the pH was pH 7. Instead of the

downward curvature that was observed at pH 4.5, the lines curved upward as they

approached the y axis. An example of this is shown in Figure 4.8(B).

The reactions with oNPGlc and cellobiose were similar. They gave normal linear

Eadie-Hofstee plots at low substrate concentrations (Figure 4.9(A)) and the rates of oNP

or glucose production also decreased at high substrate concentrations (pH 4.5). As was the

case with pNPGlc there were also upward curving plots at pH 7 (Figure 4.9(B)).

4.6 Analysis of reaction products

4.6.1 Thin layer chromatography

TLC analysis of the β-glucosidase reaction with 5 mM pNPGlc showed that it was

hydrolyzed to glucose and pNP (Figure 4.10(A)). (pNP is not detected by charring but

other experiments showed that it was formed.) There was no evidence of the presence of

any other product. TLC analysis of the β-glucosidase reaction with 40 mM pNPGlc

indicated that glucose and 2 "transient" unknown compounds are produced (Figure
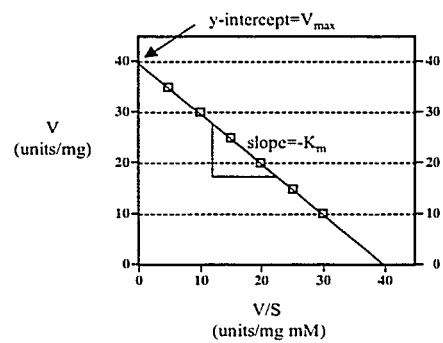
Figure 4.7. A hypothetical example of an Eadie-Hofstee plot with a slope equal to $-K_m$ and a y-intercept equal to the V_{max}. One unit is equivalent to one μmol of product per min.
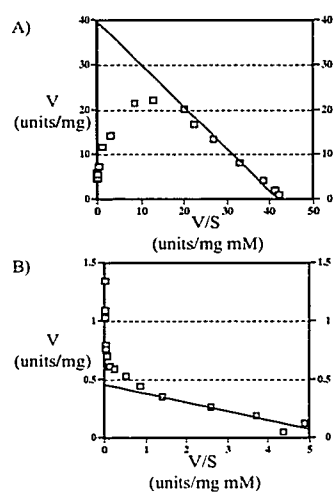
Figure 4.8. A) The Eadie-Hofstee plot that resulted from the β-glucosidase
reaction with pNPGlc at pH 4.5, 25°C. B) The Eadie-Hofstee plot
that resulted from the β-glucosidase reaction with pNPGlc at pH
7, 25°C. The enzyme concentration was 1.4 μg/mL. One unit is
equal to the production of one μmol of product per min. The lines
are drawn through the points that result in a straight line (before
the downward or upward curvatures).
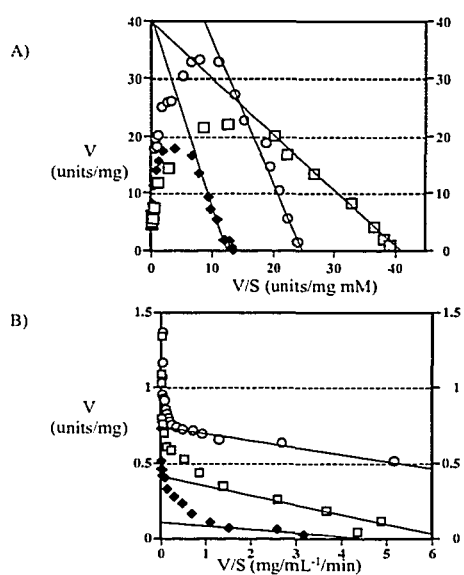
Figure 4.9. Eadie-Hofstee plots of the β-glucosidase reaction with pNPGlc

(□), oNPGlc (◆), and cellobiose (O) at: (A) pH 4.5, and (B) pH 7.

Assays were carried out as described in Section 3.6.5 (pNPGlc and

oNPGlc) and in Section 3.6.7.1. (cellobiose) at 25°C. The enzyme

concentration used was 1.4 µg/mL (pNPGlc and oNPGlc) and 8

µg/mL (cellobiose). One unit is equal to the production of one

µmol of product per min. The lines are drawn through the points

that result in a straight line (before the downward or upward

curvatures).

125

Figure 4.10. Thin layer chromatography of the β-glucosidase reaction with
pNPGlc at a concentration of : A) 5 mM and B) 40 mM. Enzyme
was added (1.4 μg/mL) to the reaction mixture at zero time and
an aliquot (1 μL) of the reaction was immediately spotted onto
the plate. Spotting was repeated at the times (min.) indicated on
the bottom of the plate. The chromatography solvent used was n-
butanol:ethyl acetate:2-propanol:acetic acid:water in the ratio of
1:3:2:1:1. Products were detected by dipping dried plates in 2%
sulphuric acid (w/v) in methanol and charring. The standards
(1μL) were: pNPGlc (P) ((A) 5mM , (B) 40 mM), 20 mM
glucose (G), 10 mM gentiobiose (Ge), and 10 mM cellobiose
(C).

A)

P G C Ge 0 1 2 4 6 8 10 12 15 20 25 30 35 40 50 60
                    (min.)

B)

P 0 1   2   4 6 8 10 12 14 16 18 20 24 28 32 G Ge C
                    (min.)
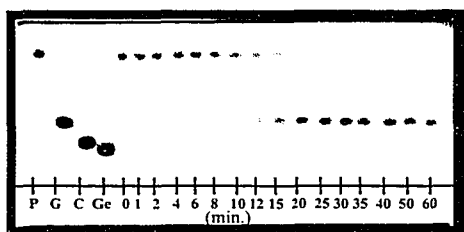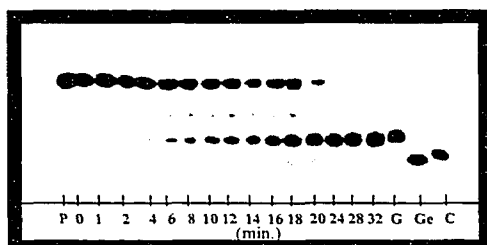
4.10(B)). One of the unknowns had an R_f of about 0.5 and it was detectable after the reaction had proceeded for 4 min. There was a little less of this unknown than of glucose. The concentration of this unknown increased and then decreased. It was never present in amounts as high as glucose and after 24 min. it could no longer be detected by TLC. The second unknown had an R_f value of 0.2, which was similar to the R_f for gentiobiose. It appeared later (at about 12-16 min). Analysis of many reactions showed that the formation of this second unknown depended on the presence of significant amounts of free glucose. The concentration of this unknown also increased and then decreased. The two unknown compounds were purified. NMR showed that the unknown compound with the R_f value of 0.5 was β-D-glucopyranosyl-(1→6)–β-D-glucopyranosyl-p-nitrophenol (pNPGe) while the compound with the R_f of 0.2 was gentiobiose. The NMR values obtained for the unknown sugars are reported in Appendix III.

TLC analysis of the enzyme reaction with cellobiose at a high concentration (95 mM) also showed that sugars in addition to glucose (Figure 4.11) were formed. A compound with an R_f of 0.4 that migrated just a little more slowly than cellobiose itself (it was hard to separate from cellobiose) was produced. Because of the crowding it is difficult to determine when it first appears but it is seen as a distinct spot after about 33 min. Glucose and a second unknown compound, with an R_f of 0.3, could be detected at early times (1 min.). This second unknown seemed to be present at concentrations higher than the glucose concentrations. The compound with the $R_f = 0.4$ (in this case the TLC plates were run in the solvent twice - hence the R_f values are different from those in the pNPGlc study) was purified and was shown (by NMR) to be gentiobiose. The unknown compound
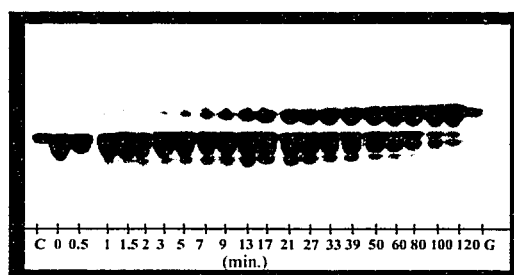
Figure 4.11. Thin layer chromatography of the β-glucosidase reaction with cellobiose (95 mM). Enzyme (8 μg/mL) was added to the reaction mixture at zero time and an aliquot (1 μL) was immediately spotted onto the plate. Spotting was repeated at the times indicated (in min.) on the bottom of the plate. The chromatography solvent used was n-butanol:ethyl acetate:2-propanol: acetic acid :water in the ratio of 1:3:2:1:1. The plates were developed twice (they were totally dried between developing). Products were detected by dipping the dried plates in 2% sulphuric acid in methanol and charring. The standards used (1μL) were: 10 mM cellobiose (C) and 10 mM glucose (G).

with the R_f value of 0.3 was also purified and found (by NMR) to be β-D-glucopyranosyl-
(1→6)-β-D-glucopyranosyl-(1→4)-D-glucopyranose (or 4-O-gentiobiosyl glucose-see
Appendix III for NMR values, Appendix II for sugar structures). This will be referred to
as either gentiobiosyl-glucose or trisaccharide in the rest of the thesis. The concentration
of the trisaccharide was highest between about 13 and 21 min. and only a small amount
could still be detected at 80 min. A significant amount of gentiobiose was still present at
120 min. but its concentration was decreasing and GLC studies (to be discussed later)
showed that it would have been totally degraded to glucose if given enough time. A
slower migrating band ($R_f = 0.25$) was also present but its identity was not determined. It
formed in the greatest concentration at times when the trisaccharide concentration was at
its zenith. From this fact and from its migration position as well as the finding that this
enzyme forms mainly β(1-6) bonds when it carries out transglucosidic reactions, it is
probably a tetrasaccharide; (β-D-glucopyranosyl-(1→6)-β-D-glucopyranosyl-(1→6)-β-D-
glucopyranosyl-(1→4)-β-D-glucopyranose).

Reactions of the β-glucosidase with cellooligosaccharides of increasing length
were also analyzed by TLC. Only the main degradation products will be discussed but
close inspection of the TLC plates showed that reactions with all of the
cellooligosaccharides also produced transglucosidic products with the substrates and
intermediates acting as acceptors. The main degradation products when cellotriose was
the substrate were cellobiose and glucose (data not shown). Mainly cellotriose and
glucose were produced initially when cellotetraose was the substrate (data not shown).
After 20 min. the cellotriose began to be broken down to cellobiose and glucose and

eventually the only product was glucose. The TLC analysis of the reaction with

cellopentaose (8 mM) is shown in Figure 4.12. Cellotetraose was produced soon after the

reaction began. Glucose was not detected on the plate at the early stages but some must

have been produced. Since cellotetraose has 4 glucose units one would expect some

difficulty in detecting glucose in comparison to the cellotetraose product (spots would be

4 times less intense). Production of cellotriose (3 min.) and finally cellobiose (6 min.)

followed. All of the intermediate cellooligosaccharides were present transiently. The

reaction progression was determined by densitometer analysis and is shown in Figure

4.13. Eventually all of the cellopentaose was accounted for by the glucose.

The reaction of β-glucosidase with 4-O-gentiobiosyl-glucose (one of the two

transglucosidic products produced upon reaction with cellobiose - see above) was also

analyzed by TLC. The R_f values of the products indicated that only cellobiose ($R = 0.34$)

and glucose ($R = 0.53$) were formed initially (data not shown). Of significance is that

gentiobiose was not produced. Eventually the only product was glucose.

TLC analysis of the reaction with gentiobiose showed that glucose and a

trisaccharide were produced (data not shown). No cellobiose was formed. The

trisaccharide was a transient product again, as eventually only glucose was present. A

small amount of the trisaccharide was purified. When it was reacted with the enzyme, the

main products were gentiobiose and glucose.

TLC was also utilized to analyze the activity of the enzyme with dinitrophenyl-β-

cellobiose (DNPC). The results (Figure 4.14) showed that an unknown ($R = 0.68$) was

formed at early times. Other experiments of longer duration showed that the unknown was

131
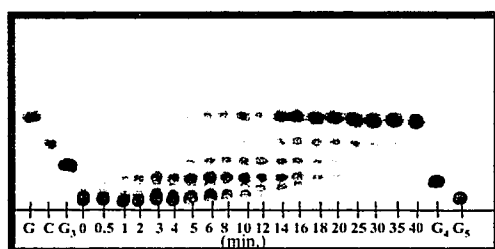
Figure 4.12. Thin layer chromatography of the β-glucosidase reaction with cellopentaose (8 mM). Enzyme was added (8 µg/mL) to the reaction mixture at zero time and an aliquot (1 µL) was immediately spotted onto the plate. Spotting was repeated at the times indicated in minutes on the bottom of the plate. The chromatography solvent used was n-butanol, ethyl acetate, 2-propanol, acetic acid and water in the ratio of 1:3:2:1:1. Products were detected by dipping the dried plates in 2% sulphuric acid in methanol and charring. The standards (1 µL) were: 10 mM glucose (G), 5 mM cellobiose (C), 10 mM cellotriose (G$_3$), 10 mM cellotetraose (G$_4$), and 8 mM cellopentaose (G$_5$).

Concentration (mM)

Time (min.)
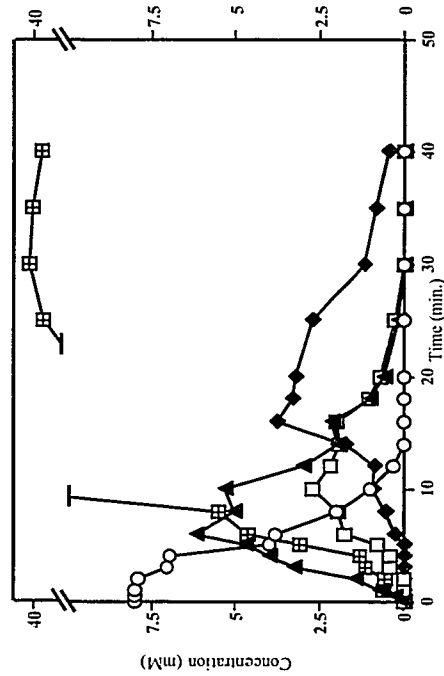
Figure 4.13. Densitometer analysis of the reaction with cellopentaose shown in Figure 4.12. The scanning parameters are given in Section 3.7.1.1. The β-glucosidase reacted with cellopentaose (O) and produced cellotetraose (▲), cellotriose (□), cellobiose (◆), and glucose (⊞). In order to show the final glucose production a portion (between 8 and 35 mM) was removed from the y-axis.
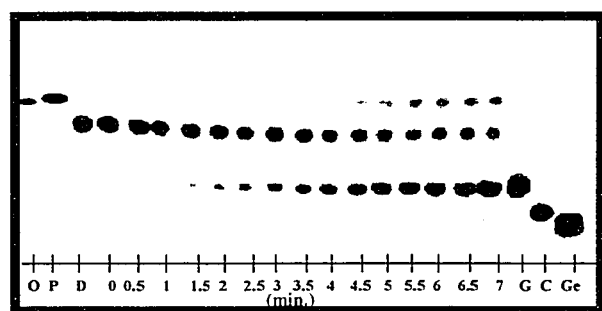
Figure 4.14. Thin layer chromatography of the β-glucosidase reaction
with 3,4-dinitrophenyl cellobioside (5 mM). Enzyme was added
(1.4 µg/mL) to the reaction mixture at zero time and an aliquot
(1 µL) was immediately spotted onto the plate. Spotting was
repeated at the times indicated in minutes on the bottom of the
plate. The chromatography solvent used was n-butanol:ethyl
acetate:2-propanol:acetic acid :water in the ratio of 1:3:2:1:1.
Products were detected by dipping the dried plates in 2%
sulphuric acid in methanol and charring. The standards used on
the plate were 5 mM oNPGlc (O), 5 mM pNPGlc (P), 5 mM
3,4-DNPC (D), 10 mM glucose (G), 10 mM cellobiose (C) and
20 mM gentiobiose (Ge).

134

transient. It was also found that the formation of significant amounts of dinitrophenol only occurred after a significant amount of the unknown had built up. The unknown interacted with the TLC fluor to fluoresce under UV light. This indicated that it contained the dinitrophenyl group. The compound migrated (Figure 4.14) to a similar location ($R_f =$ 0.69) as pNPGlc ($R_f = 0.68$) and oNPGlc ($R_f = 0.66$). This showed that it is probably 3,4-dinitrophenyl-β-D-glucopyranoside. A small amount of gentiobiose was also formed at later times, after a significant pool of glucose had built up. Similar results (data not shown) were obtained with p-nitrophenyl gentiobiose (pNPGe). pNPGlc and glucose were the initial reaction products, followed by pNPGlc breakdown.

4.6.2 Gas-liquid chromatography (GLC)

The production of glucose, gentiobiose and the gentiobiosyl-glucose (trisaccharide) was followed and quantified by GLC. GLC allowed for quantitative results to be obtained. It was not possible to use the GLC to analyze for oligosaccharides that were larger than 3 glucose units. Very high temperatures would be needed for their elution. That would have destroyed the columns. Analyses of oligosaccharides greater than 3 glucose units were, therefore, done by TLC.

4.6.2.1 Effect of pH at high and low cellobiose concentrations

The reaction of the enzyme with different concentrations of cellobiose at pH 4.5 and 7 was investigated by GLC (Figure 4.15 and 4.16). Mainly glucose was produced when the enzyme reacted with 5 mM cellobiose at pH 4.5 (Figure 4.15(A)). This showed that hydrolysis was the main reaction. Essentially no gentiobiose or trisaccharide were produced. The reaction proceeded quickly until the cellobiose was depleted. The reaction

A)

Sugar
Concentration
(mM)

Time (minutes)

B)

Sugar
Concentration
(mM)

Time (minutes)
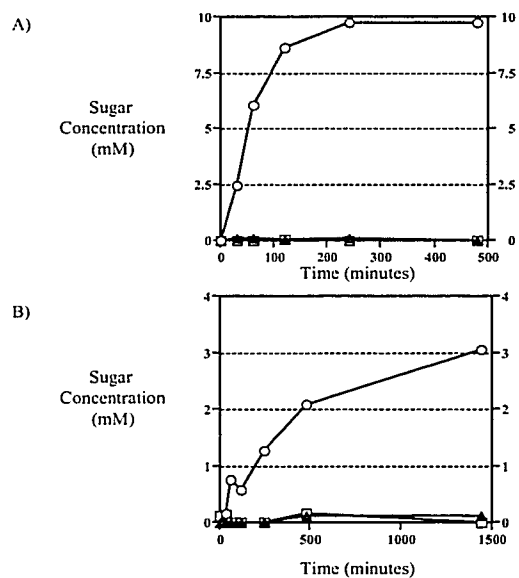
Figure 4.15. Analysis of the β-glucosidase reaction over time with 5 mM
cellobiose: at (A) pH 4.5, and (B) pH 7 by GLC. The reaction was
monitored by following the production of: glucose (O),
gentiobiose (▲), and gentiobiosyl-glucose (□). β-Glucosidase (2
μg/mL) was added to the reaction mixture which was incubated at
25°C. Experimental details are given in Section 3.8.1.

A)

Concentration
of
sugar
(mM)

Time (min.)

B)

Concentration
of
sugar
(mM)

Time (min.)
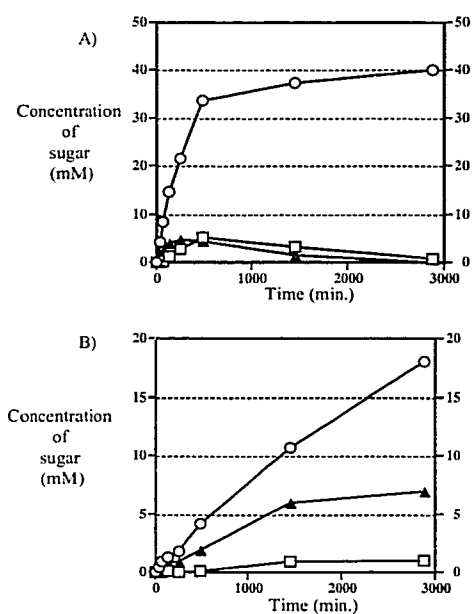
Figure 4.16. Analysis of the β-glucosidase reaction over time with 50 mM
cellobiose at (A) pH 4.5 and (B) pH 7 by GLC. The reaction was
monitored by following the production of glucose: (O),
gentiobiose (□), and gentiobiosyl-glucose (▲). Experimental
details are given in Section 3.8.1. The experiments were carried
out with 8 μg/mL β-glucosidase at 25°C.

of the enzyme with 50 mM cellobiose at pH 4.5 is shown in Figure 4.16(A). Glucose production proceeded rapidly for approximately 6 hours (360 min.). The rate of glucose production then decreased and the amount of glucose plateaued. This occurred even though there was still a large amount of cellobiose available. With 50 mM cellobiose, large amounts of trisaccharide were formed at early times. This is somewhat difficult to see because of crowding of the points. However, close examination shows that there is almost as much trisaccharide as glucose formed at early times. This shows that the reaction was almost entirely transglucosidic (since one glucose is released each time a trisaccharide forms). Further evidence that there is one to one production of glucose and trisaccharide at high cellobiose will be presented later. The presence of the trisaccharide was transient and its concentration decreased with time. Gentiobiose was also produced, but it only formed after a significant amount of glucose had accumulated. Again crowding at the early times makes this difficult to see but other experiments to be presented later will show this. The presence of gentiobiose was also transient.

The hydrolytic reaction with 5 mM was slower at pH 7 (Figure 4.15(B)) than at pH 4.5. Higher amounts of trisaccharide and gentiobiose were produced than was the case at pH 4.5. The rate of production of glucose with 50 mM cellobiose was also much slower at pH 7 than at pH 4.5 (Figure 4.16(B)). Again, nearly as much trisaccharide as glucose was produced, showing that the reaction was mainly transglucosidic (note that one glucose is produced each time that a trisaccharide is made). Gentiobiose was also produced but there was a delay and the gentiobiose was only produced after a significant amount of glucose had accumulated.

4.6.2.2 The initial rates of the hydrolytic and transglucosidic reactions as a function of the cellobiose concentration

The β-glucosidase was assayed with cellobiose concentrations ranging from 1 mM to 80 mM (pH 4.5). Aliquots of the assay mixture were removed at specific times at each concentration and treated as discussed in Section 3.8.1. The initial rates of the production of glucose and trisaccharide were determined from the slopes. The experiment was designed to determine the cellobiose concentration at which the transglucosidic reaction to form trisaccharide becomes important and to determine the initial rates of the transglucosidic reactions at each of the cellobiose concentrations. The results are shown in Figure 4.17. The trisaccharide rates started to increase between 5 and 15 mM. A drop in the rate of glucose hydrolysis also occurred at this point. The graph also shows that at high cellobiose concentration, transglucosylation is essentially the only reaction that occurs. The rate of glucose and trisaccharide production are approximately equal (at high cellobiose concentration) and since one glucose is produced for each trisaccharide this indicates that no hydrolysis occurs. It should be noted that these are initial rates and at later times the trisaccharide does degrade (see Figure 4.16(A)).

4.6.3 Nuclear magnetic resonance analyses of the β-glucosidase reactions

An NMR reaction profile of the β-glucosidase reaction with pNPGlc (75 mM) was generated as another way to analyze the formation and breakdown of the products over time (Figure 4.18). As expected, the results were very similar to those found by TLC and GLC. A much lower level of enzyme (4 μg/mL) was used and thus the reaction times were greater. pNPGe production could be detected at 2 hrs. The concentration of pNPGe reached
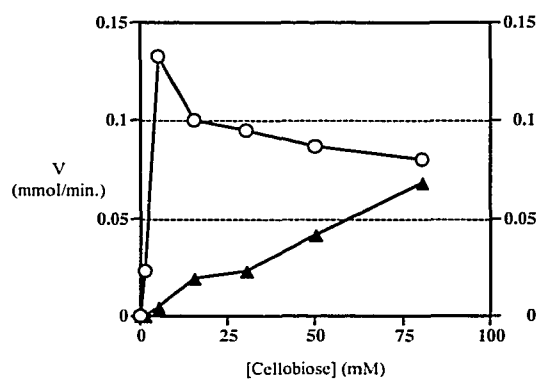
Figure 4.17. GLC analysis of the β-glucosidase initial reaction rates. The
reaction was monitored by following the production of glucose,
and gentiobiosyl-glucose as a function of time and determining
the rates from the slopes. The rates were determined at 6
cellobiose concentrations (25°C). These are plotted here as
Michaelis-Menten-like plots. The assays were each carried out
with 8 μg/mL β-glucosidase. (O) glucose production rates; (▲)
trisaccharide production rates. Experimental details regarding the
reactions and the preparation of the samples for GLC are given in
Section 3.8.1.

Figure 4.18. NMR ^{13}C time profile of pNPGlc degradation by β-
    glucosidase. The assay was done with 75 mM pNPGlc in
    the presence of 4 μg of enzyme. The reaction was followed
    to completion (74 hrs). The NMR spectrum for the reaction
    was measured and recorded when the enzyme was first
    added, and then at 2 hrs. After that a spectrum was recorded every
    4 hours. The spectra at 0, 2, 6, 18, 30, 42, and 74 are indicated in
    the Figure. The carbons from the glucose are referred to with a
    G and a number, the anomeric carbon is number one. Once the
    glucose is cleaved from the nitrophenol, the anomeric carbon can
    have either an α or β hydroxyl group at the anomeric carbon
    position. Therefore, the carbons are referred to as α or β. All sugar
    structures are illustrated in Appendix II. Shifts of the glucose
    peaks are indicated in addition to some of the peaks for
    gentiobiose and pNPGe (the transglucolytic peaks).The reactions
    were done in 150 mM acetate buffer at pH 4.5 and 25°C, using a
    10 mm probe.

free nitrophenol peaks

α,β forms of free glucose

β₁ α₁

74

β₁β₂

β₂

α₁

α₁α₃

β₄α₄

β₄α₅

74

42

gentiobiose

30

18

Free glucose

6

pNPGe

pNPGe

2

G₃   G₅   G₂   G₄   G₆

0

78   76   74   72   70   68   66   64   62   60

NMR ¹³C spectrum of pNPGlc

Nitrophenyl-group carbons

Glucose peaks

G₁

G₃   G₅G₂ G₄   G₆

0

160   150   140   130   120   110   100   90   80   70   60

¹³C ppm
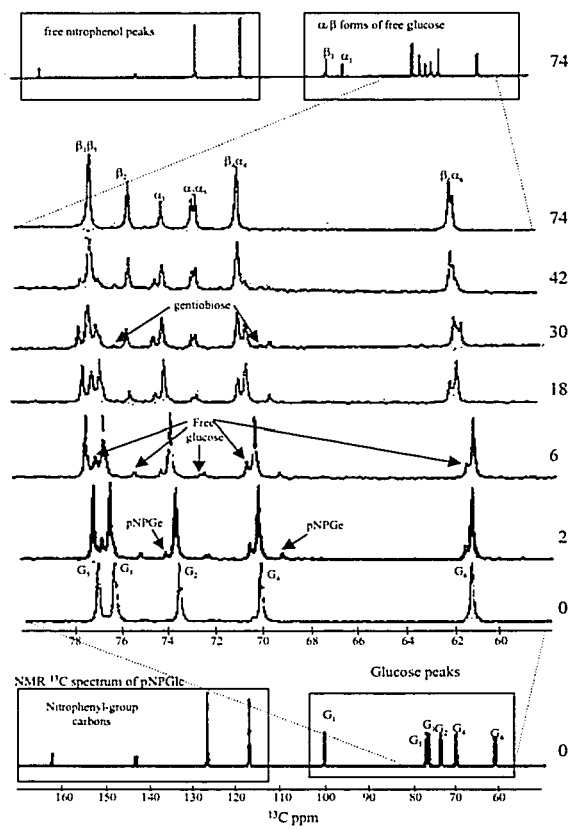
a plateau at 18 hr. The NMR data showed that gentiobiose appeared when the glucose

concentration was large (between 18 and 30 hr). The level of pNPGe decreased to

essentially zero at 54 hr. and of gentiobiose at 66 hr.

NMR analysis of the reaction of the enzyme with 80 mM cellobiose (Figure 4.19)

with low enzyme (2 μg/mL) showed that 4-O-gentiobiosyl glucose and glucose were

already present at 6 hrs. The formation of gentiobiose required the presence of free glucose

and it only began to appear later. 4-O-Gentiobiosyl glucose was completely degraded after

50 hrs whereas there were still small amounts of gentiobiose present even at 94 hrs. It was

very difficult to see the trisaccharide or gentiobiose peaks since they overlap so much with

glucose and cellobiose. Additionally, there is relatively much more substrate (cellobiose)

initially, and much more final product (glucose) and the smaller amounts of the transient

species were difficult to detect in comparison. For these reasons TLC and GLC were

utilized rather than NMR for the rest of the studies of the β-glucosidase reaction products.

Nevertheless, the results with NMR showed the same general trends as did the

chromatographic data and thus confirm those results.

4.7 pH profiles

The results obtained from the pH profiles were complex. In order to understand

them it is important to appreciate the proposed mechanism for the β-glucosidase (Figure

4.20). The initial part of the β-glucosidase reaction is always the same (colored green in

Figure 4.20), regardless of the glucosyl substrate involved. The substrate is bound to the

enzyme and the reducing end of the molecule is removed (k_2) leaving the non-reducing

end bound to the enzyme (probably covalently). Either hydrolysis (k_3) (purple) or

Figure 4.19. NMR ^{13}C time profile of cellobiose degradation by β-
glucosidase. The assay was done with 80 mM cellobiose in
the presence of 2 μg of enzyme. The reaction was followed
to completion (94 hrs). The NMR spectrum for the reaction
was measured and recorded when the enzyme was first
added, at 2 hrs and then every 4 hours. The spectra at 0, 10, 14,
18, 34, 58, and 94 hours are indicated in the Figure.
The time zero spectrum is essentially the spectrum for
cellobiose and the spectrum from 94 hours is essentially that of
glucose. The carbons from the non-reducing end glucose
are referred to with a G and a number. The reducing end glucose
can have either an α or β hydroxyl group at the anomeric carbon
position. Therefore, the carbons are referred to as α or β. The
anomeric carbon from free glucose can also be either α or β.
Therefore the free glucose peaks are labeled similar to the
reducing-end glucose of cellobiose. The sugar structures are
all illustrated in Appendix II. Important shifts of the cellobiose
peaks are indicated in addition to peaks produced by
gentiobiosyl-glucose and gentiobiose. The reactions were done in
150 mM acetate buffer at pH 4.5 and 25°C, using a 10 mm probe.
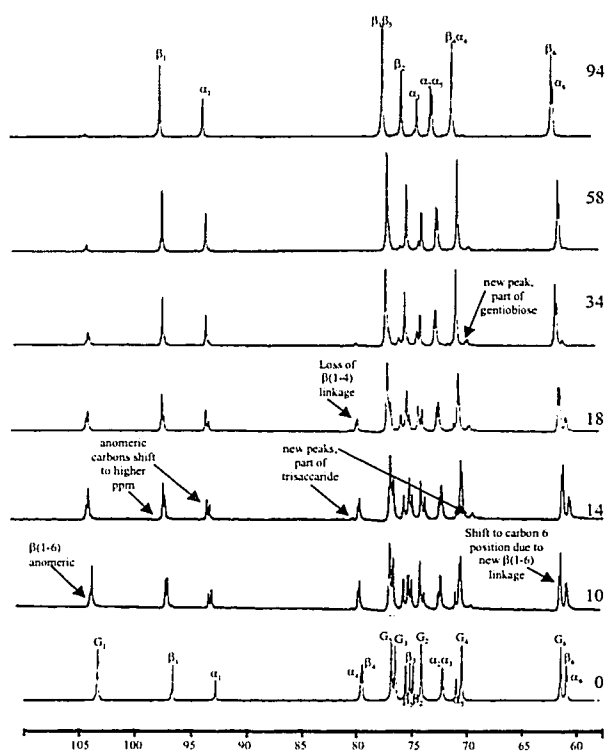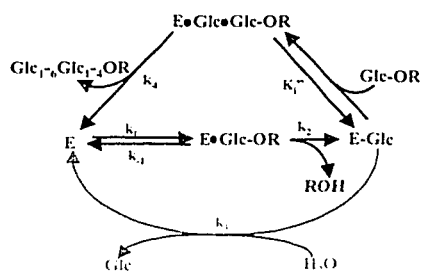
Figure 4.20. Proposed mechanism for the hydrolytic and transglucosidic

reactions of β-glucosidase. The enzyme is represented by E,

substrate is shown as Glc-OR, where the R group can be another

sugar or it can be a nitrophenyl group. $Glc_1-_6Glc_1-_4OR$ is the

transglucosidic product. Glc is glucose. Solid dots represent an

enzyme complex. The hyphen between the E and the Glc indicates

that this is probably a covalent intermediate. The common parts of

the mechanism are shown in green, the hydrolytic reaction is pink

and the transglucosidic reaction is blue. Production of gentiobiose

occurs as shown by the transglucosidic portion of the mechanism,

but in that case the acceptor is just glucose.

transglucosylis (k_4) (blue) can remove the remaining Glc. Transglucosylis occurs when a second molecule of substrate competes with a water molecule for a position in the active site and acts as an acceptor. Thus, the enzyme is forced through the transglucosidic pathway. If the rate of transglucosidic reaction (k_4) is slower than hydrolysis (k_3), the whole reaction will appear to be slowed down at high cellobiose concentration. Alternatively, if the transglucosidic reaction rate is faster than the hydrolysis step (k_3), the whole reaction will appear to speed up. The rates will tend to decrease or increase with increasing substrate concentration depending on the relative rates of the transglucosidic and hydrolytic reactions. Since water is only replaced by cellobiose at high substrate concentration, the straight lines of the Eadie-Hofstee plots at low substrate concentration represent hydrolysis.

Kinetic analyses were done at a series of pH between 3 and 7. Figure 4.21 shows Eadie-Hofstee plots of the purified enzyme on pNPGlc at 5 "representative" pH values. Only 5 are shown since the plot becomes too crowded when all are presented. The straight dashed lines at low substrate concentrations represent the hydrolytic portion of the reaction. These plots and the others done at the other pH values (that are not shown) indicate that the rates decreased at high substrate concentrations at each pH between 3 and 6 while the rates increased at high substrate concentrations at pHs above 6. These decreases and increases (shown as solid lines in Figure 4.21) are due to the transglucosidic reactions that result in the creation of pNPGe (see Figure 4.17).

The downward and upward curvatures observed on the Eadie-Hofstee plots were extrapolated to infinite substrate concentrations (the y-axis). Surprisingly, they gave

147
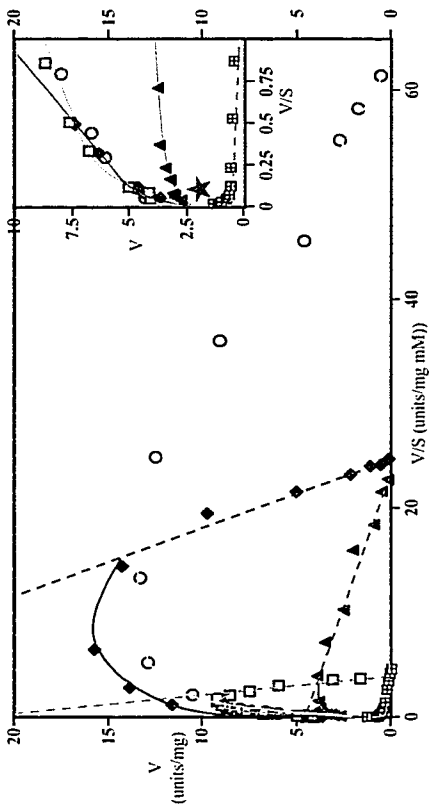
Figure 4.21. The β-glucosidase reaction with pNPGlc at 5 different pH's. Dotted lines illustrate the hydrolytic reaction and solid lines indicate the effect of the transglucosidic reactions. The different pH's are shown as: blue for pH 3 (□), purple for pH 4 (◆), peach for pH 5 (○), green for pH 6 (▲) and pink for pH 7 (⊞). The inset enlarges the portion of the plot where the extrapolated lines all converge on the y-axis. The convergence point, represented by the V_{max}(trans) values, is indicated by a red star. One unit is equal to 1 μmol of product per min.

148

nearly the same value (referred to as V_{max}(trans) values) at every pH tested (shown in greater detail - inset to Figure 4.21). These results suggested that the effects of the rate decreasing at low pH and increasing at high pH are probably due to a pH 'dependent' hydrolytic reaction component and a pH 'independent' transglucosidic reaction component. Therefore, at low pH the hydrolytic reaction (k_3) was faster than the transglucosidic reaction (k_4) and the plots curved downward at high substrate concentration. On the other hand, at high pH the hydrolytic reaction was slower than the transglucosidic reaction and the plots curved upward at high substrate concentration.

The points involved in the linear parts of the plots were obtained at substrate concentrations below 5 mM (see Figure 4.17). As already stated, these lines are representative of the hydrolytic reaction. These were utilized to determine the hydrolytic V_{max} and K_m values for the β-glucosidase. V_{max}(trans) values were determined using an SAS statistical program (see section 3.6.3.1).

Reactions with oNPGlc and cellobiose yielded plots (data not shown) that behaved in the same way as pNPGlc except that the curves converged at different points on the y-axis for each different substrate. pH profiles of the K_m and V_{max} values of pNPGlc, oNPGlc and cellobiose for the linear hydrolytic portions of the curves are shown in Figure 4.22. The pH optima for the V_{max} values were in the pH 4 - 4.5 range in each case (Figure 4.22(A)). The K_m values were high at low pH but decreased to much lower values at high pH (Figure 4.22(B)). The rates at infinite substrate concentration {V_{max}(trans) values}, calculated by SAS, for pNPGlc, oNPGlc and cellobiose varied by only a small amount (within error) with pH (over the range studied) (Figure 4.22(C)). As already mentioned,

Figure 4.22. Changes in β-glucosidase hydrolytic activity with pNPGlc($\bigcirc$),
oNPGlc ($\blacktriangle$), and cellobiose ($\square$). A) V_{max}; B) K_m; C) V_{max}(trans).
The red lines indicate the median value of the V_{max}(trans) for
each substrate. Assays were completed as described in Section
3.6.5 for pNPGlc and oNPGlc and in Section 3.6.7.1 for
cellobiose. All V_{max} and K_m values were calculated from the
linear portions of the Eadie-Hofstee plots. The V_{max}(trans) values
were determined by using the statistical analysis program, SAS.
Details regarding the SAS program can be found in Section
3.6.3.1.

A)

B)

C)

151
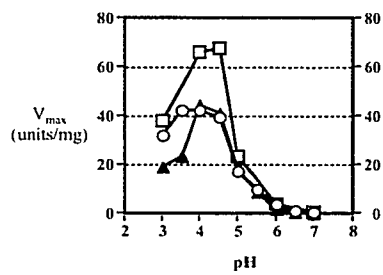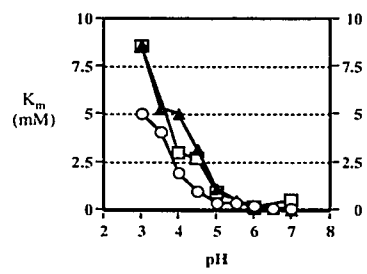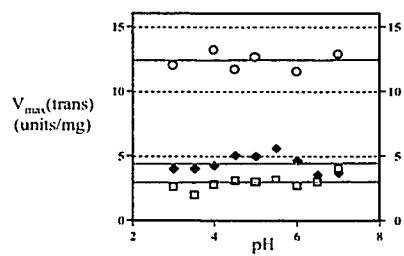
the actual plots for pNPGlc for some representative pH values are shown in Figure 4.21 (inset).

4.8 Assays with other disaccharides

       Gentiobiose, isotrehalose, sophorose, laminaribose, cellotriose, cellotetraose, cellopentaose, and 4-O-gentiobiosyl-glucose were all good substrates at pH 4.5. Each showed the decreased reaction rates at high substrate concentrations (pH 4.5), similar to the observations with oNPGlc, pNPGlc and cellobiose. The V_{max}, K_m and V_{max}/K_m values for the hydrolytic portion of all of the reactions are reported in Table 4.1. The V_{max}(trans) values for each of the substrates tested are also shown in Table 4.1. These V_{max} (trans) values were obtained using the non-linear line-fitting program, SAS. The V_{max}(trans) values calculated by this program were again similar to the extrapolated y-intercept values observed on the Eadie-Hofstee plots in every case. Relative to the studies with the natural sugar substrates, the nitrophenol substrates had very low V_{max}(trans) values. The V_{max}(trans) values for isotrehalose and laminaribose were relatively large. For each of the other substrates the V_{max}(trans) values were similar to the V_{max}(trans) value of cellobiose.

Table 4.1. Kinetic constants obtained for β-glucosidase at pH 4.5. (25°C) with a variety of glucosyl substrates.

Substrate*	V_{max} (units/mg)	V_{max} (trans) (units/mg)	K_m (mM)	V_{max}/K_m (units/mg mM)
pNPGlc	40	3	1.0	42
oNPGlc	41	5	3.2	13
cellobiose	68	12	2.7	25
cellotriose	56	12	0.5	108
cellotetraose	36	17	0.3	135
cellopentaose	34	10	0.5	75
isotrehalose	41	30	4.6	9
sophorose	35	13	3.9	9
laminaribose	57	44	1.5	37
gentiobiose	43	12	1.3	32
gentiobiosyl-glucose	46	10	1.5	30

*All structures of the β-glucosidase substrates in this thesis are illustrated in Appendix I.

4.8.1 3,4-Dinitrophenyl cellobiose (DNPC)

The enzyme reaction with DNPC was unusual (Figure 4.23). The rates increased with time (at each substrate concentration) instead of remaining constant (or decreasing). Also, at substrate concentrations higher than 0.5 mM, the overall rates decreased with increasing substrate concentration.
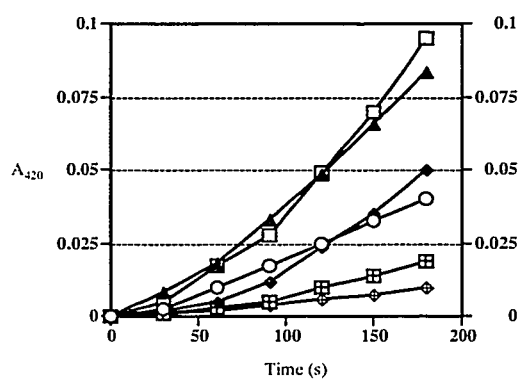
Figure 4.23. Reaction of β-glucosidase with 3,4-dinitrophenyl-β-D-
cellobioside (DNPC). The concentrations of substrate shown
are: 0.05 mM (O), 0.1 mM (▲), 0.5 mM (□), 1 mM (◆), 2 mM (⊞)
and 3 mM (⊕). All reactions were done at 25°C, pH 5. The enzyme
concentration used was 1.4 μg/mL.

4.9 Activity with metals

       Kinetic analyses were done in the presence of divalent metals and EDTA. Divalent metal ions (Mg^{2+}, Zn^{2+}, Mn^{2+} and Ca^{2+}) and EDTA did not significantly affect either the V_{max} or the K_m values of the enzyme (data not shown).

4.10 Anomeric preference

       The β-glucosidase reaction with cellobiose was monitored to determine the enzyme's anomeric preference at the reducing end of cellobiose. Additionally, the products were analyzed to determine if the glucose produced was in the β or α form. The reaction was carried out very rapidly with 5 mM cellobiose at pH 4.5 with an enzyme concentration of 40 µg/mL. Aliquots were removed at specific times and processed for GLC according to Section 3.8.1. Extreme care was taken to make sure the tubes in the lyophilizer were kept very cold in order to prevent as much mutarotation as possible. Figure 4.24 illustrates that the β-glucosidase utilized the β-anomer of cellobiose over the α-anomer and that a much greater amount of β-D-glucose was produced than of α-D-glucose.

4.11 Inhibitor studies

       The inhibitory effects of several sugars were determined (Table 4.2). Galactose, xylose, allose and mannose were essentially non-inhibitory ($K_i >> 100$). Gluconolactone had Ki values that were about 5-10 times lower than those of cellobiose, gentiobiose or glucose. This study shows that the enzyme is specific for glucosyl substrates.
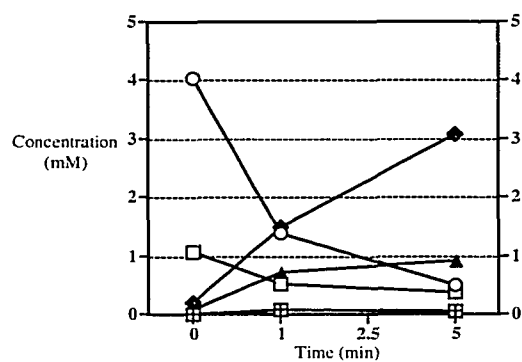
Figure 4.24. GLC analysis of β-glucosidase preference for the β-anomer of
   cellobiose. The cellobiose reaction was completed with a high
   enzyme concentration (40 μg/mL) and 5 mM cellobiose, at pH
   4.5, 25°C. The samples were very quickly frozen and the tubes
   were kept very cold until the lyophilization was complete.
   Experimental details of the silylation reactions are given in
   Section 3.8.1. (□) α-anomer of cellobiose, (O) β-anomer of
   cellobiose (reducing end), (▲) α-anomer of glucose, (♦) β-
   anomer of glucose, and (⊞) gentiobiose.

Table 4.2. K, values obtained by β-glucosidase in the presence of pNPGlc.

Inhibitor	K_i (mM) Sigma β-glucosidase
Glucose	3.1
Mannose	260
Allose	260
Galactose	670
Xylose	340
Gluconolactone	0.4
Cellobiose	3.3
Gentiobiose	4.7
Maltose	320
Cellotriose	0.4

4.12 Acceptor studies

The above data showed that cellobiose and glucose can compete with water and act as acceptors. Studies were therefore done to determine if other sugars or alcohols also acted as acceptors. TLC was used to follow the reactions with pNPGlc in the presence of D-glucose, D-galactose, D-xylose, isotrehalose, sophorose, lactose, sucrose, mannose and maltose. All except mannose and maltose produced transglucosidic products (data not shown). D-Xylose formed only a small amount of transglucosidic product at a very slow rate. When glucose was added, the product was gentiobiose. Several alcohols were also tested (glycerol, sorbitol, arabitol, methanol, ethanol, propanol, butanol and ethylene

glycol). They all formed transglucosidic products. Methanol, in particular, formed large amounts of transglucosidic adduct (Figure 4.25). The effects of the alcohols on the V_{max} and K_m values were also assayed (pNPGlc). The alcohols had only small affects on the rates of the reaction (data not shown).

4.13 Analysis of the glycoprotein nature of β-glucosidase

The NMR proton spectra of the β-glucosidase indicated the presence of sugars as part of the protein structure (Figure 4.26). The spectral peaks suggested the presence of α and/or β-D mannose, fucose, and N-acetylglucosamine. Thus the enzyme is probably a glycoprotein. An SDS-PAGE analysis of the PNGase F treated β-glucosidase is shown in Figure 4.27. The change in molecular weight of the native enzyme indicated that a significant amount of carbohydrate were normally present. It is noteworthy to mention that N-linked glycosylation occurs at sites with characteristic sequence motifs (Asn-X-Ser or Asn-X-Thr, X can be any residue except Pro). Fourteen of these sites can be identified in the β-glucosidase primary sequence.

The nature of the glycosidic bond was investigated by incubating the β-glucosidase with peptide N-glycosidase F (PNGase F), which cleaves N-linked oligosaccharides from glycoproteins. Generally, 3 types of N-linked oligosaccharides can be released from glycoproteins; I) oligomannose; II) complex; and III) hybrid (Glyko, 1994). Type I oligosaccharides (oligomannose) contain only mannose residues that are connected by α linkages to a core sugar structure. The core structure generally contains N-acetyl-glucosamine, mannose and sometimes fucose. Type II oligosaccharides (complex) contain a mixture of neutral and charged sugars (sialic acid) attached to a core
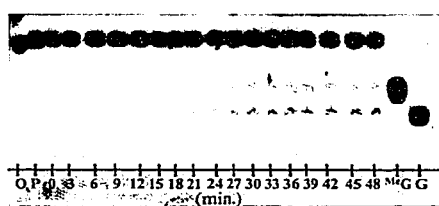
Figure 4.25. Thin layer chromatography of the β-glucosidase reaction with
pNPGlc (10 mM) in the presence of methanol (1 M). Enzyme
was added to the reaction mixture at zero time and an aliquot (1
μL) was immediately spotted onto the plate. Spotting was
repeated every 3 min. The chromatography solvent used was n-
butanol: ethyl acetate:2-propanol:acetic acid: water in the ratio of
1:3:2:1:1. Products were detected by dipping the dried plates in
2% sulphuric acid in methanol and charring. The standards used
on the plate were10 mM oNPGlc (O), 10 mM pNPGlc (P), 20
mM methyl-β-D-glucose (^{Me}G), and glucose 20 mM (G).
Reaction times are indicated in minutes along the bottom of the
plate.

Figure 4.26. Proton NMR spectrum of the β-glucosidase (Sigma). The region

indicative of sugars is shown (0-6ppm). Both α and β anomeric

protons can be identified in the 4.4 to 5.6 ppm range. In addition,

α (1-6) linkages between mannose units, which are common in N-

linked oligosaccharides, are also expected to produce a signal in

this range. The carbohydrate skeleton protons generally produce a

signal between 3.4 and 4.2 ppm. Protons from the CH_3 group of

fucose, a sugar commonly found on glycoproteins, were also

noticeable. The large peak at ~2 ppm is from the acetate of the

buffer. The large peak marked by the asterisk is the signal from

water. The smaller peaks surrounding the buffer peak probably

result from the acetyl group of N-acetyl-D-glucosamine (either α

or β since both sugars acetyl groups show proton shifts in the same

range). The peak close to 0 ppm is the standard DSS (2,2

dimethyl-2-silapentane-5-sulfonate))(5μL), which was used to

orientate the spectrum. The purified Sigma protein was

lyophilized and resuspended in 0.5 mL D_2O. The final

concentration of β-glucosidase was 10 mg/mL. The NMR proton

spectrum was recorded using a Bruker AMX-500

spectrophotometer operating at 500.139Mhz as described in

Section 3.9.3

acetate
CH₃ protons

anomeric sugar
protons

acetyl group
protons from
N-acetyl-(β/α)-D-
glucosamine

DSS
standard

sugar skeleton
protons

α   β

Fucose-CH₃
protons

6   5   4   ppm 3   2   1   0

219—
170—
116—
76—
53—

1   2   3   4   5
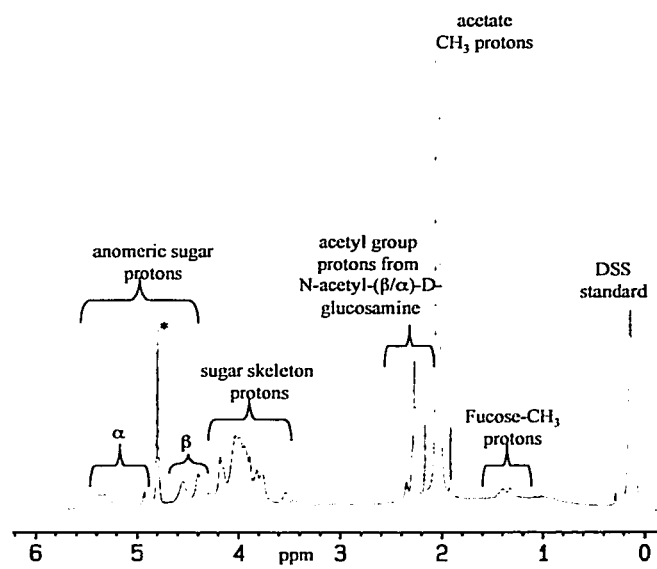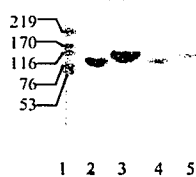
Figure 4.27. SDS-PAGE of the Sigma β-glucosidase in its deglycosylated and glycosylated forms. The gel lanes are shown on the bottom of the gel. The molecular weight marker is shown in lane 1 with corresponding molecular weights indicated on the left, sizes are shown in kDa. Lane 2 contains the β-glucosidase after a 16 hr treatment with peptide N-glycosidase F (PNGase F). Lane 3 contains the untreated β-glucosidase. Lane 4 , like lane 2, contains PNGaseF treated β-glucosidase. However, this treatment utilized 10-fold less β-glucosidase than in lane 2 to ensure the treatment went to completion. Lane 5 is similar to lane 3 but contains 10 fold less β-glucosidase.The Phast gel™ system with an 8-25 % gradient gel and Coomassie staining was used for SDS-PAGE analysis.

structure. Type III oligosaccharides (hybrid) are a mixture of mannose and other neutral sugars as well as sialic acid attached to a core structure.

To identify the sugar composition of the β-glucosidase N-linked oligosaccharides the residues liberated by PNGase F were labeled with the fluorophore ANTS (8-aminonaphthalene-1,3,6-trisulfonic acid). The labeled oligosaccharides were then separated on a FACE$^\circledR$ (Fluorophore-Assisted-Carbohydrate-Electrophoresis) gel. The resultant profiling gel is shown in Figure 4.28. The FACE$^\circledR$ system contains some reagents that are not stable and therefore it was necessary to analyze the profiling control. The profiling control, trypsin inhibitor, gave the expected pattern (compare lane 6 in Figure 4.28(A) with lane 2 of Figure 4.28(B)), which indicated that the reagents were working correctly. The oligosaccharides obtained from the β-glucosidase displayed 7 individual bands, which suggested the presence of 7 unique oligosaccharides (Figure 4.28 (A), lane 6). The quantitation control, which is equivalent to a particular concentration of oligosaccharide chains, is only accurate when the FACE$^\circledR$ imaging system is used. However, the control was included on the gel, to get a rough estimate of the number of oligosaccharides in each band of the profiling gel.

The oligosaccharides were analyzed on the profiling gel by comparing their mobility with the mobility of the glucose polymers in the oligosaccharide ladder. The degree of polymerization (DP) is defined by the glucose ladder and is equivalent to the number of glucose in each polymer (Glyko, 1994). Therefore, a tetrasaccharide band would have a DP of 4. A band that migrated between a trisaccharide and tetrasaccharide band would have a DP of 3.5. Unfortunately, the DP number is only relative to the
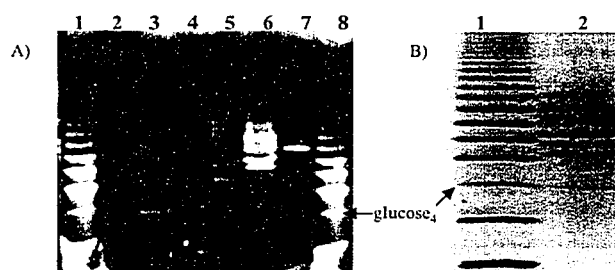
Figure 4.28. Analysis of the oligosaccharide component of the Sigma β-
glucosidase enzyme. A) The carbohydrate profiling gel obtained
after cleavage of the oligosaccharides from the β-glucosidase with
peptide N-glycosidase F. The gel shows a glucose ladder in lanes
1 and 8. The band that represents a polymer of 4 glucose is
indicated. A quantitation control that can be used with the FACE[x]
imaging system is present in lane 3. The N-linked
oligosaccharides isolated from 300 μg of β-glucosidase are shown
in lane 5. Oligosaccharides from trypsin inhibitor served as a
control and are shown in lane 6. Lane 7 contains the tracking dye,
and lanes 2 and 4 were empty. B) A gel reproduced from the
Glyko profiling manual (Glyko, 1994). Lane 1 contains the
glucose ladder and the second lane contains the trypsin inhibitor
control. This control shows the same pattern as lane 6 and
validates the results obtained in (A).

number of glucose monomers. Furthermore, if an oligosaccharide is charged, as occurs with the negatively charged sialic acid, it will have greater mobility than a neutral oligosaccharide of the same size. In addition to the complexities caused by the presence of charged sugars, not all neutral monosaccharides have the same mobility as glucose. The effects of different sugars on the mobility of an oligosaccharide, relative to glucose, are shown in Table 4.3.

As stated above, analysis of the profiling gel shows that there are 7 unique oligosaccharide bands. The mobility of each band was measured and compared with the mobility of the bands in the glucose polymer ladder. The calculated DP values (from the largest oligosaccharide to the smallest) were 9.7, 9.0, 8.3, 7.9, 5.8, 5.3, and 4.4. However, these values only give the number of equivalent glucose residues in the oligosaccharide and not the identity of the sugars. Only when the data obtained from both profiling gel and sequencing gel is put together can the composition of the different oligosaccharides be obtained.

In order to determine the identity and order of the sugars present on the β-glucosidase, each of the bands visualized on the profiling gel was excised and digested with the exoglycosidases in Table 4.4.

The digestion products for each of the bands were then electrophoresed on an N-linked oligosaccharide sequencing gel. The sequencing gel for one of the bands on the profiling gel is shown in Figure 4.29. It was the top band and had a DP of 9.7.

Table 4.3. The effect of monosaccharides, other than glucose, on the DP (degree of polymerization) number of an oligosaccharide.

Monosaccharide	Mobility Shift (in DP units*)
Neuraminic acid	-1
Galactose	+1
N-acetyl glucosamine	+0.75
Mannose	+0.75
Fucose	+0.6
N-acetyl glucosamine (bisecting)	+0.5

* DP units are measured relative to a ladder of glucose polymers.

Table 4.4. Carbohydrate degrading enzymes utilized for N-linked sequencing of the β-glucosidase oligosaccharides.

Enzyme Name	Specificity	Glyko Abbreviation
Neuraminidase	(α2-3,6,8) linked N-acetylneuraminic acid	NANase III
β-Galactosidase	(β1-4) linked galactose	GALase III
β-N-Acetylhexosaminidase	(β1-2,3,4,6) N-acetylglucosamine	HEXase III
α-Mannosidase	(α1-2,3,6)linked mannose	MANase II
α-Mannosidase	(α1-6) linked mannose	MANase VI

Only the addition of MANase II and MANase VI resulted in digestion of the DP-9.7-oligosaccharide to a core structure. The core structure is not digested by the exoglycosidases because it contains mannose(β1-4)-N-acetyl-glucosamine (β1-4)[fucose(α1-6)]-N-acetyl glucosamine. The mannose of this structure can not be cleaved by MANase II or MANase VI because neither enzyme is capable of cleavage of a β(1-4) bond (Glyko, 1994). The digested oligosaccharide migrated to the same position as the
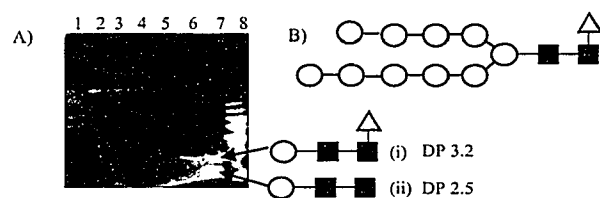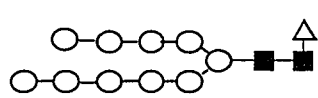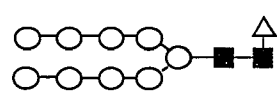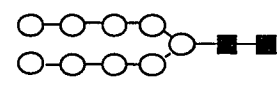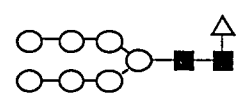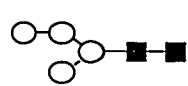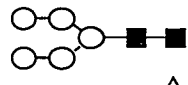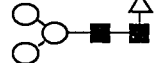
Figure 4.29. A) Carbohydrate sequencing gel resulting from reaction of the
DP 10 oligosaccharide from the carbohydrate profiling gel with a
variety of exoglycosidases. The lanes are 1) the β-glucosidase
oligosaccharide profile, 2) the purified oligosaccharide band from
the profiling gel, no enzymes added 3) digestion with NANase III
(removes α2-3,6,8 linked N-acetylneuraminic acid if present), 4)
digestion with NANase III and GALase III (removes the above in
addition to (β1-4) linked galactose), 5) digestion with NANase
III, GALase III, and HEXase III (in addition to above digestion of
β1-2,3,4,6 linked N-acetyl glucosamine residues occurs), 6)
digestion with NANase III, GALase III, HEXase III, MANase II
and MANase VI (in addition to previous digestion, the α1-2,3,6
linked mannose and α1-6 linked mannose are digested) 7) the
core standard containing i) Mannose (β1-4) N-acetyl glucosamine
(β1-4)[fucose(α1-6)]-N-acetyl glucosamine or ii) Mannose (β1-
4) N-acetyl glucosamine (β1-4)-N-acetyl glucosamine, lane 8 is
the glucose polymer ladder. B) A schematic representation of the
possible structure of the isolated oligosaccharide. Mannose is
represented by (O); N-acetyl glucosamine (■); and fucose (△).
The nature of the linkages between the mannose residues could
not be determined under the conditions of the sequencing
reactions.

fucosylated core standard. The core structure has a DP of 3.2 (Glyko, 1994). Therefore,

the remaining DP to be accounted for is 6.5. The exoglycosidase results in Figure 4.29

suggest that the oligosaccharide only contains mannose. From Table 4.3 mannose has a

mobility equal to 0.75 DP, therefore, the oligosaccharide must contain 9 mannose. Each

of the isolated oligosaccharides displayed the same type of result. They were only

digested when the MANase II and MANase VI enzymes were present. Since the profiling

gel suggested the sugars differed by 1 or 2 sugar residues, the only difference between the

oligosaccharides isolated was the number of mannose residues present. Two of the

oligosaccharides aligned with the lower band of the core structure standard (DP-2.5).

This indicated that the core structure of those oligosaccharides was not fucosylated.

However, they were still only digested by MANase II and MANase VI. The probable

structures for the isolated oligosaccharides are shown in Figure 4.30. Unfortunately

digestion by MANase II and MANase VI provides almost no information about the

linkages between the mannose residues except that they are not α1-4.

Figure 4.30. Probable structures for the oligosaccharides released from the
Sigma β-glucosidase. Digestion of the purified oligosaccharide
chains with exo-glycosidases indicates that the oligosaccharides
are mainly composed of mannose. The linkage between the
mannose residues could not be determined from the data since the
MANase enzymes cleaved α 1-2,3,and 6 linkages. The
oligosaccharides were digested down to core structures and
showed the same mobility as the core standards that were
composed of Man(β1-4)GlcNAc(β1-4)[Fucα1-6)]GlcNAc or
Man(β1-4)GlcNAc(β1-4)GlcNAc, where Man is mannose,
GlcNAc is N-acetyl glucosamine and Fuc is fucose. The number
of mannose in each oligosaccharide is calculated from the change
in the degrees of polymerization (DP units-see Table 4.3). DP(ex)
are the experimental DP values obtained from the profiling gel.
DP (calc) are values determined using Table 4.3. Sugars are
represented by (O) mannose; (■) N-acetyl glucosamine; and
(△) for fucose.

	DP(ex)	DP(calc)
	9.7	9.9
	9.0	9.2
	8.3	8.5
	7.9	7.7
	5.7	5.5
	5.3	5.5
	4.4	4.7

5. Results

Part II:

Characterization of Recombinant *Aspergillus niger* β-Glucosidase

The gene for the β-glucosidase from *Aspergillus niger* that was obtained from Dr. Shoseyov's group at the Hebrew University in Jerusalem was inserted into a plasmid and expressed in *Pichia pastoris*. This "recombinant β-glucosidase" was used to compare the unsubstituted and substituted β-glucosidases (see Chapter 6). The properties of the recombinant β-glucosidase are detailed here and compared to the properties of the Sigma cellulase powder β-glucosidase that are described in Chapter 4.

5.1 Comparison of sequences, sizes and kinetics

CNBr cleavage of the Sigma β-glucosidase was carried out. Several candidate CNBr bands were obtained on the gel. One was chosen and sequenced by Dr. McKay, University of Calgary. The chosen band produced a clear sequence, MAAYYKVGRDRLW, which matched exactly with residues 319-331 in the β-glucosidase sequence provided by Dr. Shoseyov. The sequence alignments of the β-glucosidases from family 3 indicate that this is not a highly conserved region (Figure 5.1) of the protein, except in *Aspergillus* species. This indicates that the β-glucosidase purified from the Sigma mixture is the same as the β-glucosidase that is being studied by Dr. Shoseyov. The possibility that there could be such a good match by chance alone (especially since both enzymes were obtained from *A. niger*, is very small). Additionally, the native and subunit protein sizes of the two enzymes were essentially identical (Figure 5.2). The V_{max} and K_m values of the enzymes as well as the unique profiles of the Eadie-

171

Figure 5.1. A portion of an *A. niger* β-glucosidase (cyan) alignment with fungal glycosyl hydrolase family 3 enzymes. The peptide sequence obtained from the *A. niger* protein purified from the Sigma cellulase mixture is indicated in the blue box, and highlighted in cyan. The main similarities occur with other *Aspergillus* sp β-glucosidases. The microorganism from which each sequence was obtained is indicated to the left with the accession number for the sequence. Bright red indicates very conserved residues, pink represents partially conserved residues, pale pink refers to slightly conserved residues and no shading indicates no sequence conservation. A Blosum 55 matrix was used for the alignment, with pairwise grouping parameters of: 1 for k-tuple (word) size, 20 for the maximum gap length, a gap penalty of 4 and the number of top diagonals to use was 5. The multiple sequence alignment parameters were: 9 for gap creation and 8 for gap extension.

A)   B)

Figure 5.2. A) SDS-PAGE of the β-glucosidase isolated from the Sigma

cellulase mixture (lane 2) and the β-glucosidase isolated by Dr.

Shoseyov's group (lane 3). The first lane is the molecular weight

marker. B) Overloaded Native PAGE of the two β-glucosidases.

Lane 1 is the enzyme provided by Dr. Shoseyov's group, while lane

2 is the Sigma enzyme. Molecular weights of the marker proteins

are indicated in each case. The proteins used in lane 1 of the gel in

(A) were: myosin (212 kDa), α2-macroglobulin (170 kDa), β-

galactosidase (116 kDa), transferrin (76 kDa) and glutamic

dehydrogenase (53 kDa). The proteins used in lane 3 of the gel in

(B) were thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232

kDa), lactate dehydrogenase (140 kDa) and albumin (66 kDa).

Hofstee plots with both pNPGlc (Figure 5.3) and cellobiose (not shown) at pH 4.5 provide

further evidence that the enzymes are identical.

5.2 Production of the pPicZAα plasmid with the β-glucosidase gene

The pPICZAα:BGL plasmid was created as described in section 3.11.2.

Sequencing of the appropriate parts of the plasmid was done. The sequencing showed that

the β-glucosidase gene was successfully inserted into the pPicZAα plasmid and that the

gene was in frame with the α-secretion sequence and the His-tag sequence.

5.3 Transformation into *Pichia pastoris* and expression of β-glucosidase

The purified plasmid pPICZAα:BGL was transformed into the X-33, GS115, and

KM71 strains of *Pichia pastoris*. Controls containing no DNA and plasmid DNA only,

were also done. After 4 days the controls displayed no growth (as was expected). The *P.*

pastoris wild type strain, X-33, showed the most transformants, followed by the KM71

strain and finally the GS115 strain.

The transformants were analyzed to determine their Mut phenotype. Small sections

of the transformants were streaked onto MMH plates and then onto MDH plates. A Mut⁺

and a Mut⁻ control were also streaked onto the plates for comparison. The KM71

transformants were also streaked even though they would all be expected to be Mut⁻. If the

strains are Mut⁻ they will grow equally well on both types of plates. If the strains are Mut⁺

they will grow well on the MDH plate but poorly on the MMH plate. As expected, all of

the KM71 transformants grew normally on the MDH plates but grew poorly on the MMH

plates, indicating that they were Mut⁻. The X-33 transformants and the GS115

transformants grew normally on both types of plates indicating that they were all Mut⁺.
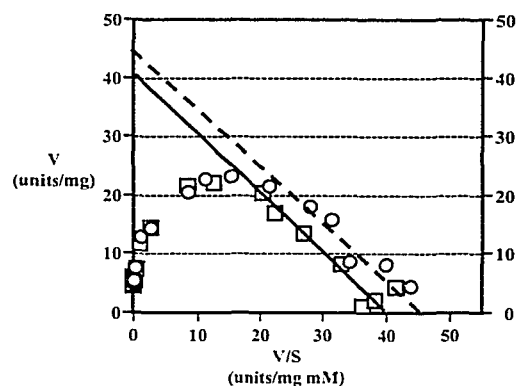
Figure 5.3. A comparison of the β-glucosidase activities of the two enzymes with pNPGlc at pH 4.5, 25°C. The β-glucosidase isolated from the Sigma cellulase mixture is represented by (□), while the β-glucosidase that Dr. Shoseyov's group isolated is represented by (O). The enzyme provided by Dr. Shoseyov was sent as a lyophilisate that was resuspended into an appropriate amount of buffer so that the same enzyme concentration was used for the assays. The enzyme concentration used was 1.4 μg/mL. One unit is equal to one μmol of product per min.

Expression of the β-glucosidase gene was examined for six X-33 transformants,
two GS115 transformants, and four KM71 transformants. Controls for secreted expression
and background expression were also set up. The β-glucosidase gene was expressed in all
of the transformants. However, the X-33 strains displayed the greatest expression in the
fastest time. The GS115 strains expressed less β-glucosidase. The KM71 strain expressed
more β-glucosidase than the GS115 strains but less than the X-33 strain. The control
strains indicated that the secretion signal was working and that no β-glucosidase was
endogenously secreted by the *P. pastoris* strains. The expression results were also
visualized by SDS-PAGE. All of the transformants displayed bands of the expected size
for the β-glucosidase protein. The GS115 protein bands were very faint, indicating that the
protein was more poorly expressed in those strains (Figure 3.12). The background control
strain showed no β-glucosidase band, even when the gel was silver stained (not shown).

5.4 Purification of recombinant β-glucosidase

The β-glucosidase was easily purified from the BMMY media (Section 3.16.4).
SDS-PAGE indicated that very few other proteins were secreted by *P. pastoris* and the
protein was usually found to be pure after elution from the Sephacryl S-300 column. In
some cases, the concentrated protein was yellowish in color. In those cases the yellow
material was removed by application of the sample to the FPLC Superose 6 and 12
columns (connected in series). After concentration, the collected fractions were analyzed
by SDS-PAGE and determined to be pure (>98%) by staining with Coomassie Blue and/or
silver stain.

5.4.1 Molecular weight of recombinant β-glucosidase

The recombinant β-glucosidase migrated the same distance as the Sigma enzyme on SDS and on native PAGE (data not shown). It is also eluted at the same volume as the Sigma enzyme from the size exclusion columns (FPLC). Light scattering was also utilized to determine the size of the recombinant enzyme (Figure 5.4). Light scattering in conjunction with the SDS PAGE data suggested that the recombinant enzyme is a dimer (as was also the case with the Sigma enzyme). The light scattering radius of 8-9 nm corresponds to a molecular weight of 196 kDa, which is approximately twice that of the SDS PAGE molecular weight. The light scattering results also indicated that there is aggregation. Aggregation is known to occur readily when recombinant enzymes are concentrated. In this case the aggregation seems to be reversible. When size exclusion was used for the purifications there was no evidence of aggregation. In those studies the protein was usually not concentrated beyond 1 mg/mL. In addition there was no evidence of aggregation on native PAGE gels. Thus, aggregation does not seem to be a problem at low protein concentration. The solutions used to assay the enzyme had very low protein concentration (0.002 to 0.04 mg/mL) and thus it is highly unlikely that the proteins were aggregated during the assays.

5.5 Temperature stability of the recombinant β-glucosidase

The thermal stability of the recombinant enzyme was assayed (data not shown) in the same manner as for the Sigma enzyme (Section 4.6.6). The two enzymes had very similar stabilities (data not shown).

Distribution (%)

50
45
40
35
30
25
20
15
10
5
0

0.01   0.1   1   10   100   1000   10000

Hydrodynamic Radius (R$_h$) (nm)
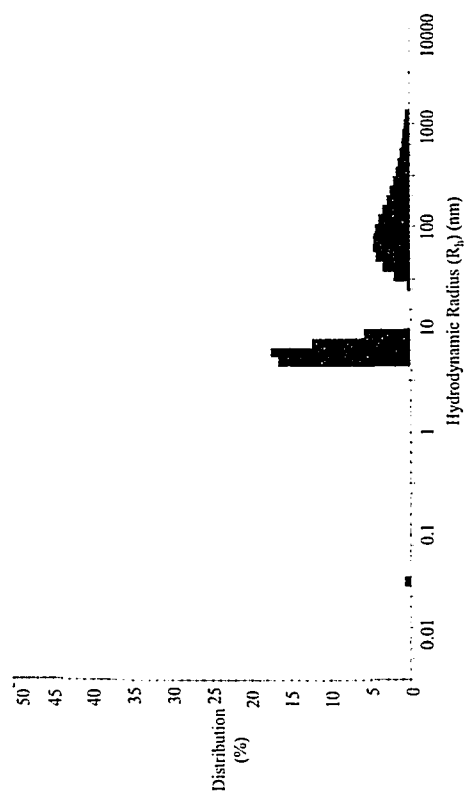
Figure 5.4 Distribution of macromolecule size in a 2.6 mg/mL solution of β-glucosidase (recombinant enzyme expressed in *Pichia pastoris*). Macromolecule sizes are shown in terms of the hydrodynamic radii. The small peak that occurs at an approximate radius of 0.015 nm results from the sodium acetate buffer. The large peak between 6-10 nm results from the protein. The broad "peak" probably results from protein aggregation.

5.6 Fluorescence characterization of the recombinant β-glucosidase

The emission spectra obtained for the recombinant enzyme at excitation wavelengths of 275, 285 and 295 nm were all similar in shape to the native enzyme (Figure 5.5(A), (B), and (C), respectively). The excitation spectrum for the recombinant enzyme that was obtained at an emission wavelength of 339 nm was also almost identical to the Sigma enzyme spectrum (Figure 5.5(D)).

5.7 Kinetic characterization of the recombinant β–glucosidase

5.7.1 pNPGlc pH profiles

The pH profiles for V_{max}, K_m, and V_{max}/K_m for pNPGlc were obtained (Figure 5.6). The trends illustrated in the figure are similar to those of the Sigma enzyme. The main exception was that the recombinant β-glucosidase had V_{max} values that were somewhat higher than those of the Sigma enzyme.

5.7.2 Glucosyl substrates

The V_{max} values for the action of the recombinant enzyme on glucosyl substrates were somewhat larger than those for the Sigma enzyme (Table 5.1) in each case. The increased activity may have resulted from the fact that the recombinant enzyme was purified quickly and used immediately while the Sigma enzyme required more strenuous purification. In addition, the powder was presumably much older as it was prepared and processed by Sigma. The majority of the K_m values obtained were very similar for the two enzymes. The largest difference in K_m values occurred when gentiobiose was used as a substrate. The Km of gentiobiose with the recombinant enzyme was 2 times higher than that of the Sigma enzyme.
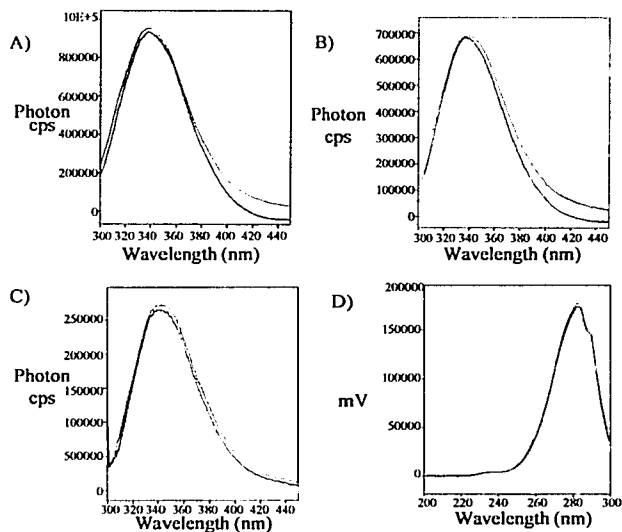
Figure 5.5. Fluorescence emission and excitation spectra of β-glucosidase from the Sigma mixture (blue) and the recombinant enzyme (pink). Intensity is measured as photon counts per second (cps) for the emission spectra and in millivolts (mV) for the excitation spectrum. Emission spectra were collected at excitation wavelengths of: A) 275 nm, B) 285 nm, and C) 295 nm. The excitation spectrum (D) was obtained at an emission wavelength of 339 nm. Both enzyme samples were diluted to 17 μg/mL with acetate buffer (pH 4.5, 150 mM). Spectra were collected at 25°C. A spectrum of the buffer alone was subtracted from each protein spectrum.

180

Figure 5.6. Effect of pH (pNPGlc) on the recombinant (O) and the Sigma
(□) β-glucosidases. A) V_{max}; B) K_m; C) V_{max}/K_m. The assays at
the various pHs were carried out as described in section 3.6.5.
The enzyme concentration used was 1.4 μg/mL. All assays were
completed at 25°C in 150 mM sodium acetate/sodium
phosphate buffer. One unit is equal to one μmol of product per
min.

A)

B)

C)

182
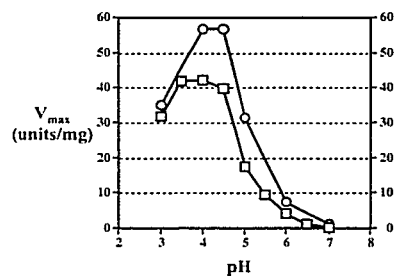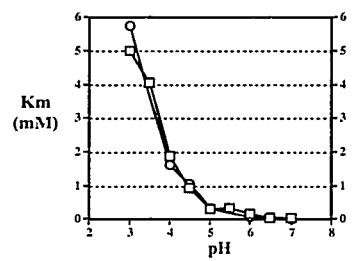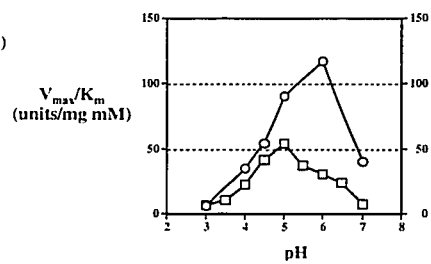
Table 5.1 A comparison of kinetic constants of the ß-glucosidase isolated from the Sigma cellulase powder (WT(S)) and the recombinant (WT(P)) β-glucosidases.

Substrate	WT(S)-V_{max}[a]	WT(S)-K_m[b]	WT(P)-V_{max}[a]	WT(P)-K_m[b]
pNPGlc	40	1.0	57	1.0
oNPGlc	41	3.2	60	3.0
Cellobiose	68	2.7	78	2.9
Gentiobiose	43	1.3	46	2.2
Cellotriose	56	0.5	94	0.7
Cellotetraose	36	0.3	66	0.6
Cellopentaose	34	0.4	40	0.4

[a] V_{max} values are calculated in units of µmol/mg/min.
[b] K_m values are calculated in units of mM.

5.7.3 Inhibitor Studies

The recombinant β-glucosidase was also competitively inhibited to about the same extent as the Sigma enzyme by each inhibitor tested (Table 5.2). Glucose was a good inhibitor, while its epimers: galactose, mannose, and allose were not. Gluconolactone, cellobiose, gentiobiose and cellotriose were all good inhibitors (particularly gluconolactone and cellotriose). Maltose and xylose were poor inhibitors. It should be stated that quite a lot of error occurs when determining K_i values, especially when the inhibition is poor (large K_i values). Despite this the K_i values for the two enzymes with the various inhibitors were quite similar.

Table 5.2. A comparison of inhibitor dissociation constants (K_i) for the Sigma and the recombinant β-glucosidases.

Inhibitor	K_i (mM) Sigma β-glucosidase	K_i (mM) Recombinant β-glucosidase
Glucose	3	3
Mannose	260	280
Allose	260	220
Galactose	670	670
Xylose	340	290
Gluconolactone	0.4	0.4
Cellobiose	3.3	3.4
Gentiobiose	4.7	4.8
Maltose	320	270
Cellotriose	0.4	0.4

5.8 Analysis of substrate breakdown

5.8.1 Thin layer chromatography

TLC showed that the recombinant enzyme degraded 5 mM cellobiose to glucose without producing detectable amounts of gentiobiosyl-glucose or gentiobiose (Figure 5.7(A)). When the enzyme was reacted with 50 mM cellobiose a relatively large amount of gentiobiosyl-glucose was produced (Figure 5.7(B)) even at early times. It could be detected after 2-4 min. Gentiobiose was not separated from the cellobiose on the TLC plate and was probably masked by the large spots for cellobiose. These results are similar to those found with the β-glucosidase from the Sigma cellulose powder.

A)

B)

Figure 5.7. Thin layer chromatography of the recombinant β-glucosidase
reaction with: A) 5 mM cellobiose and B) 50 mM cellobiose.
Enzyme (8 μg/mL) was added to the reaction mixture at zero time
and an aliquot (1 μL) was immediately spotted onto the plate.
Spotting was repeated at the times indicated in minutes on the
bottom of the plate. The chromatography solvent used was n-
butanol:ethyl acetate:2-propanol:acetic acid:water in the ratio of
1:3:2:1:1. Products were detected by dipping the dried plates in 2%
sulphuric acid in methanol and charring. The standards (1μL) were:
10 mM glucose (G), 10 mM cellobiose (Ce), 10 mM gentiobiose
(Ge), and 5 mM gentiobiosyl-glucose (GG).

The recombinant enzyme was also reacted with pNPGlc at 5 and 50 mM and analyzed by TLC (Figure 5.8). The enzyme reacted as expected and produced only glucose and nitrophenol at 5 mM (the nitrophenol does not char and so is not seen on these plates). Larger sugars (pNPGe) were observed at 50 mM pNPGlc. Gentiobiose was not visible at either high or low pNPGlc concentrations but that is probably because the reaction did not proceed long enough to produce large enough amounts of glucose for it to act as an acceptor.

5.8.2 Gas-liquid chromatography

The production of glucose, gentiobiose and gentiobiosyl-glucose from cellobiose by the recombinant β-glucosidase was also analyzed utilizing GLC. Only a small amount of enzyme was added and the reactions were allowed to proceed for a long time in order to follow the reactions better. At pH 4.5, the production of glucose from 5 mM cellobiose occurred in a similar fashion for the two enzymes (Figure 5.9). Very little gentiobiose or trisaccharide were produced by either the recombinant or the Sigma enzyme with 5 mM cellobiose (amounts were barely above background). When the enzyme activity was monitored in the presence of 50 mM cellobiose at pH 4.5 the results were also similar to those of the Sigma enzyme (Figure 5.10). The initial rate of glucose production at 50mM cellobiose was smaller than at 5 mM (as expected). Glucose accumulation plateaued at about 45 mM for the Sigma enzyme and at about 55 mM for the recombinant enzyme. The amount of gentiobiose increased and then decreased as the enzymes degraded it (Figure 5.10(B)). In the case of both enzymes there was a lag before the gentiobiose production rate became significant. This is because the amount of glucose had to be built up to a large
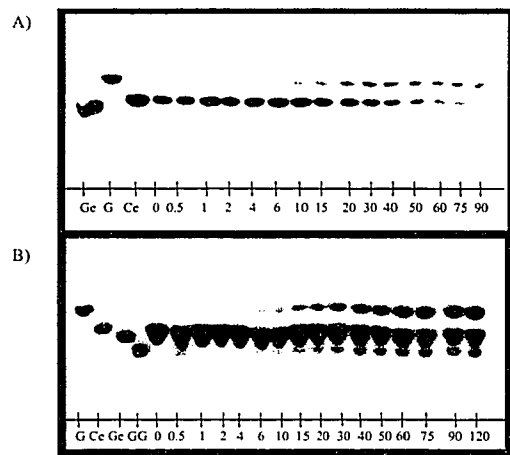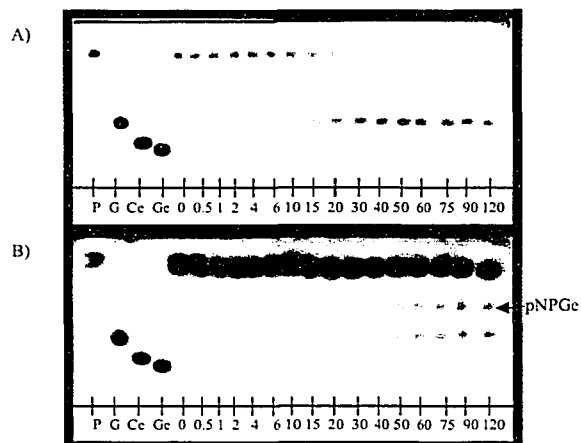
Figure 5.8. Thin layer chromatography of the recombinant β-glucosidase
reaction with: A) 5 mM pNPGlc and B) 50 mM pNPGlc. Enzyme
(1.4 µg/mL) was added to the reaction mixture at zero time and an
aliquot (1 µL) was immediately spotted onto the plate. Spotting was
repeated at the times indicated in minutes on the bottom of the
plate. The chromatography solvent used was n-butanol:ethyl
acetate:2-propanol:acetic acid:water in the ratio of 1:3:2:1:1.
Products were detected by dipping the dried plates in 2% sulphuric
acid in methanol and charring. The standards (1µL) were: 5 mM
pNPGlc (P), 10 mM glucose (G), 10 mM cellobiose (Ce), and 10
mM gentiobiose (Ge). pNPGe is p-nitrophenyl-gentiobiose.

Figure 5.9. Analysis of the glucose produced by the reactions of the β-glucosidase from Sigma and the recombinant β-glucosidase with 5 mM cellobiose. Gentiobiose and gentiobiosyl-glucose were also monitored but the amounts present were only slightly above background and are not shown. The symbol used to represent the Sigma enzyme is (□) and the recombinant enzyme is (O). The concentration of the recombinant enzyme was 2 μg/mL and a concentration for the Sigma enzyme was chosen that resulted in approximately the same rate of glucose production. The complete experimental details are given in Section 3.8.1.

188

Figure 5.10. Analysis of the products of the reactions of the Sigma β-
glucosidase and the recombinant β-glucosidase with 50 mM
cellobiose at pH 4.5. The products that were monitored were: A)
glucose, B) gentiobiose, and C) gentiobiosyl-glucose. The
symbol used to represent the Sigma enzyme is (□) and the
recombinant enzyme is (O). The enzyme concentration of the
recombinant enzyme was 2 μg/mL and a concentration for the
Sigma enzyme was chosen that resulted in approximately the
same rate of glucose production as for the recombinant enzyme.
The activities of the two enzymes used was actually a little
different and this is seen by the different profiles for gentiobiose
and trisaccharide. The complete experimental details are given in
Section 3.8.1.

A)

[Glucose]
mM

Time (min)

B)

[Gentiobiose]
mM

Time (min)

C)

[Gentiobiosyl-
Glucose]
mM

Time (min)

190
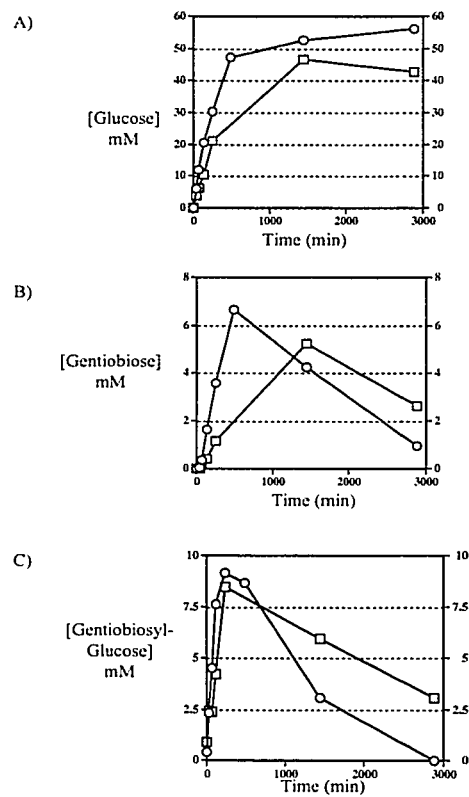
enough amount before synthesis of gentiobiose began. The maximum amount of gentiobiose occurred at different times for the two enzymes because the rates of glucose production were not fully matched. The production of the trisaccharide also rapidly increased and then decreased. There was, however, no lag in its production rate. The build up and loss of the trisaccharide occurred in a similar fashion for both enzymes although again, because the enzyme activity was not fully matched for the two enzymes, there were different degradation rates. TLC showed that the trisaccharide produced was also utilized by the enzyme to produce longer glucose chains. These larger sugars would not be visible using the GLC method as sugars larger than trisaccharides would require temperatures that would leach the column coatings.

The production of the three sugars was also analyzed at pH 7. With 5 mM cellobiose (Figure 5.11) the hydrolysis to glucose was slow by both enzymes. Very little gentiobiose or trisaccharide were produced. When 50 mM cellobiose (Figure 5.12) was used, the rate of glucose production at pH 7 was more rapid than at 5 mM cellobiose (as expected). The rate at which the trisaccharides and gentiobiose were produced were similar for the two enzymes. The trisaccharide comprised a large proportion of the product when it is considered that one glucose is produced every time that a trisaccharide forms. Again, there was a delay in gentiobiose production that was dependent on an initial build-up of glucose. The delay was very noticeable in this case.

The recombinant enzyme was also analyzed by GLC with pNPGlc (Figure 5.13). When 5 mM pNPGlc was used, the main reaction was hydrolysis and the amount of glucose accumulated until it plateaued at about 5 mM. When the concentration of pNPGlc
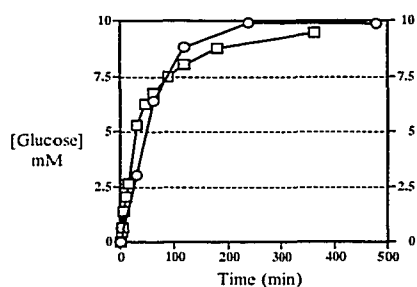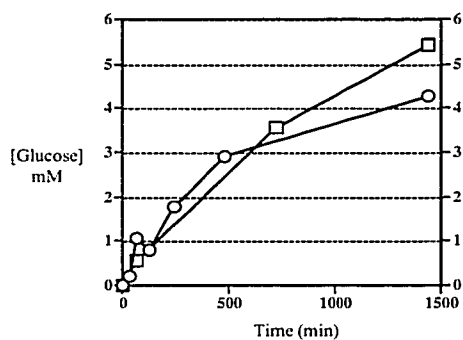
Figure 5.11. Analysis of the glucose produced by the reaction of the β-
glucosidase from Sigma and the recombinant β-glucosidase with
5 mM cellobiose at pH 7. Gentiobiose and gentiobiosyl-glucose
were also monitored but the amounts present were only slightly
above background and are not shown. The symbol used to
represent the Sigma enzyme is (□) and the recombinant enzyme
is (o). The concentration of the recombinant enzyme was 2
μg/mL and a concentration for the Sigma enzyme was chosen
that resulted in approximately the same rate of glucose
production. The complete experimental details are given in
Section 3.8.1.

Figure 5.12. Analysis of the products from the reactions of the β- glucosidase
from Sigma and the recombinant β-glucosidase with 50 mM
cellobiose at pH 7. The products that were monitored were: A)
glucose, B) gentiobiose, and C) gentiobiosyl-glucose. The
symbol used to represent the Sigma enzyme is (□) and the
recombinant enzyme is (○). The concentration of the
recombinant enzyme was 2 μg/mL and a concentration for the
Sigma enzyme was chosen that resulted in the same rate of
glucose production. The complete experimental details are given
in Section 3.8.1.

A)

B)

C)

194

A)

B)
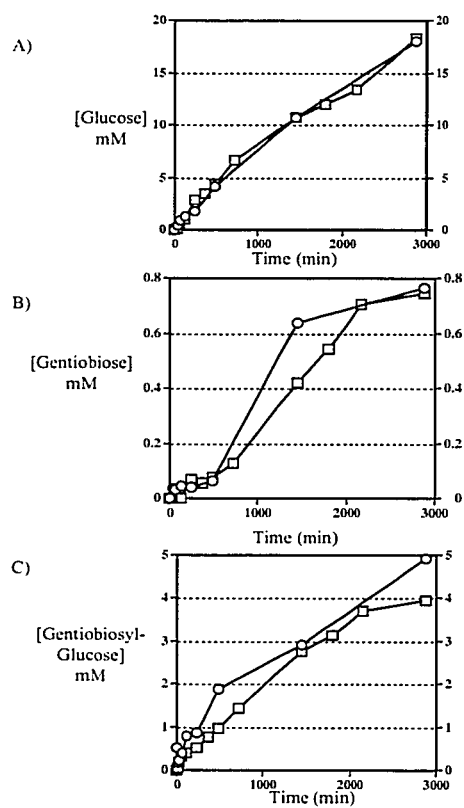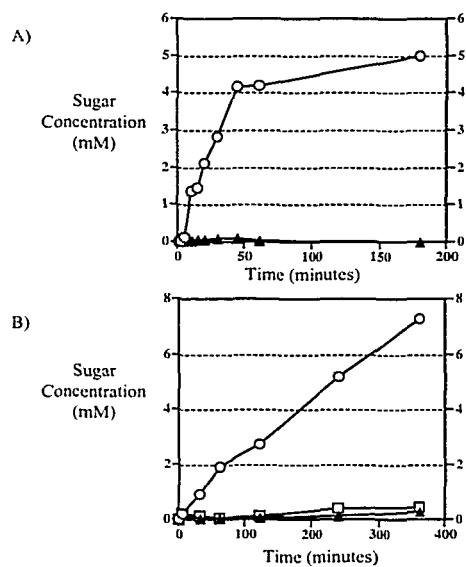
Figure 5.13. Analysis of products resulting from the recombinant β-
glucosidase reaction with pNPGlc, pH 4.5, by GLC. Products
were monitored with: (A) 5 mM pNPGlc and (B) 50 mM
pNPGlc. The reaction was monitored by following the
production of glucose (O), gentiobiose (▲) and, in (B) only,
pNPGe (□). Experimental details are given in Section 3.8.1.
The enzyme concentration used was 2 μg/mL.

was 50 mM (Figure 5.13). the reaction rate was slow in comparison to the rate with 5 mM

pNPGlc. This is expected from the kinetic plots (see Figure 5.3). pNP-Gentiobiose was

produced in the presence of 50 mM pNPGlc but only a very small amount of this

transglucosidic product was observed. It should be noted that the V_{max}(trans) value for

pNPGlc is very small and thus lowered rates of glucose production are expected even

when only a portion of the reaction forms the transglucosidic product. That is, the

substrate acting as an acceptor directs a large proportion of the reaction from hydrolysis to

the formation of pNPGe and thus decreases the amount of glucose produced and since the

V_{max} (trans) value is very small, the amount of trisaccharide produced is very small.

Gentiobiose was also produced. Again, the glucose concentration had to be high relative to

the substrate concentration for gentiobiose formation to occur and thus there was a delay

(it was only formed in significant amounts after the reaction had proceeded for 4 hours).

The breakdown of 50 mM gentiobiose by the recombinant β-glucosidase was also

analyzed using GLC (Figure 5.14). As expected, the glucose production rate was

somewhat slower than the rate when cellobiose was reacted (see Table 5.1). A

trisaccharide was also produced (probably gentiobiosyl-β(1-6)-glucose). It should form

when a gentiobiose joins with a glucose by a β(1-6) bond. Less trisaccharide seemed to be

produced than with 50 mM cellobiose. However, close observation shows that the

proportion of trisaccharide that was produced at early times is quite large. The

trisaccharide probably degrades rapidly and thus it only appears as if not much was

formed.

196
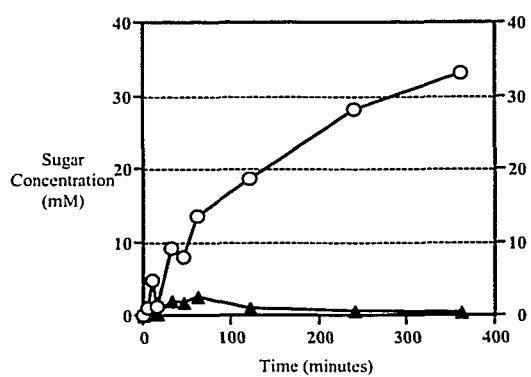
Figure 5.14. Analysis of the products resulting from the recombinant β-
glucosidase reaction with 50 mM gentiobiose by GLC. The
reaction was monitored by following the production of: (O)
glucose, and (▲) trisaccharide. Experimental details are given
in Section 3.8.1. The enzyme concentration used was 2 µg/mL

GLC experiments were also done to determine initial velocities of glucose and trisaccharide production (Figure 5.15) as a function of the cellobiose concentration. These were done to find the concentration of cellobiose that is required for it to be able to act as an acceptor for the transglucosidic reactions. A similar experiment was done with the Sigma enzyme (see Figure 4.17). There were some differences between the two enzymes but the general trend was the same. The initial rates of glucose production increased rapidly as a function of the concentration of cellobiose. At the point where the initial rates of trisaccharide production began to increase (at 15 – 30 mM cellobiose), the initial rate of glucose production declined. As the concentration of cellobiose was increased, the rate of glucose production decreased to the point where it was almost accounted for totally by the transglucosidic reaction (the production of each trisaccharide requires the release of one glucose).

5.9 Mutagenesis strategy for the recombinant β–glucosidase

5.9.1 Sequence analysis

Dr. Shoseyov reported that the β-glucosidase that he was working with is a member of the family 3 glycosyl hydrolases (Personal communication and Dan *et al*. (2000)). The enzyme contained the known active site motif of [LIVM](2)-[KR]-x-[EQK]-x-(4)-G-[LIVMFT]-[LIVT]-[LIVMF]-[ST]-D-(W)-x-[SGADNI]. This confirms that it is a member of the family 3 glycosyl hydrolases (http://expasy.cbr.nrc.ca/cgi-bin/nicedoc.pl?PDOC00621). Dr. Shoseyov's group indicated that they were interested in investigating the properties of the catalytic residues of the protein in addition to some other conserved regions (D261, D501, D145, E140, the region from V382 to N386, and a second
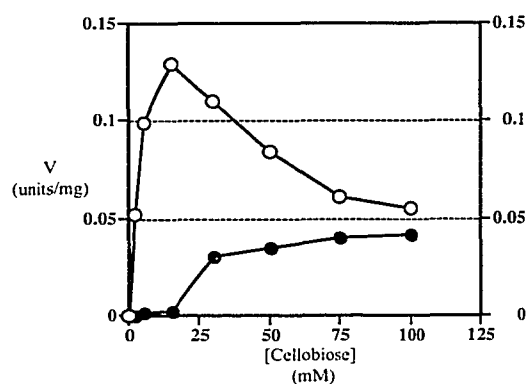
Figure 5.15. Initial rates of production of glucose and trisaccharide by the recombinant β-glucosidase. The initial rates were determined at various cellobiose concentrations. The amounts of glucose and trisaccharide produced were plotted against time and the initial rates of production (V) were determined from the slopes. The rates are plotted above: (O) initial rate of glucose production; (●) initial rate of trisaccharide production. The assays were carried out with 10 μg/mL β-glucosidase. During the reaction the samples were incubated at 25°C. Experimental details regarding the preparation of the samples for GLC are given in Section 3.8.1.

region from Y627 to Y640). I agreed to examine different sites, which limited the number

of conserved residues available for study. Studies by the Shoseyov lab indicated that the

enzyme activity was completely destroyed by reaction with N-bromosuccinimide, 2,3-

butanedione, and tetranitromethane, but that sodium cyanoborohydride and

diethylpyrocarbonate had no effect on enzyme activity (personal communication). If one

assumes that the chemical reagents had free access to all of the side chains (which is not

necessarily the case) this suggests that Trp, Arg, and Tyr may be important for enzyme

activity while Lys and His are not. Sequence alignments were done utilizing only the

fungal family 3 glycosyl hydrolase members in an effort to discover a greater number of

conserved residues. The sequence alignments indicated that a significant number of Trp

were conserved in the fungal family 3 glycosyl hydrolases (Figure 5.16). The sequences

used in the alignment were obtained from GenBank and their accession numbers are

indicated in Appendix #IV.

Tryptophan is important for sugar binding in many glycosyl hydrolases (Quiocho,

1993). Since the evidence presented in Chapter 4 indicates that β-glucosidase has several

glucose binding subsites, it is possible that at least some of the conserved Trp would be

positioned in the subsites. Therefore, I decided to investigate the effect of substitutions for

Trp on the sugar-binding and catalytic properties of the enzyme. From the conservation

found within the sequences it was decided that the Trps at positions 21, 49, 139, 262, 430,

551, 806, and 813 might play some role in binding substrate. Leu-426 was also chosen for

mutagenesis because of the importance of a Trp at an equivalent position at the active site

in a family 3 barley (*Hordeum vulgare*) β-D-glucan exohydrolase whose sequence is quite

Figure 5.16 Alignment of fungal family 3 glycosyl hydrolase β-glucosidases.
The alignment was done with the Textco Gene Inspector
program (version 1.5). The regions of high conservation are
indicated in red, lower conservation in pink or pale pink. No
conservation is indicated by no shading. All sequences were
obtained from GenBank, except the *A. niger* sequence which was
provided by Dr. Shoseyov. Sequence accession numbers are
given in Table 1 in Appendix IV. A Blosum 55 matrix was used
for the alignment, with pairwise grouping parameters of: 1 for k-
tuple (word) size, 20 for the maximum gap length, a gap penalty
of 4 and the number of top diagonals to use was 5. The multiple
sequence alignment parameters were: 9 for gap creation and 8 for
gap extension. The motif for family 3 glycosyl hydrolases is
indicated in purple ([LIVM](2)-[KR]-x-[EQK]-x(4)-G-
[LIVMFT]-[LIVT]-[LIVMF]-[ST]-DW-x-[SGADNI]). The motif
is centered on the conserved aspartic acid residue (shown in
yellow in the purple portion of the alignment) that has been
shown to be involved in the catalytic mechanism of *Aspergillus
wentii* β-glucosidase A3 (Bause and Legler, 1980) and
Aspergillus niger (Dan *et al.*, 2000). The residues chosen for
mutagenesis are indicated with blue boxes and in blue in the *A.
niger* sequence, which is highlighted in cyan. The 19 amino acid
leader sequence is shown in green, and the N-terminal amino
acid of the mature protein is indicated by a green 1 (Dan *et al.*,
2000).

W21   W49   W139   W762

A. oryzae P48825	1	
A. niger AJ132386	1	
H. annuus P06425	1	
K. marxianus P07337	1	
S. fibuligera P22506	1	
S. fibuligera P22507	1	
T. reesei (U09580)	1	
A. wentii (peptide)	1	

W551

W806 W813

similar (Figure 5.17) to the *A. niger* β-glucosidase sequence. The structure of the barley enzyme is the only family 3 glycosyl hydrolase that has been completed and the barley enzyme structure could therefore be used to propose the position and/or role of the Trps (and Leu). The x-ray structure determined by Dr. Fincher's lab (Varghese *et al.*, 1999) is shown in Figure 5.18 and the Trps that were investigated are indicated. Trp 262, Trp 430, and the Trp that is substituted for Leu 426 were shown to be located in the active site of that enzyme. Trp 49, Trp 21 and Trp 551 are some distance from the active site. Trp 49 is buried inside the enzyme molecule but is not very far from the active site. Trp 139 is in between the active site and Trp 551. There are no equivalents to Trp 806 and 813 in the barley enzyme.

Due to the explosion in sequencing that has occurred since I made the original alignments, the number of family 3 glycosyl hydrolase sequences has expanded to over 100 members. Additional work by Dr. G.B. Finchers group at the University of Adelaide, Australia has shown that the family 3 sequences can be subdivided into 6 major branches based on an unrooted radial phylogenetic tree (Harvey *et al.* 2000) (Figure 5.19). The *Aspergillus niger* sequence was not included in the results, but based on sequence analysis it is a member of branch #4, while the barley enzyme is a member of branch #1. The branch #4 subdivision contains 16 other members and a new alignment with these protein sequences is shown in Figure 5.20. As mentioned above, the *H. vulgare* β-glucan exohydrolase is the only family 3 enzyme whose structure is known.

Figure 5.17 Alignment of the *Aspergillus niger* β-glucosidase sequence (BGL-

NCBI-A.n.) with the *Hordeum vulgare* β-D-glucan exohydrolase

(Barley). Both proteins belong to the family 3 glycosyl hydrolases

and there was 37% identity. The sequences were aligned using

Gene Inspector™ version 1.5 by Textco. The scoring matrix used

was Blosum 55. The alignment parameters used were a k-tuple

(word) size of 1, a maximum gap length of 5, a gap penalty of 3,

the number of top diagonals to use was 5, a gap creation of 12 and a

gap extension of 10. The Trp and Leu residues that were substituted

are circled with blue. They are also identified in Figure 5.16.

```
% matched: 37                    10        20        30        40        50
BGL-NCBI-A.n_    1   ELA SPPVV  SPUANQQGD  QAVO AVB  USO     B U
Barley           1   VUL KDATK  ------     UED UA   LI GR     IQQ TQIERL
BGL-NCBI-A.n_   42   --------   M T        L ELCU    -QTGGUP
Barley          39   UATPOULRON FIQS S     SUPRKOATAK   -DHVOGF QKACHST
BGL-NCBI-A.n_   66   UE MCLQ SP L ROSDVHS  A  AGHHVA   K H AVLR   K  GGO SD
Barley          88   I H1YG  AU H QNHUVGAT  I  HNUGLG  FK P HUKRI K T AL URA
BGL-NCBI-A.n_  116   K AD1QLG A AGPLG  DG   LU GF  PALSGULFA  T K 1
Barley         138   T  QYAFA C  1A-UC  RH  LV  E   RR1UQSH-T  L  P   QDUP
BGL-NCBI-A.n_  162   ----DAN UAT   HV1 VEQ  E GQAPEAQ   FGFH SG  SAHLDOKTIHH
Barley         186   KDFTS PPFU   NKU ACA  K UG---DG  TUDO  HH  TI-INREGLH
BGL-NCBI-A.n_  209   ELVL PFAD  IRA AGA    HQ1 HSVG  CQ SVTLNKL   AE G H
Barley         232   NIHHPAYKH  HDK UST    SSU GUKH  HA QLUTGV   DTK
BGL-NCBI-A.n_  258             ---H  HAGUSGA      PGDUD   SGT YHOT
Barley         282   I  EGIDR  ITTPAGSDYS  YSUKAS1      PH-K  QQF1
BGL-NCBI-A.n_  296   H IS LN T  U QL  HA  U HAAVYKU  RDR TP H  FSS TRDEVO
Barley         328   GH CG U  I TS1  AU  T LRVKFTH     FEN
BGL-NCBI-A.n_  346   VKV  USEGP  VK UNOVUNU  QRH SE  F  1GAD      DGA
Barley         363   ----ADPAH  A GLGK---   -QE RDA    AARK     GKTSTDAPL
BGL-NCBI-A.n_  390      TG ERLU  ALI ED GSH  P GAH CSDR  QCDH TLXTG  SSG ANFPY
Barley         404     PK APK1  LUA SH DHL  G OCG      -LH 1E  GD GRTTU
BGL-NCBI-A.n_  440   LU PEQ ISH  E LKHKNG  TATDHHA1DQ  IEALA TASU  LUF HADSQ
Barley         444   GT ILE UKA  A DPSTUU   AENPD----  -AEFU SGGF  VA1 AU-
BGL-NCBI-A.n_  490   E Y1HVDGHL   RR     HR  H DHV1K A  SHCHHT  V   HDUG L HE
Barley         485   - EHPVTETK   EL  1PE  P  STUO UC  GGURCAT  L  - SRGU OP
BGL-NCBI-A.n_  540   HYDHPHUTA1  L G D DRS  HSLAB U V  RUHPGAKS F  GH TREAYQ
Barley         533   LLAASDALUA  AH  S     QGUT A F   DFGFTGRL R  U SU
BGL-NCBI-A.n_  590   V UTER HQ  NO POEDFUE  GVF1DYROFD  KRHETH VEF   SVTTFH
Barley         576   QC  T U  GO HVD---                -LFRL   T
BGL-NCBI-A.n_  640   VSNLEUQ LS  APAVEPASGE  TEAAPTFGEU  DHASDYLYPS  GLQR1TKF1V
Barley         599
BGL-NCBI-A.n_  690   PHLHOTDLEA  SSGDASYGQO  SSDYLPEGAT  DQSAQP1LPA  QQOPQQHPRL
Barley         599
BGL-NCBI-A.n_  740   VDEL1RVSVT  1KHTQKUAGO  EVPQLVVSLQ  GPHEPK1VLR  QFER1TLQPS
Barley         599
BGL-NCBI-A.n_  790   EETKHS TL   RRDLAH HVE  KUH E1TSVP  KHVFV0SSSR  KLPLRASLPT
Barley         599               HA        KK-
BGL-NCBI-A.n_  840   VH
Barley         605   -Y
```

Figure 5.18 Barley β-D-glucan exohydrolase structure with residues that correlate to the β-glucosidase substitutions done in this study shown in red. The numbers refer to the *H. vulgare* enzyme but the equivalents in the *A. niger* β-glucosidase are: W49 (W70), W139 (Y160), W262 (W286), L426 (W430), W430 (W434) and W551(W544). Trp-21 had no equivalent in the *H. vulgare* enzyme. The catalytic nucleophile (D285-*H.vulgare*, D261-*A. niger*) and putative catalytic acid/base (E491) are shown in cyan. The other possible acid/base catalyst is shown in yellow (E220-*H.vulgare*, E195-*A. niger*). The structure was obtained from the RCSB protein data bank (http://www.rcsb.org/pdb/) using the code 1EX1. The structure was reported by Varghese *et al.*, (1999).

208

Figure 5.19. Unrooted radial phylogenetic tree of family 3 glycosyl hydrolases produced by Harvey *et al.* (2000). Alignment of the sequences with the *Aspergillus niger* β-glucosidase indicate that this enzyme would be in branch 4 (Refer to Figure 5.20 for alignment). This was verified by Dr. Maria Hrmova (personal communication). The tree was constructed with the EprotPars program of the University of Wisconsin GCG package. The *H. vulgare* β-glucan exohydrolase is in branch 1. The proteins that have been modeled using the β-glucan exohydrolase are shown in bold type in each branch. This figure is available on the Web at http://planta.waite.adelaide.edu.au/labs/gbf/modeling.htm.

Saccharomyces fibuligera -Glucosidase 2
Saccharomyces fibuligera -Glucosidase
Saccharomyces avenaria -Glucosidase
Phanerochaete chrysosporium -Glucosidase
Phanerochaete chrysosporium -Glucosidase
Agelomyces capsulatus H antigen
Coccidioides immitis(b)
Aspergillus aculeatus -Glucosidase
Aspergillus kawachi -Glucosidase
Trichoderma reesei(a)
Phanerochaete chrysosporium -Glucosidase
Botryotinia fuckeliana -Glucosidase
Botryotinia fuckeliana -Glucosidase
Septoria lycopersici -Glucosidase
Gaeumannomyces graminis(a)
Gaeumannomyces graminis(b)
-Glucan Glucohydrolase
Coccidioides immitis(a)
Cochliobolus heterostrophus -Glucosidase

Ruminococcus ORF
Microlispora bispora(1,3)-
Streptomyces ORF
Streptomyces coelicolor(b) -Xylosidase
Neurospora crassa reesei(b) -Xylosidase
Trichoderma reesei niger -Xylosidase
Aspergillus nidulans -Xylosidase
Aspergillus oryzae -Xylosidase
Aspergillus nidulans(h)
Arabidopsis thaliana(h)
Arabidopsis thaliana(a)

Arabidopsis thaliana(f)

Streptomyces coelicolor(a)
Thermotoga maritima(b)
Thermotoga neapolitana(c)
Thermoanaerobacter brockii -Glucosidase/Xylosidase
Thermoanaerobacter ethanolicus -Xylosidase/Arabinosidase
Streptomyces coelicolor(c)
Streptomyces lividans -Xylosidase
Streptomyces roseiviolaor(d)

Escherichia coli(s)
Salmonella typhimurium -Glucosidase
Chryseobacterium meningosepticum -Glucosidase
Bacteroides fragilis -Glucosidase
Caldocellum saccharolyticum -Glucan Glucohydrolase
(1,4)-
Dictyostelium discoideum -Glucan
Arabidopsis thaliana (g)
Hordeum vulgare -Gluca
Exohydrolase ExoI
Tropaeolum majus -Glucosidase
Nicotiana tabacum -Gluca
Exohydrolase
Hordeum vulgare -Gluca
Exohydrolase ExoI
Arabidopsis thaliana(b)
Arabidopsis thaliana(c)
Arabidopsis thaliana(d)
Arabidopsis thaliana(e)
Arabidopsis thaliana(f)
Arabidopsis thaliana(g)
Haemophilus influenzae ORF
Mycobacterium tuberculosis(a)
Bacteroides fragilis(e)

Thermotoga neapolitana(c)
Pfollia comosa -Glucosidase

Prevotella ruminicola -Glucosidase
Clostridium thermocellum -Glucosidase
Clostridium stercorarium -Glucosidase
Ruminococcus albus(a)

Schizosaccharomyces pombe -Glucosidase
Kluyveromyces fragilis -Glucosidase
Agrobacterium tumefaciens -Glucosidase
Cellulomonas biazotea -Glucosidase
Saccharopolyspora erythraea -Glucosidase
Streptomyces coelicolor(g)
Streptomyces venezuelae ORF
Streptomyces antibioticus Glycosidase
Caldocellum saccharolyticum Glucosidase
Cellvibrio gilvus -Glucosidase
Mycobacterium leprae(a)
Mycobacterium tuberculosis(b)
Glycine max -Glucosidase
Pseudomonas aeruginosa -Glucosidase
Schizophyllum commune -Glucosidase
Aspergillus wentii -Glucosidase
Debaryomyces hansenii bgl protein

Butyrivibrio fibrisolvens -Glucosidase

Ruminococcus albus(b)

Listeria monocytogenes -Glucosidase

Acetivibrio cellulolyticus(b) -Glucosidase

Erwinia chrysanthemi -Glucosidase/Xylosidase A
Bacillus sp -Glucosidase
Alteromonas sp. strain O-7 N-
Acetyl -Glucosaminidase A
Bacillus subtilis ORF
Streptomyces thermoviolaceus N-Acetyl -Glucosaminidase II
Streptomyces coelicolor(b)
Synechocystis sp ORF
Streptomyces coelicolor(c)
Thermotoga maritima(c)
Borrelia burgdorferi(b)
Deinococcus radiodurans ORF
Neisseria meningitidis(a)
Neisseria meningitidis(b)
Escherichia coli(s)
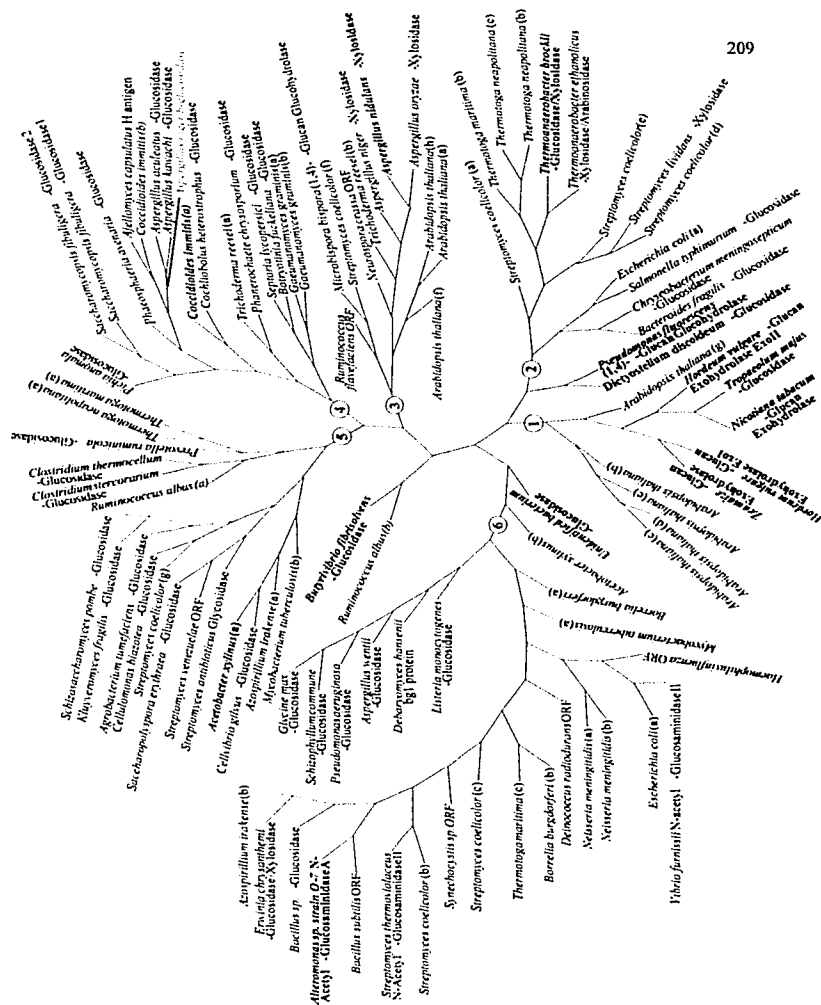Vibrio furnisii N-acetyl -Glucosaminidase II

Figure 5.20 Alignment of family 3 glycosyl hydrolases from branch 4.

The alignment was done with the Textco Gene Inspector
program (version 1.5). The regions of high conservation are
indicated in red, lower conservation in pink or pale pink. The
green box is the area at which the Sigma CNBr peptide was
shown to be the same as in the sequence provided by Dr.
Shoseyov. No conservation is indicated by no shading. All
sequences were obtained from GenBank. Sequence accession
numbers are given in Table 1 in Appendix IV. A Blosum 62
matrix was used for the alignment, with pairwise grouping
parameters of: 1 for k-tuple (word) size, 5 for the maximum gap
length, a gap penalty of 3 and the number of top diagonals to use
was 5. The multiple sequence alignment parameters were: 10 for
gap creation and 10 for gap extension. The motif for family 3
glycosyl hydrolases is indicated in purple ([LIVM](2)-[KR]-x-
[EQK]-x(4)-G-[LIVMFT]-[LIVT]-[LIVMF]-[ST]-DW-x-
[SGADNI]). The residues chosen for mutagenesis are indicated
with blue boxes in the A. niger sequence, which is highlighted in
cyan. The N-terminus of the mature A. niger β-glucosidase is
indicated by a black arrow (residue #20).

A. niger AJ132386
A. CAPSULATUS (U20346)
A. ACULEATUS (D64088)
A. KAWACHI (AB003470)
B. FUCKELIANA (AJ130890)
C. IMMITIS (B) (AF022893)
C. IMMITIS (A) (U87005)
C. HETEROSTROPHUS (AF027687)
G. GRAMINIS (A) (U35463)
G. GRAMINIS (B) (U17568)
P. AVENARIA (AJ276675)
P. CHRYSOSPORIUM (AF036872)
P. ANOMALA (X02903)
S. FIBULIGERA 1 (M22475)
S. FIBULIGERA 2 (M22476)
S. LYCOPERSICI (U24701)
T. REESEI (U09580)

Block 1 (positions 10 20 30 40 50), W139

A. niger AJ132386	130
A. CAPSULATUS (U20346)	131
A. ACULEATUS (D64088)	130
A. KAWACHI (AB003470)	130
B. FUCKELIANA (AJ130890)	112
C. IMMITIS (B) (AF022893)	120
C. IMMITIS (A) (U87005)	146
C. HETEROSTROPHUS (AF027687)	135
G. GRAMINIS (A) (U35463)	132
G. GRAMINIS (B) (U17568)	4
P. AVENARIA (AJ276675)	143
P. CHRYSOSPORIUM (AF036872)	109
P. ANOMALA (X02903)	143
S. FIBULIGERA 1 (M22475)	144
S. FIBULIGERA 2 (M22476)	146
S. LYCOPERSICI (U24701)	132
T. REESEI (U09580)	130

Block 2

A. niger AJ132386	180
A. CAPSULATUS (U20346)	181
A. ACULEATUS (D64088)	180
A. KAWACHI (AB003470)	180
B. FUCKELIANA (AJ130890)	162
C. IMMITIS (B) (AF022893)	190
C. IMMITIS (A) (U87005)	196
C. HETEROSTROPHUS (AF027687)	185
G. GRAMINIS (A) (U35463)	182
G. GRAMINIS (B) (U17568)	34
P. AVENARIA (AJ276675)	193
P. CHRYSOSPORIUM (AF036872)	248
P. ANOMALA (X02903)	193
S. FIBULIGERA 1 (M22475)	194
S. FIBULIGERA 2 (M22476)	196
S. LYCOPERSICI (U24701)	182
T. REESEI (U09580)	180

Block 3

A. niger AJ132386	224
A. CAPSULATUS (U20346)	225
A. ACULEATUS (D64088)	224
A. KAWACHI (AB003470)	224
B. FUCKELIANA (AJ130890)	204
C. IMMITIS (B) (AF022893)	224
C. IMMITIS (A) (U87005)	233
C. HETEROSTROPHUS (AF027687)	222
G. GRAMINIS (A) (U35463)	223
G. GRAMINIS (B) (U17568)	74
P. AVENARIA (AJ276675)	235
P. CHRYSOSPORIUM (AF036872)	279
P. ANOMALA (X02903)	243
S. FIBULIGERA 1 (M22475)	239
S. FIBULIGERA 2 (M22476)	243
S. LYCOPERSICI (U24701)	224
T. REESEI (U09580)	211

Block 4, W262

A. niger AJ132386	274
A. CAPSULATUS (U20346)	275
A. ACULEATUS (D64088)	274
A. KAWACHI (AB003470)	274
B. FUCKELIANA (AJ130890)	254
C. IMMITIS (B) (AF022893)	274
C. IMMITIS (A) (U87005)	283
C. HETEROSTROPHUS (AF027687)	272
G. GRAMINIS (A) (U35463)	273
G. GRAMINIS (B) (U17568)	124
P. AVENARIA (AJ276675)	285
P. CHRYSOSPORIUM (AF036872)	329
P. ANOMALA (X02903)	293
S. FIBULIGERA 1 (M22475)	289
S. FIBULIGERA 2 (M22476)	293
S. LYCOPERSICI (U24701)	274
T. REESEI (U09580)	261

		10	20	30	40	50

A. niger AJ132386   319
A. CAPSULATUS (U20346)   320
A. ACULEATUS (D64088)   319
A. KAWACHI (AB003470)   319
B. FUCKELIANA (AJ130890)   302
C. IMMITIS (B) (AF022093)   319
C. IMMITIS (A) (U87805)   329
C. HETEROSTROPHUS (AF027687)   320
G. GRAMINIS (A) (U35463)   320
G. GRAMINIS (B) (U17568)   171
P. AVENARIA (AJ276675)   330
P. CHRYSOSPORIUM (AF036072)   374
P. ANOMALA (X02903)   332
S. FIBULIGERA 1 (M22475)   336
S. FIBULIGERA 2 (M22476)   340
S. LYCOPERSICI (U24701)   324
T. REESEI (U09580)   305

A. niger AJ132386   367
A. CAPSULATUS (U20346)   368
A. ACULEATUS (D64088)   367
A. KAWACHI (AB003470)   367
B. FUCKELIANA (AJ130890)   351
C. IMMITIS (B) (AF022093)   367
C. IMMITIS (A) (U87805)   377
C. HETEROSTROPHUS (AF027687)   360
G. GRAMINIS (A) (U35463)   370
G. GRAMINIS (B) (U17568)   221
P. AVENARIA (AJ276675)   378
P. CHRYSOSPORIUM (AF036072)   418
P. ANOMALA (X02903)   379
S. FIBULIGERA 1 (M22475)   385
S. FIBULIGERA 2 (M22476)   389
S. LYCOPERSICI (U24701)   374
T. REESEI (U09580)   345

A. niger AJ132386   414
A. CAPSULATUS (U20346)   414
A. ACULEATUS (D64088)   414
A. KAWACHI (AB003470)   414
B. FUCKELIANA (AJ130890)   391
C. IMMITIS (B) (AF022093)   413
C. IMMITIS (A) (U87805)   424
C. HETEROSTROPHUS (AF027687)   415
G. GRAMINIS (A) (U35463)   411
G. GRAMINIS (B) (U17568)   262
P. AVENARIA (AJ276675)   424
P. CHRYSOSPORIUM (AF036072)   453
P. ANOMALA (X02903)   427
S. FIBULIGERA 1 (M22475)   433
S. FIBULIGERA 2 (M22476)   437
S. LYCOPERSICI (U24701)   423
T. REESEI (U09580)   376

A. niger AJ132386   458
A. CAPSULATUS (U20346)   458
A. ACULEATUS (D64088)   450
A. KAWACHI (AB003470)   458
B. FUCKELIANA (AJ130890)   437
C. IMMITIS (B) (AF022093)   457
C. IMMITIS (A) (U87805)   468
C. HETEROSTROPHUS (AF027687)   459
G. GRAMINIS (A) (U35463)   460
G. GRAMINIS (B) (U17568)   311
P. AVENARIA (AJ276675)   468
P. CHRYSOSPORIUM (AF036072)   497
P. ANOMALA (X02903)   471
S. FIBULIGERA 1 (M22475)   479
S. FIBULIGERA 2 (M22476)   482
S. LYCOPERSICI (U24701)   470
T. REESEI (U09580)   422

A sequence alignment figure. Row labels (left column) with sequence residue positions:

Species	Block 1	Block 2	Block 3	Block 4
A. niger AJ132386	506	553	603	646
A. CAPSULATUS (U20346)	507	554	604	647
A. ACULEATUS (D64088)	506	553	603	648
A. KAWACHI (AB003470)	506	553	603	648
B. FUCKELIANA (AJ130890)	481	530	579	629
C. IMMITIS (B) (AF022893)	506	553	603	646
C. IMMITIS (A) (U87805)	510	560	609	651
C. HETEROSTROPHUS (AF027687)	501	550	600	642
G. GRAMINIS (A) (U35463)	502	550	599	648
G. GRAMINIS (B) (U17568)	353	400	448	490
P. AVENARIA (AJ276675)	516	563	613	656
P. CHRYSOSPORIUM (AF036872)	545	592	642	682
P. ANOMALA (X02903)	520	567	617	657
S. FIBULIGERA 1 (M22475)	527	574	624	667
S. FIBULIGERA 2 (M22476)	531	578	628	671
S. LYCOPERSICI (U24701)	512	560	609	654
T. REESEI (U09580)	469	516	566	602

215

A multiple sequence alignment figure.

Ruler positions: 10, 20, 30, 40, 50

Markers: W806, V813

Block 1

Label	Start
A. niger AJ132386	692
A. CAPSULATUS (U20346)	694
A. ACULEATUS (D64088)	692
A. KAWACHI (AB003470)	692
B. FUCKELIANA (AJ130890)	669
C. IMMITIS (B) (AF022993)	693
C. IMMITIS (A) (U87805)	699
C. HETEROSTROPHUS (AF027687)	690
G. GRAMINIS (A) (U35463)	683
G. GRAMINIS (B) (U17568)	527
P. AVENARIA (AJ276675)	703
P. CHRYSOSPORIUM (AF036872)	720
P. ANOMALA (X02903)	705
S. FIBULIGERA 1 (M22475)	717
S. FIBULIGERA 2 (M22476)	721
S. LYCOPERSICI (U24701)	692
T. REESEI (U09580)	641

Block 2

Label	Start
A. niger AJ132386	742
A. CAPSULATUS (U20346)	744
A. ACULEATUS (D64088)	742
A. KAWACHI (AB003470)	742
B. FUCKELIANA (AJ130890)	669
C. IMMITIS (B) (AF022993)	740
C. IMMITIS (A) (U87805)	740
C. HETEROSTROPHUS (AF027687)	737
G. GRAMINIS (A) (U35463)	683
G. GRAMINIS (B) (U17568)	527
P. AVENARIA (AJ276675)	751
P. CHRYSOSPORIUM (AF036872)	723
P. ANOMALA (X02903)	748
S. FIBULIGERA 1 (M22475)	759
S. FIBULIGERA 2 (M22476)	763
S. LYCOPERSICI (U24701)	692
T. REESEI (U09580)	642

Block 3

Label	Start
A. niger AJ132386	789
A. CAPSULATUS (U20346)	791
A. ACULEATUS (D64088)	789
A. KAWACHI (AB003470)	789
B. FUCKELIANA (AJ130890)	709
C. IMMITIS (B) (AF022993)	787
C. IMMITIS (A) (U87805)	789
C. HETEROSTROPHUS (AF027687)	786
G. GRAMINIS (A) (U35463)	723
G. GRAMINIS (B) (U17568)	567
P. AVENARIA (AJ276675)	798
P. CHRYSOSPORIUM (AF036872)	753
P. ANOMALA (X02903)	779
S. FIBULIGERA 1 (M22475)	807
S. FIBULIGERA 2 (M22476)	811
S. LYCOPERSICI (U24701)	732
T. REESEI (U09580)	674

Block 4

Label	Start
A. niger AJ132386	837
A. CAPSULATUS (U20346)	839
A. ACULEATUS (D64088)	837
A. KAWACHI (AB003470)	837
B. FUCKELIANA (AJ130890)	758
C. IMMITIS (B) (AF022993)	835
C. IMMITIS (A) (U87805)	839
C. HETEROSTROPHUS (AF027687)	836
G. GRAMINIS (A) (U35463)	772
G. GRAMINIS (B) (U17568)	615
P. AVENARIA (AJ276675)	846
P. CHRYSOSPORIUM (AF036872)	802
P. ANOMALA (X02903)	816
S. FIBULIGERA 1 (M22475)	855
S. FIBULIGERA 2 (M22476)	859
S. LYCOPERSICI (U24701)	791
T. REESEI (U09580)	722

5.9.2 Rationale for the Amino Acid Substitutions

Tryptophans usually have two types of interactions with sugars. Hydrogen bonding can occur between the Trp nitrogen and the hydroxyl groups of sugars. Secondly, Van der Waals interactions occur through nonpolar-nonpolar associations, a phenomenon referred to as stacking. Stacking occurs because most sugars have one face that is less polar than the other and can therefore, pack against the face of Trp, Phe and Tyr residues. The interaction occurs not only from the complementary nonpolar surfaces, but is also a result of the proximity between the aliphatic protons in the sugar ring and the π-electron cloud of the aromatic ring (Weis and Drickamer, 1996). The aliphatic protons carry a net positive charge because the oxygens of the hydroxyls attached to the same carbons are electron withdrawing.

Initially 4 highly conserved Trp (49, 262, 430 and 551 from Figure 5.16) were substituted with Phe and Ala. Later, substitutions for additional Trp as well as Leu 426 were also created. It was hoped that changes of substrate binding and/or activity that occurred as a result of these substitutions could be used to suggest a role of the Trp for sugar binding. If additional substituted enzymes were subsequently created, degenerate primers were constructed that were capable of producing a variety of alternative substituted amino acids.

Trp 21 was substituted with Tyr, Leu and His. Tyr is similar to Trp in size, although it is less hydrophobic. However, it is still commonly involved in stacking interactions because of its aromatic structure. Leu is hydrophobic. However, it does not have the ring structure necessary for stacking interactions and changes to binding and/or activity that

217

resulted from this substitution would suggest positions where stacking was important. His is smaller than Trp. This could result in a cavity in the protein structure. Additionally, His is much more hydrophilic, which could indicate the importance of a hydrophobic residue at conserved Trp positions. However, His contains an imidazole ring that might be able to form H-bonds similar to those that could be formed with Trp. Trp 49 was initially substituted with an Ala and a Phe. Substitution of Trp residues with a Phe should indicate the importance of hydrogen bonding. Phenylalanine is capable of interactions with sugars via stacking but would not be able to form hydrogen bonds. Substitution of Trp with an Ala would abolish both of these types of interactions and would maintain only a slight hydrophobic environment. Substitution with an Ala would also leave a cavity in the position formerly occupied by the Trp. Later, Trp 49 was also substituted by Asp, Asn and Gly. Dr. Shoseyov's group substituted Trp 139 with a Leu. Trp 262 was substituted with Ala, Cys, Phe, Leu, and Pro. The reasoning for the Ala, Phe and Leu substitutions is the same as was mentioned above. Both Cys and Pro are very different from Trp. Cys is quite small and is considered to be hydrophobic in some hydropathy scales. It is capable of forming hydrogen bonds. Pro is much smaller than Trp and can cause bends in protein chains and influence the protein architecture. It would be expected to cause large differences in substrate binding and enzyme activity. Trp 430 was substituted with Ala, Phe and Leu for the reasons discussed above. Trp 551 was substituted with Ala and Phe. The Leu at position 426 was substituted with Gly, Ala and Val. Gly has only a hydrogen side chain, Ala has only a methyl group, while Val is most like Leu since it only differs by

a methylene group. Only the Val substitution would maintain a hydrophobic environment similar to that of Leu. Both Gly and Ala would leave cavities in the protein structure.

The general philosophy was to study each of the substituted enzymes in detail in order to determine if they play any significant role in this enzyme. However, only preliminary studies were done with Trp-139. Enzymes with substitutions for Trp 806 and 813 were created but since preliminary studies showed that they were not very interesting, those studies were abandoned. The results with Trp 49 and Trp 262 were of most interest since they suggested that these Trp had significant affects. The substitutions of the other Trp did not have such dramatic affects but, in some cases, further study to follow up on some potentially interesting properties is warranted.

6. Results

Part III:

Substitutions for Some Conserved Residues in *Aspergillus niger* β-Glucosidase.

6.1 Characterization of the substituted enzymes

The substituted enzymes were characterized physically and kinetically. In most cases, the experiments described below were done on each of the substituted enzymes. The object was to discover relatively large differences in the hydrolytic and/or transglucosidic activity and of the binding properties that occurred as a result of the substitutions that could lead to new information about the roles of the substituted residues in the enzyme. I also looked for effects that could potentially be exploited industrially.

1) SDS and/or native PAGE were used to determine the size and purity of all of the enzymes. Some examples are shown in Figure 6.1. The changes in protein composition during purification are shown on the gels. In most cases, SDS-PAGE gels with overloaded protein were carried out in addition to the ones shown to show the purity of the enzymes.

2) The gel filtration that was used for purification also worked to establish whether the substituted enzymes eluted from the columns at the same volume as the wild type enzyme. This was done to determine if the substituted enzymes were the same size as the wild type enzyme and to determine if they were still associated as dimers.

3) Fluorescence (emission and excitation) was utilized to determine if major structural changes occurred as a result of the substitutions. There are a large number (21) of Trp in this protein (Dan *et al.* 2000) and thus, although Trp were replaced, significant
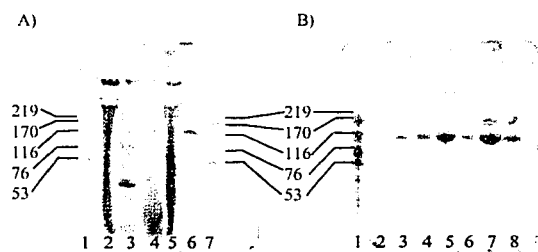
Figure 6.1. SDS-PAGE of some of the substituted β-glucosidases. (A):
enzymes after the ammonium sulfate cuts of 40% and 60%
saturation of the centrifuged liquid media. Lane 1 and 7: molecular
weight markers, Lane 2: W49A, Lane 3: W49F, lane 4: W262A,
lane 5: W262C and lane 6: wild type. In each case the band at
about 116 kDa was the enzyme. (B): enzymes after gel filtration
through the Sephacryl-300 column. Lane 1: molecular weight
marker, lane 2: W211, lane 3: W49A, lane 4: W262A, lane 5:
L426A, lane 6: W430A, lane 7: W551A and Lane 8: wild type. If
impurities were observed on the gel at this stage, the enzyme was
purified further utilizing FPLC size exclusion. Sizes of the
molecular weight markers are indicated on the side of the gel
(kDa). Gels were run with the Phast gel™ system from Pharmacia.
Gradient gels (8-25%) were used. Gels were stained with
Coomassie blue. The apparent higher molecular weight bands in
lanes 7 & 8 of gel (B) are pieces of polyacrylamide.

spectral changes should not occur if the substitutions do not significantly influence the protein structure.

4) Isolation and sequencing of the genes of each of the substituted enzymes from the *P. pastoris* genome was done to confirm that the β-glucosidase gene was actually integrated into the *P. pastoris* genome and to show that the substitution was present.

5) Kinetic studies were done with the β-glucosidases having substitutions for Trp and one Leu using the same substrates and inhibitors utilized with the wild type enzyme in order to obtain clues about the role of the Trp or Leu. Since the substituted enzymes were all expressed in *Pichia pastoris* they were compared with the recombinant wild type enzyme rather than the β-glucosidase isolated from the Sigma *A. niger* cellulase mixture. The substrates used for the kinetic analyses were pNPGlc, oNPGlc, cellobiose, cellotriose, cellotetraose, cellopentaose, and gentiobiose. Such a large number of substrates were tested because information about the hydrophobic nature of the binding could be obtained and because I hoped to find information about the glucose sub-sites. pNPGlc was used to investigate the pH dependence of the substituted enzymes. Competitive inhibition studies were done. Changes in K_i values that resulted from inhibitors could suggest that there were changes to the glucose binding sites in the substituted enzymes. Thus, the natural substrates, cellobiose and gentiobiose (a transferolysis product) were tested as competitive inhibitors. Inhibition by D-glucose and the D-monosaccharides that are epimers of D-glucose (D-mannose, D-allose, D-galactose) as well as D-xylose which is like glucose but has only 5 carbons and therefore, lacks the C_6-OH, were also tested to see if the specificity of the enzyme for glucosyl substrates was changed by the substitutions. Additionally,

gluconolactone and maltose, were used for inhibition studies (see Appendix II for sugar structures). Differences in inhibition by gluconolactone could indicate some involvement of the substitutions in the mechanism of the enzyme because gluconolactone is a transition state analog inhibitor. If the disaccharide maltose was able to act as a good competitive inhibitor, this might indicate large changes in the sugar binding sites of the β-glucosidase as one would not expect maltose to be a good inhibitor. Some of the K_i values obtained were only done at one inhibitor concentration. Such K_i values are subject to more error than if more concentrations are used since the slopes and intercepts (equation # 3, section 3.6.8) of only two lines were used for the calculation. In addition, since the calculation of K_i values is dependent on differences between the V_{max} and K_m values with no inhibitor, and the $^{app}V_{max}$ and $^{app}K_m$ values in the presence of inhibitors, and since each of these kinetic constants is subject to error, the error is magnified by these calculations and is additive. Thus, some caution must be used in the interpretation of the K_i values. This was especially true when the inhibition was poor since the slopes and intercepts in those cases were not very different in the presence or absence of inhibitors and thus the differences between the V_{max} and K_m values and the $^{app}V_{max}$ and $^{app}K_m$ values would be small and could potentially result in relatively large errors. In general, I searched for patterns in the changes of the V_{max}, K_m and K_i values. If there were not any consistent patterns, the Trp or Leu was either not involved or only indirectly involved in interactions between β-glucosidase and its substrates.

6) TLC was utilized to analyze the products that resulted from the reactions of the substituted β-glucosidases with 50 mM cellobiose. In some cases, additional TLC

experiments were done at lower cellobiose concentrations. TLC allowed for the rapid analysis of the effects of the substitutions and it was also the only way that analysis for oligosaccharides of larger than 3 glucoses was done.

7) The breakdown of cellobiose was investigated by GLC. Initially, the breakdown of cellobiose at 5 and 50 mM cellobiose was monitored. The production of glucose, gentiobiose and gentiobiosyl-glucose was followed to determine the affects of the substitutions on the transglucosidic reactions. If warranted, initial velocity experiments of the rates of glucose and trisaccharide production at various cellobiose concentrations (1 mM to 100 mM) were also completed to determine the concentration of cellobiose necessary for transglucosylation to occur.

6.2 Tryptophan 21

Trp-21 was substituted with a His, a Leu and a Tyr using degenerate primers and the Quik-Change mutagenesis kit. The plasmids resulting from the mutagenesis were sequenced to find those with the above substitutions (only the β-glucosidase gene was sequenced). The sequencing also showed that there were not any substitutions at other positions. The plasmids were transformed into *Pichia pastoris* and the proteins were expressed. Integration of the β-glucosidase genes into the *P. pastoris* genome was verified through PCR analysis of the genomic DNA and sequencing (refer to Section 3.15 for details). The three *P. pastoris* organisms with the substituted genes were able to grow normally and expressed the substituted enzyme (the amount expressed was similar to wild type expression). The proteins were purified from the media. All of the enzymes with substitutions for Trp-21 eluted from the gel filtration columns in the same volume of buffer as the wild type enzyme. SDS-PAGE showed that the enzymes were pure and were the same size as wild type. The substituted enzymes were only utilized for kinetic studies if they were >95% pure.

6.2.1 Fluorescence characterization

Substitutions for Trp 21 caused almost no change to the shape of the emission spectra recorded at 275, 285 and 295 nm excitation (parts A, B, and C of Figures 6.2, 6.3, 6.4). In each case, the maxima occurred at approximately the same wavelength as for the wild type enzyme.

A)

Photon
cps

B)

Photon
cps

C)

Photon
cps

D)

mV
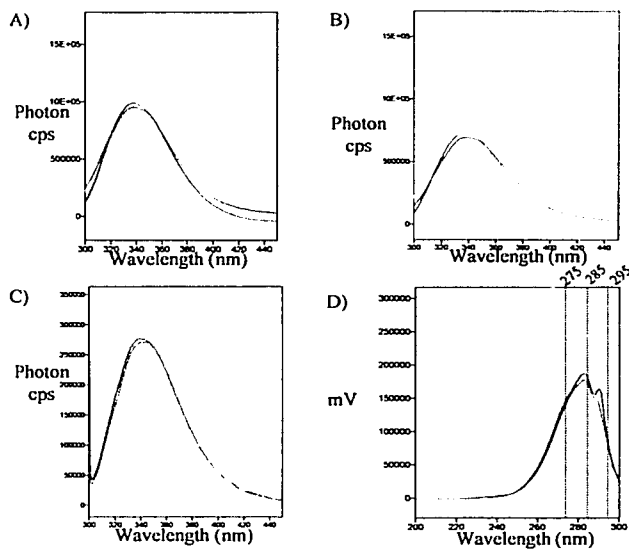
Figure 6.2. Fluorescence emission spectra of W21H (blue) and wild type
(red/pink) β-glucosidase at: A) 275 nm excitation, B) 285 nm
excitation, C) 295 nm excitation and D) excitation spectra obtained
at 339 nm emission. Intensities for emission spectra were in
photon counts per second (cps). Intensity for the excitation spectra
was in millivolts (mV). All spectra were collected at 25°C at a rate
of 1 nm/s. The slit width for emission spectra was 1 nm and for
excitation spectra was 5 nm. The excitation wavelengths used to
produce the emission spectra are indicated by dotted lines in (D).
A_{280} readings of concentrated enzymes were obtained. The
enzymes were then diluted to an A_{280} of 0.029. Measurements
were taken (1/nm) for all spectra.
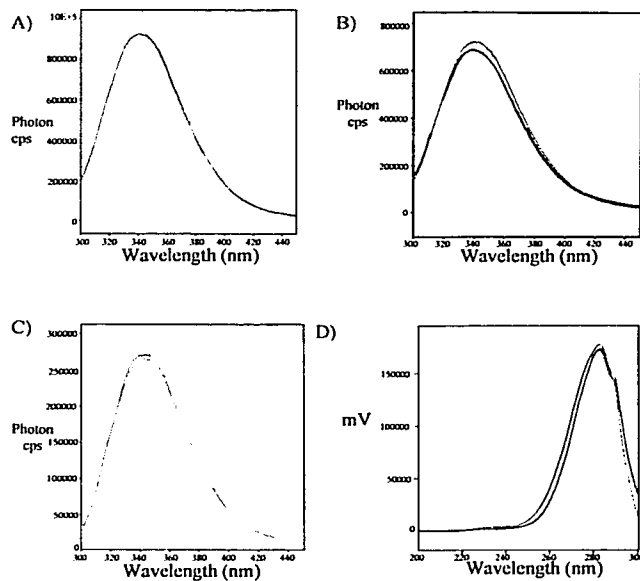
Figure 6.3. Fluorescence emission spectra of W21L (blue) and wild type
(red/pink) β-glucosidase at: A) 275 nm excitation, B) 285 nm
excitation, C) 295 nm excitation and D) excitation spectra obtained
at 339 nm emission. Intensities for emission spectra were in
photon counts per second (cps). Intensity for the excitation spectra
was in millivolts (mV). All spectra were collected at 25°C at a rate
of 1 nm/s. The slit width for emission spectra was 1 nm and for
excitation spectra was 5 nm. A_{280} readings of concentrated
enzymes were obtained. The enzymes were then diluted to an A_{280}
of 0.029. Measurements were taken (1/nm) for all spectra.

A)

B)

C)

D)
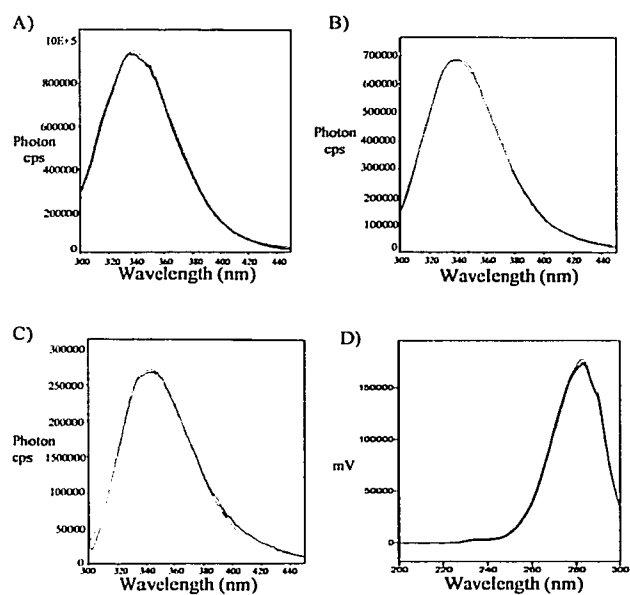
Figure 6.4. Fluorescence emission spectra of W21Y (blue) and wild type
(red/pink) β-glucosidase at A) 275 nm excitation, B) 285 nm
excitation, C) 295 nm excitation and D) excitation spectra obtained
at 339 nm emission. Intensities for emission spectra were in
photon counts per second (cps). Intensity for the excitation spectra
was in millivolts (mV). All spectra were collected at 25°C at a rate
of 1 nm/s. The slit width for emission spectra was 1 nm and for
excitation spectra was 5 nm. A_{280} readings of concentrated
enzymes were obtained. The enzymes were then diluted to an A_{280}
of 0.029. Measurements were taken (1/nm) for all spectra.

The excitation spectra of the three substituted enzymes are shown in part D of Figures 6.2, 6.3, and 6.4. W21H β-Glucosidase was the most different. The shoulder that was found at 288 nm was a small peak rather than just a shoulder (as is the case for the wild type). The excitation spectra of the W21L and W21Y β-glucosidases were almost identical to the excitation spectrum of the wild type enzyme.

6.2.2 Kinetic characterization

6.2.2.1 pNPGlc and oNPGlc

Overall, the K_m values with pNPGlc and oNPGlc (Figure 6.5 and Table 6.1) were not dramatically different from those of the wild type enzyme. There is likely to be some error in the K_m values (especially if the V_{max} values are low) because it is a little difficult to know which points contribute to hydrolysis and which to transglucosidic reactions. This is illustrated in Figure 6.5, which shows the Eadie-Hofstee plots of the substituted enzymes with pNPGlc. One can see that it is somewhat difficult to determine the exact slope of the line for the hydrolysis component of the reactions. Considering the possibility of error, Table 6.1 shows that the K_m values for pNPGlc and oNPGlc are not affected much by substitution for Trp-21.

The V_{max} values were, however, changed quite a lot (Figure 6.5 and Table 6.1. Those of W21L and W21Y β-glucosidases were smaller than for the wild type enzyme while W21H β-glucosidase had a higher V_{max} value for pNPGlc, but a lower V_{max} value for oNPGlc. The biggest difference from the wild type enzyme occurred with W21L β-glucosidase. It had V_{max} values approximately 4-5 times less than those of the wild type enzyme did.
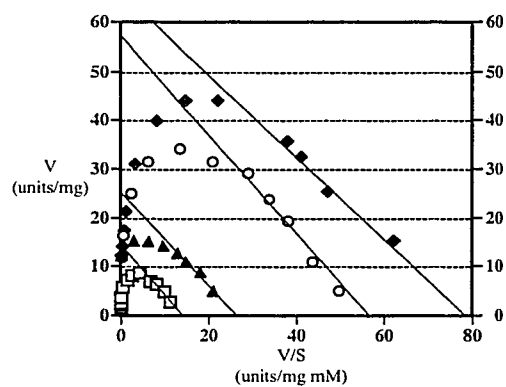
Figure 6.5. Eadie-Hofstee plots of the reactions of the β-glucosidases with substitutions for Trp 21with pNPGlc, pH 4.5, 25°C. Data was acquired by the assay methods described in Section 3.6.2. (O) WT, (◆) W21H, (□) W21L, and (▲) W21Y. The enzyme concentration used was 8μg/mL. The lines illustrate the portion of the curves used to determine the hydrolytic kinetic values. All assays were completed with a final enzyme concentration of 1.4 μg/mL. One unit is equivalent to the production of one μmol of product per min.

The Eadie-Hofstee profiles (Figure 6.5) of the Eadie-Hofstee plots for the substituted enzymes all had similar decreases in rate at high substrate concentrations as did the wild type enzyme. The downward curvature occurred at about the same pNPGlc concentration as with wild type enzyme.

Table 6.1. Kinetic constants for the wild type (WT) and the enzymes with substitutions for Trp-21 with pNPGlc and oNPGlc.

Substrate	WT		W21H		W21L		W21Y	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
pNPGlc	57	1.0	69	0.9	12	0.8	28	1.1
oNPGlc	60	3.0	44	2.2	16	1.9	31	2.6

[a] V_{max} values were in units of μmol/mg/min.
[b] K_m values were in mM.

6.2.2.1.1 Effect of pH

The effects of pH were only investigated for W21Y β-glucosidase. It had similar pH profiles as did the wild type enzyme (Figure 6.6). The main differences were in the magnitudes of the V_{max} and V_{max}/K_m values. The Eadie-Hofstee plots of W21Y β-glucosidase had the same unique profiles as for wild type enzyme. The plots curved downward at low pH (pH 5 and below) and upwards at high pH (pH 6 and above). The V_{max}(trans) values were not measured but appeared to be the same at each pH.

6.2.2.2 Reaction with natural glucosyl substrates

These studies were only done for the complete series of cellooligosaccharides in the case of W21Y β-glucosidase. Cellotetraose and cellopentaose are very expensive. The K_m values for cellobiose were about the same for all three substituted enzymes as for wild

231

Figure 6.6. The effect of pH on the hydrolysis reaction with pNPGlc at 25°C. The wild type β-glucosidase is shown by (O) and W21Y β-glucosidase is shown by (▲). A) V_{max}. B) K_m. C) V_{max}/K_m. Eadie-Hofstee plots (at each pH) were done to obtain the V_{max} and K_m values. Assays were completed according to Section 3.6.5. The enzyme concentration used for each assay was 1.4 μg/mL. One unit is equivalent to the μmol of product produced per min.

A)

B)

C)
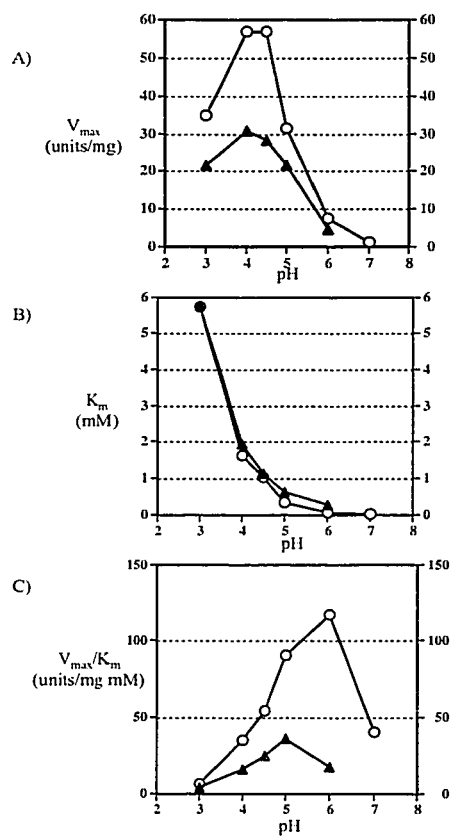
type β-glucosidase (Figure 6.7 and Table 6.2) except that they tended to be about 3 times

larger for cellotriose. However, the K_m values decreased as a function of substrate

length as occurs with wild type β-glucosidase. The trend that was observed with the wild

type enzymes, larger oligosaccharides have a lower K_m, also seems to occur. The K_m

values for gentiobiose were smaller with the substituted enzymes than with the wild type

enzyme. In general, the V_{max} values were decreased by about 1/3 for W21H β-glucosidase

but more for W21L and W21Y β-glucosidases.

Table 6.2. Hydrolytic V_{max} and K_m values for natural glucosyl substrates.

Substrate	WT		W21H*		W21L*		W21Y	
	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]
Cellobiose	78	2.9	55	3.0	24	2.8	36	2.9
Cellotriose	94	0.7	47	1.7	16	1.3	12	2.0
Cellotetraose	66	0.6	ND	ND	ND	ND	20	1.2
Cellopentaose	13	0.6	ND	ND	ND	ND	11	1.1
Gentiobiose	46	2.2	42	1.2	9	1.8	24	1.8

*ND indicates the experiment was not done.
[a] The units for the V_{max} values were μmol/mg/min.
[b] The units for the K_m values were mM.

The Eadie-Hofstee plots for each of these substrates with the substituted enzymes

curved downwards at high substrate concentration (see Figure 6.7 for cellobiose – rest not

shown). This strongly suggests that transglucosidic reactions with the substrates acting as

acceptors occur at high concentrations just as they do with wild type. The curvatures

occurred at about the same concentrations of substrate as they did for the wild type
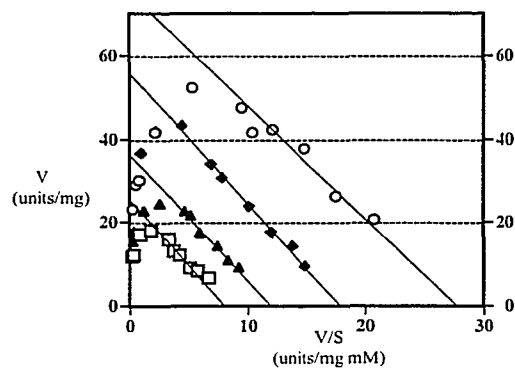
Figure 6.7   Eadie-Hofstee plot illustrating the reactions of the β-glucosidases
with cellobiose. The reaction was done as described in Section
3.6.7 at pH 4.5 and 25°C. Wild type β-glucosidase (O), W21H
(◆), W21L (□), and W21Y (▲). The enzyme concentration
used for the above assays was 8 μg/mL. One unit is equivalent to
one μmol of product produced per min. The lines indicate the
portions of the plot that were used to determine the kinetic values
for hydrolysis.

enzyme, suggesting that these substrates bind to the acceptor sites of the substituted enzymes with about the same affinity as they do to wild type.

6.2.2.3. Competitive inhibition studies

The results are shown in Table 6.3. As stated in Section 6.1 there is some error involved in these values. Taking that into account, there is no definite pattern of inhibition changes that results from the substitutions except that W21Y β-glucosidase bound these inhibitors better than the wild type or the other substituted enzymes. Also, the substituted enzymes seem to bind gentiobiose better than the wild type enzyme.

Table 6.3. K_i values* obtained by the β-glucosidase W21 substituted enzymes in the presence of pNPGlc.

Inhibitor	WT	W21H	W21L	W21Y
Glucose	2.7	4.5	3.1	1.8
Mannose	220	215	360	150
Allose	290	230	ND	130
Galactose	280	180	180	110
Xylose	670	590	290	270
Gluconolactone	0.4	0.4	0.7	0.2
Cellobiose	3.4	2.1	4.0	0.9
Gentiobiose	4.8	2.2	1.9	0.9
Maltose	270	200	230	86

*All K_i values are reported in mM.
ND indicates the experiment was not done.

6.2.3 Analysis of cellobiose degradation with thin layer chromatography

The β-glucosidases with substitutions for Trp-21 all produced gentiobiose and gentiobiosyl-glucose in addition to glucose when they were reacted with cellobiose. TLC plates of the reactions are shown in Figure 6.8 A, B, and C, for W21H, W21L and W21Y-β-glucosidases, respectively. The results, except for differences in V_{max}, were generally the same for all three substituted enzymes and overall were not very different from the results with wild type (Figure 5.7). Glucose and trisaccharide could be detected quite soon after the reaction started. Gentiobiose production only became significant after a large amount of glucose was produced. Both trisaccharide and gentiobiose were transient products and were broken down if the reaction was followed to completion.

6.2.4 Gas-liquid chromatography

The β-glucosidases with substitutions for Trp 21 were analyzed for the rates of glucose, gentiobiose and gentiobiosyl-glucose production at initial cellobiose concentrations of 5 and 50 mM at pH 4.5. It was attempted to match the glucose production rates to the rate with the wild type enzyme so that the relative affects of the substitution on the production of gentiobiose and/or gentiobiosyl-glucose could be illustrated. In the presence of 5 mM cellobiose, the cellobiose was quickly hydrolyzed to give 10 mM glucose (Figure 6.9). As was the case with the wild type β-glucosidase, only negligible amounts of gentiobiose and trisaccharide were produced at this low cellobiose concentration. W21L β-Glucosidase was not investigated in the presence of 5 mM cellobiose.

Figure 6.8 Thin layer chromatographic analysis of the reactions of the W21
substituted β-glucosidases with 50 mM cellobiose. A) W21H B)
W21L and C) W21Y. Enzyme was added (8 μg/mL final
concentration) to the reaction mixture at zero time and an aliquot
(1 mL) was immediately spotted onto the plate. Spotting was
repeated at the times indicated in minutes on the bottom of the
plates. The chromatography solvent used was n-butanol:ethyl
acetate:2-propanol:acetic acid :water in the ratio of 1:3:2:1:1. The
plates were developed twice and were dried in between solvent
exposures. Products were detected by dipping the dried plates in
2% sulphuric acid in methanol and charring. The standards used
on the plates were 10 mM cellobiose (C), 5 mM trisaccharide (T),
10 mM gentiobiose (Ge) and 50 mM glucose (G). The standards
on the plate for W21Y β-glucosidase (i.e. C) had different
concentrations. They were 5, 1, 1, and 1 mM for cellobiose,
gentiobiose, trisaccharide and glucose, respectively. The slight
horizontal line through the produced glucose in (C) was a result of
a scratch on the TLC plate.

A)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 T Ge G
(min)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 T Ge G
(min)

C)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 T Ge G
(min)
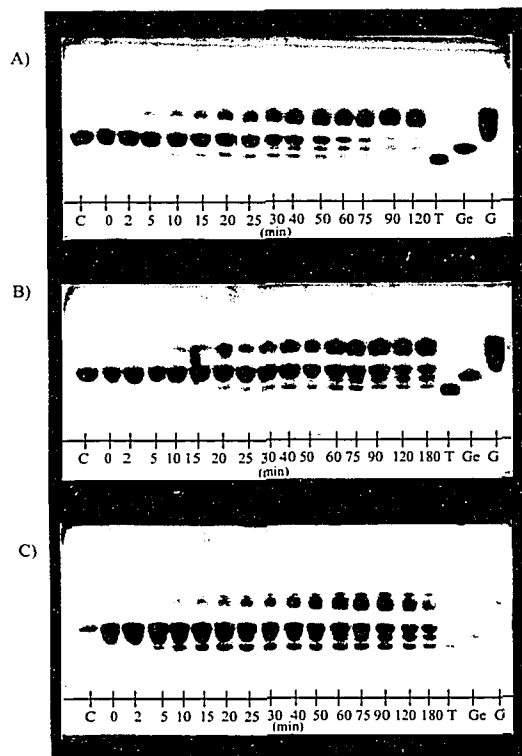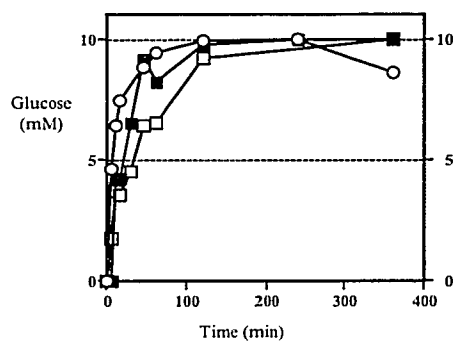
Figure 6.9. Analysis of glucose production from the reaction of the wild type
β-glucosidase and two β-glucosidases with substitutions for Trp
21 with 5 mM cellobiose over 6 hrs. The reactions were carried
out at pH 4.5, 25°C. Wild type enzyme (O) W21H (■), and
W21Y (□). The wild type enzyme concentration was 10 μg/mL
and amounts of the substituted enzymes were added to give
approximately rates approximately equal to that of wild type. The
experimental details are described in Section 3.8.1. No detectable
amounts of gentiobiose or trisaccahride were produced.

The results of the reaction of all three substituted enzymes in the presence of 50 mM cellobiose are shown in Figure 6.10. The experiment was done with a relatively large number of points over only a one hour time to obtain "initial velocity" plots. The data show that the enzymes with substitutions for Trp 21 had essentially the same initial rates for gentiobiose and gentiobiosyl-glucose production as did the wild type β-glucosidase when the glucose production rates were matched. The production of gentiobiose was delayed initially for all of these enzymes (Figure 6.10(B)). As shown earlier for wild type this delay occurs because gentiobiose can only be produced when the glucose concentration is high enough to act as an acceptor. Experiments were also done over an extended period to see the time courses of the whole reaction for each of these substituted enzymes. Again, the results (not shown) were not too different from those of wild type.

Figure 6.10. GLC experimental results with 50 mM cellobiose and an
enzyme concentration that produced an amount of glucose
equivalent to that produced by 10 μg/mL of wild type enzyme.
The reactions were done at pH 4.5 and 25°C. The samples were
prepared for GLC as described in section 3.8.1. The amount of
glucose, gentiobiose and gentiobiosyl-glucose is reported in A,
B, and C, respectively, in mM. Wild type β-glucosidase (O);
W21H (▲); W21L(□) and W21Y (◆).

A)

[Glucose] mM

Time (min.)

B)

[Gentiobiose] mM

Time (min.)

C)

[Gentiobiosyl-Glucose] mM

Time (min)
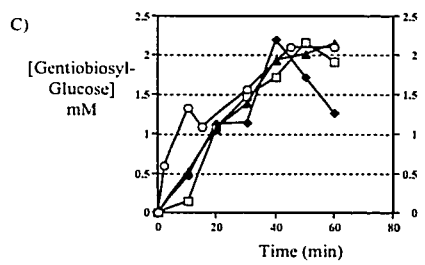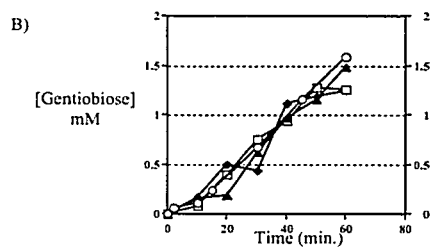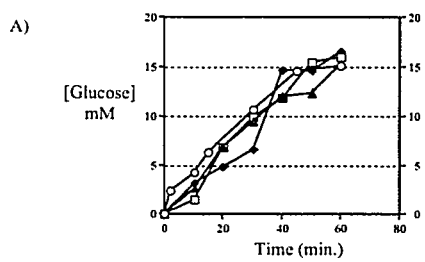
6.3 Tryptophan 49

Trp 49 of β-glucosidase was initially substituted with Ala and Phe. Later, because the substituted enzymes had properties that were quite different from those of wild type β-glucosidase, Trp 49 was also substituted with an Asp, an Asn and a Gly. These latter substitutions were only investigated[*] by GLC. The mutated gene sequences were analyzed by sequencing, to ensure that the proper substitution had occurred, and were transformed into *P. pastoris*. PCR analysis of the yeast genome confirmed the integration of the β-glucosidase gene and sequencing of the PCR product verified that the mutated enzymes were actually part of the genome. The yeast was able to grow normally and the substituted protein was expressed and secreted into the media. The protein was purified. The substituted β-glucosidases eluted from the gel filtration columns in the same volume of buffer as was required to elute the wild type enzyme. The purity of the substituted enzymes was analyzed by SDS-PAGE. All enzymes were greater than 95% pure before they were utilized for kinetic analysis.

6.3.1 Fluorescence characterization

Substitutions for Trp 49 did not result in any significant changes of the fluorescence spectra from those of the wild type enzyme. The peak maxima were at essentially the same wavelength for the enzymes with substitutions by Phe and Ala as they were for the wild type enzyme at all three excitation wavelengths (275, 285 and 295) (Figure 6.11 and 6.12). The excitation spectra for both W49A and W49F β-glucosidase were also similar to

---

[*] The studies with Asp, Asn and Gly were done in collaboration with Sandra Zabrocki (fourth year project, Biochemistry 530, University of Calgary).

A)

B)

C)

D)
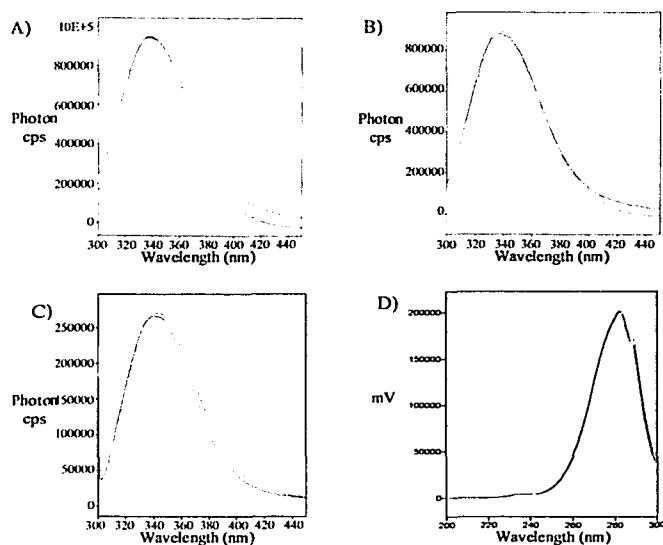
Figure 6.11. Fluorescence emission spectra of wild type β-glucosidase (pink) and
W49A β-glucosidase (blue) at: A) 275 nm excitation, B) 285 nm
excitation, C) 295 nm excitation and D) excitation spectra obtained at
339 nm emission. Fluorescence intensities for emission spectra were
in photon counts per second (cps). Intensity for the excitation spectra
was determined in millivolts (mV). All spectra were collected at a
rate of 1 nm/s. The slit width for emission spectra was 1 nm and for
excitation spectra was 5 nm. All spectra were collected at 25°C. A_{280}
readings were obtained with concentrated enzyme and the enzymes
were then diluted to give an A_{280} of 0.029. Measurements were taken
(1/nm) for all spectra.

Figure 6.12. Fluorescence emission spectra of wild type β-glucosidase (pink) and W49F β-glucosidase (blue) at: A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were in photon counts per second (cps). Intensity for the excitation spectra was determined in millivolts (mV). All spectra were collected at a rate of 1 nm/s. The slit width for emission spectra was 1 nm and for excitation spectra was 5 nm. All spectra were collected at 25°C. A_{280} readings were obtained with concentrated enzyme and the enzymes were then diluted to give an A_{280} of 0.029. Measurements were taken (1/nm) for all spectra.

the wild type spectrum (Figure 6.11 and 6.12 (D)). There was a main peak at 283 nm and a shoulder at 288 nm.

6.3.2 Kinetic characterization

6.3.2.1 pNPGlc and oNPGlc

The V_{max} and K_m values of W49A and W49F β-glucosidase were consistently lower than those of the wild type enzyme (Figure 6.13 and Table 6.4). The Eadie-Hofstee plots of the substituted enzymes were also different (Figure 6.13). The enzymes with substitutions for Trp 49 required higher concentrations of substrate for the downward curvature to occur. The curvature with the wild type enzyme begins to occur at quite low pNPGlc concentrations (at about 3 mM) while with the substituted enzymes it first begins to occur at pNPGlc concentrations of about 10 -20 mM.

Table 6.4. Kinetic values obtained with W49 substituted enzymes with aryl-glucosides.

Substrate	WT		W49A		W49F	
	V_{max}^a	K_m^b	V_{max}^a	K_m^b	V_{max}^a	K_m^b
pNPGlc	56.9	1.0	11.8	0.5	6.8	0.5
oNPGlc	60.2	3.0	13.2	1.3	3.4	1.2

[a] The V_{max} is in units of μmol/mg/min.
[b] The K_m is in units of mM.

6.3.2.1.1 Effect of pH

The pH profiles that resulted from the hydrolytic activity of wild type, W49A and W49F β-glucosidase are shown in Figure 6.14. The V_{max} profiles of the substituted enzymes followed similar trends as the wild type enzyme except the magnitude was about
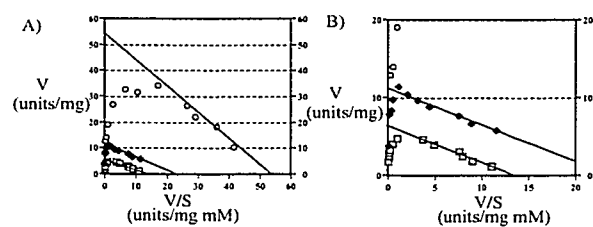
Figure 6.13. Eadie-Hofstee plots of the reaction of the wild type and of the β-glucosidases with substitutions for Trp-49 with pNPGlc, at pH 4.5 and 25°C (A). Wild type enzyme (O), W49A (◆) and W49F (□). (B) Close-up of W49A β-glucosidase and W49F β-glucosidase. Data was acquired by the assay methods described in Section 3.6.2. The lines illustrate the portion of the curves used to determine the kinetic values in Table 6.22. All assays were completed with an enzyme concentration of 1.4 μg/mL. One unit is equivalent to the production of one μmol of product per min.

Figure 6.14. The effect of pH (pNPGlc) on the hydrolysis reaction for: wild type (O), W49A (◆) and W49F (□). A) V_{max}. B) K_m. C) V_{max}/K_m. Eadie-Hofstee plots (at each pH) were used to determine the V_{max} and K_m values. All of the reactions were done using 1.4 µg/mL enzyme at 25°C. Assays were completed according to Section 3.6.5.

A)

B)

C)
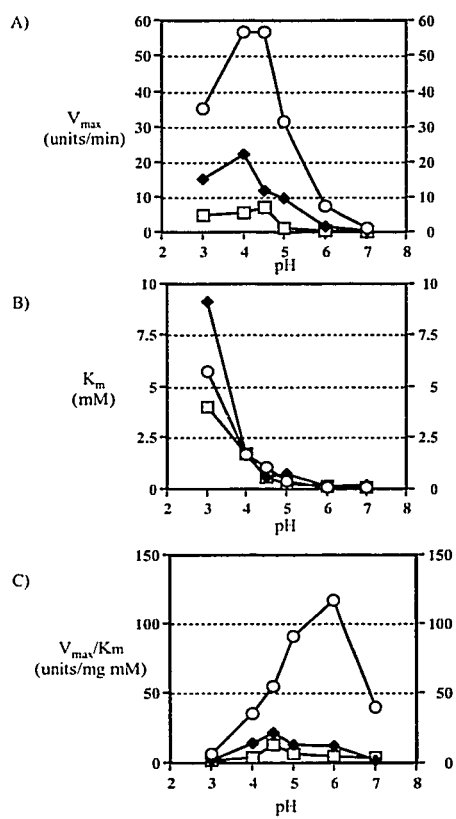
3 times lower with W49A β-glucosidase and about 10 times lower with W49F β-

glucosidase. The K_m pH profiles were quite similar except that both differed slightly from

the wild type values at low pH. The change in V_{max}/K_m with pH showed the largest

differences between the enzymes with substitutions for Trp 49 and the wild type enzyme.

The values were much higher for wild type and the peak maximum occurred at pH 6. For

the substituted enzymes the peak maxima occurred at pH 4.5.

6.3.2.2 Reaction with natural glucosyl substrates

The V_{max} values for the hydrolytic component of the reactions catalyzed by the

enzymes with substitutions for Trp-49 were much lower than they were with the wild type

enzyme (Figure 6.15 and Table 6.5).

Table 6.5. A comparison of kinetic values obtained from reaction of wild type and W49
substituted β-glucosidases with natural substrates.

Substrate	WT		W49A		W49F	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
Cellobiose	78	2.9	14	1.9	0.9	0.35
Cellotriose	94	0.7	23	3.2	1.1	0.5
Cellotetraose	66	0.6	38	3.9	1.2	0.6
Cellopentaose	13	0.6	11	1.7	0.5	0.7
Gentiobiose	46	2.2	4.9	2.1	0.8	0.6

[a] V_{max} values are reported in units of μmol/mg/min.
[b] K_m values are reported in mM.

Of significance is that the K_m values with W49A β-glucosidase were low for cellobiose

but increased rather than decreased as the number of glucose monomers in the substrate
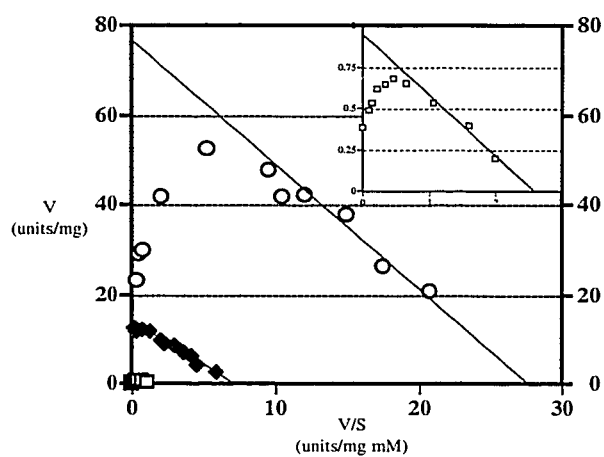
Figure 6.15. Eadie-Hofstee plots of the reaction (cellobiose) of the wild type
β-glucosidase and the enzymes with substitutions for Trp 49. Wild
type enzyme (O) W49A (◆), and W49F (□). The inset plot
shows the W49F plot on a larger scale. The assays were done as
described in section 3.6.7 at pH 4.5 and 25°C with 8 μg/mL
enzyme. One unit is equivalent to the production of 1 μmol of
product per min.

increased. This is different from wild type β-glucosidase. In that case, the K_m values decreased relative to cellobiose as the oligosaccharide length increased. The biggest decrease with wild type enzyme is from cellobiose to cellotriose (about 4 fold). For these substituted enzymes the K_m values increased about 1.5 fold for cellotriose compared to cellobiose. This is significant since it is a clue as to why this enzyme has a lowered ability to carry out transglucosidic reactions. There was only a hint of the expected downward curvature of the Eadie-Hofstee plot (at pH 4.5) with W49A β-glucosidase (Figure 6.15). In the case of W49F β-glucosidase the curvature was more obvious.

6.3.2.3 Competitive inhibitor studies

In general, the K_i values were lower for the substituted enzymes (Table 6.6) than for wild type β-glucosidase.

Table 6.6. K_i values (mM) obtained from glucosyl monosaccharides and disaccharides.

Inhibitor	WT	W49A	W49F
Glucose	2.7	2.4	0.6
Mannose	220	145	40
Allose	290	390	210
Galactose	280	420	82
Xylose	670	520	698
Gluconolactone	0.4	0.2	0.2
Cellobiose	3.4	2.8	1.0
Gentiobiose	4.8	1.8	0.9
Maltose	270	540	250

W-49F β-Glucosidase was inhibited significantly and consistently better than wild type by most of the sugars (especially the glucosyl inhibitors). Despite the error involved, the consistently lower K_i values must mean that the inhibitors bind better especially when taken together with the lower K_m values for this substituted enzyme.

6.3.3 Analysis of cellobiose breakdown by TLC

When 10 mM cellobiose was reacted with W49A β-glucosidase only glucose was produced (Figure 6.16(A)). Similar experiments at 10 mM cellobiose with the wild type enzyme resulted in the production of small amounts of both trisaccharide and gentiobiose (data not shown). In the presence of 50 mM cellobiose, only a small amount of gentiobiose and/or trisaccharide (Figure 6.16 (B)) was produced (it is hard to tell whether the spot below the cellobiose was gentiobiose or trisaccharide). Similar results at these two cellobiose concentrations were obtained with W49F β-glucosidase (Figure 6.17(A)). However, the hydrolytic reaction is quite slow with this enzyme and so the amount of glucose produced was also low. No gentiobiose or trisaccharide were produced in the presence of 10 mM cellobiose. At 50 mM cellobiose (Figure 6.17 B)) a small amount of trisaccharide and/or gentiobiose was produced (again it was hard to determine which it was). The main product was glucose. These TLC data show that the enzymes with substitutions for Trp-49 result in the production of significantly less transglucosidic product than occurs with wild type β-glucosidase.

Figure 6.16. Thin layer chromatography analysis of the reaction of W49A β-
glucosidase with cellobiose: A) 10 and B) 50 mM ; at 25°C.
Enzyme was added (8 μg/mL final concentration) to the reaction
mixture at zero time and an aliquot (1 μL) was immediately
spotted onto the plate. Spotting was repeated at the times indicated
in min. on the bottom of the plate. The chromatography solvent
used was n-butanol:ethyl acetate:2-propanol:acetic acid:water in
the ratio of 1:3:2:1:1. The plates were eluted twice and they were
dried in between elutions. Products were detected by dipping the
dried plates in 2% sulphuric acid in methanol and charring. The
standards used on the plates were 10 mM cellobiose (C), 10 mM
gentiobiose (Ge) and 20 mM glucose (G). The slight horizontal
line through the glucose in (B) was a result of a scratch on the
TLC plate.

A)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 240 Ge G
(min)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 240 Ge G
(min)
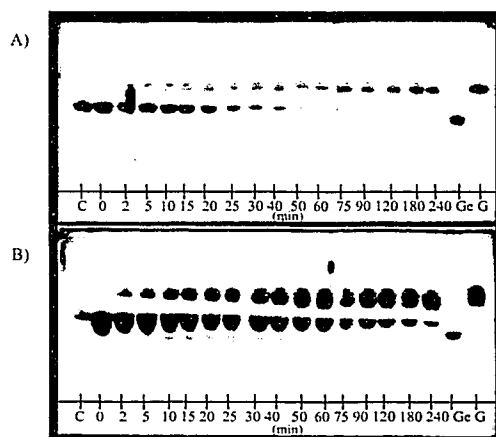
Figure 6.17. Thin layer chromatography analysis of the reaction of W49F β-
glucosidase with: A) 10 and B) 50 mM cellobiose at 25°C.
Enzyme was added (8 µg/mL final concentration) to the reaction
mixture at zero time and an aliquot (1 µL) was immediately
spotted onto the plate. Spotting was repeated at the times indicated
in minutes on the bottom of the plate. The chromatography solvent
used was n-butanol:ethyl acetate:2-propanol:acetic acid :water in
the ratio of 1:3:2:1:1. The plates were eluted twice and they were
dried in between solvent exposures. Products were detected by
dipping the dried plates in 2% sulphuric acid in methanol and
charring. The standards used on the plates were 10 mM cellobiose
(C), 5 mM trisaccharide (T), 10 mM gentiobiose (Ge) and 10 mM
glucose (G).

A)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 T Ge G
(min.)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 180 T Ge G
(min.)
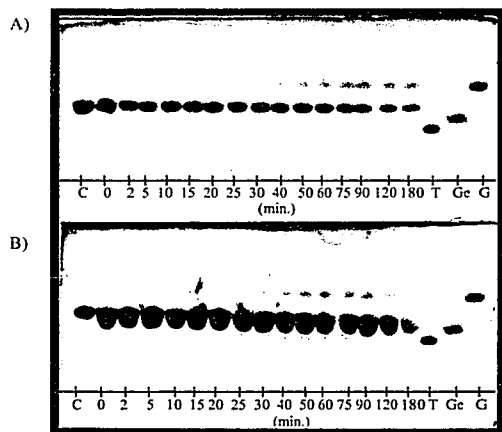
6.3.4 Gas-liquid chromatography

The enzyme reactions of the β-glucosidase with substitutions for Trp-49 with 5 and 50 mM cellobiose were followed by gas chromatography. Concentrations of the substituted enzymes were adjusted so that they produced an equivalent rate of glucose production as did wild type enzyme when its concentration was 10 µg/mL. This was done so that the amounts of transglucosidase products could be compared relative to the production of glucose. In the presence of 5 mM cellobiose both of the substituted enzymes behaved as wild type in that mainly hydrolysis was occurring. The cellobiose was completely hydrolyzed to glucose with no evidence at all for the production of trisaccharide or gentiobiose (Figure 6.18). The reactions of the enzymes with 50 mM cellobiose are shown in Figure 6.19. Significantly less gentiobiose and gentiobiosyl-glucose were produced by the β-glucosidases substituted for Trp-49 than for wild type. This is especially obvious at very early times (the first 10 minutes).

Because of the significantly lower amount of the transglucosidic reaction component compared to the hydrolytic reaction component with these enzymes, the reactions were studied in more detail. The initial rates of glucose and trisaccharide production were determined from plots of the amount of sugar produced over a short time period at cellobiose concentrations between 1 and 100 mM. These experiments were done with enzymes having 5 different substitutions. A relatively large number of points were taken over a short period of time and the slopes of the lines were determined. These initial rates were plotted against substrate concentration to produce Michaelis-Menten type plots (Figure 6.20). The plot for glucose production (Figure 6.20(A)) by wild type β-

[Glucose]
mM

10

5

0

0   100   200   300   400
Time (min)
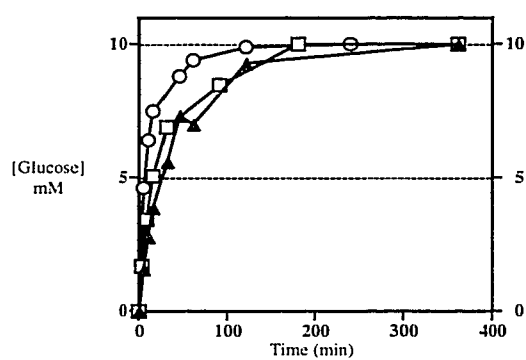
Figure 6.18. Analysis (GLC) of glucose production from the reaction of the
wild type β-glucosidase and enzymes with substitutions for Trp-
49 with 5 mM cellobiose over 6 hrs. Wild type enzyme (o),
W49A (▲), and W49F (□). Enzyme concentrations were chosen
for W49A and W49F β-glucosidase that produced an equivalent
amount of glucose as 10 μg/mL of wild type enzyme. The
reaction was done at pH 4.5 and 25°C. The samples were then
prepared for GLC as described in section 3.8.1.
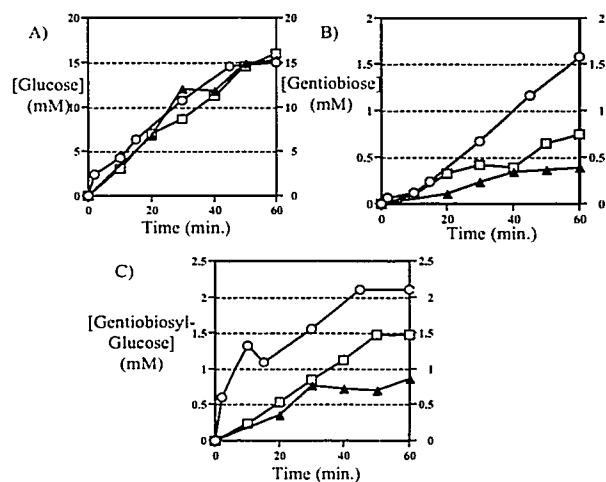
Figure 6.19 GLC results of the substituted W68 enzymes and wild type with
50 mM cellobiose. The concentration of enzyme used for the
substituted enzymes was equivalent in terms of the rate of
glucose production by the wild type enzyme (10 μg/mL). The
enzyme was added to cellobiose at time zero and aliquots were
removed every 10 min. and treated as described in Section
3.8.1. A) glucose; (B) gentiobiose; (C) gentiobiosyl glucose.
Wild type (O); W49A (▲); and W49F (□).

glucosidase was hyperbolic at low cellobiose concentration. However, when the concentration of cellobiose became higher (about 15-30 mM cellobiose), the rate of glucose production decreased and the curves took a definite dip. This happens because the cellobiose acts as an acceptor and the slower transglucosylation reaction occurs rather than hydrolysis. At the same time that the glucose production rate decreased, the rate of trisaccharide production increased (Figure 6.20(A)). At higher cellobiose concentration the rate of glucose and trisaccharide production is almost at a 1:1 ratio. Since the release of one glucose is required for production of one trisaccharide this shows that the main reaction at high concentrations of cellobiose is transglucosylation.

The results of similar experiments with β-glucosidases with substitutions for Trp-49 (Figure 6.20 (B) to (F)) are clearly different from those for the wild type enzyme. W49A β-Glucosidase (Figure 6.20(B)) has a much slower trisaccharide production rate than the wild type enzyme. There is really no cellobiose concentration that results in a noticeable increase in trisaccharide production with this enzyme. The rate of glucose production does not decrease, but displays the expected hyperbolic Michaelis-Menten plot. The substitution of Phe, and Asn for Trp 49 (Figure 6.20 (D and F)) caused similar effects. When substituted with Asp and Gly (Figure 6.20 (C and E)) there were somewhat greater rates of trisaccharide production. The rates of trisaccharide production in these cases increased somewhat at about 30-50 mM cellobiose, and this coincided with decreases in glucose production. However, the rate of trisaccharide production was smaller than for wild type, and the decreases in the rates of glucose production were less than the decrease observed with wild type β-glucosidase.

Figure 6.20. Michaelis-Menten plots for the wild type and β-glucosidases with substitutions for W49. The rates of glucose production (O) and gentiobiosyl-glucose production (▲) are shown. All rates were determined from individual plots of sugar production over time at each cellobiose concentration. (A) wild type; (B) W49A; (C) W49D; (D) W49F (E) W49G; (F) W49N. The enzyme concentrations of the substituted β-glucosidases were chosen because they had the same glucose production rate as 10 μg/mL of wild type enzyme. All reactions were done at 25°C, pH 4.5. Experimental details are given in Section 3.8.1.

A)

B)

C)
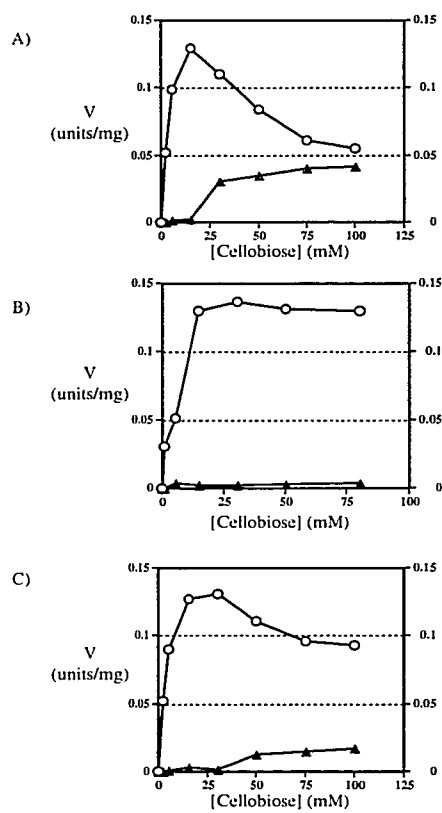
D)
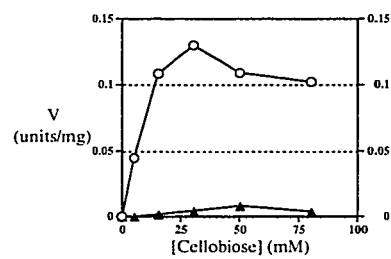
264

E)
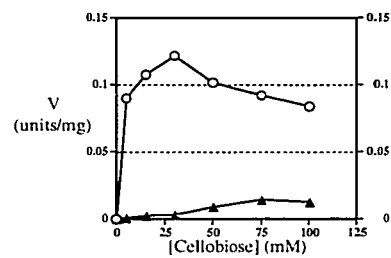
F)
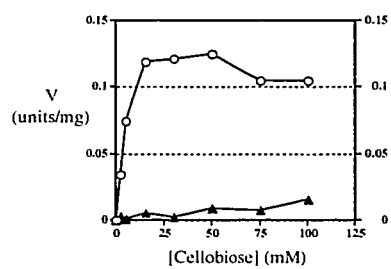
6.4. Tryptophan 139

Trp-139 was substituted with a Leu by Dr. Shoseyov's group in Israel and given to us after it was transformed into *P. pastoris* strain KM71. The yeast was grown in the same manner as the other KM71 strains of *Pichia pastoris*. The enzyme was secreted into the media by the yeast and was purified. W139L β-Glucosidase eluted with the same volume of buffer as did the wild type enzyme. Some very preliminary studies were done on this enzyme with pNPGlc (Figure 6.21), cellobiose and gentiobiose. Fluorescence, TLC and GLC experiments were not done. The kinetic data that was obtained is shown in Table 6.7, and the pH profiles for W139L β-glucosidase are shown in Figure 6.22. The V_{max} values were significantly smaller than for wild type, but the V_{max} for pNPGlc was larger than for the natural substrates. The K_m was 3 times larger than that of wild type when pNPGlc was the substrate but was 4 to 5 times lower when cellobiose and gentiobiose were substrates, respectively. This was the opposite of what was observed with the wild type enzyme where reactions with both cellobiose and gentiobiose result in higher K_m values than with pNPGlc. The pH affected the substituted enzyme in a similar way as wild type (Figure 6.22). However, at each pH the V_{max} was lower and the K_m was higher for the substituted β-glucosidase. The Eadie-Hofstee plots curved downwards in each case. This suggests that the transglycosylation reaction occurs with this substituted enzyme. In each case the V_{max}(trans) seems to be about 10-15% of the value of the hydrolytic V_{max}. If, as is likely, this is due to the transglucosidic reaction, it indicates a similar extent of transglucosidic activity as with the wild type enzyme.
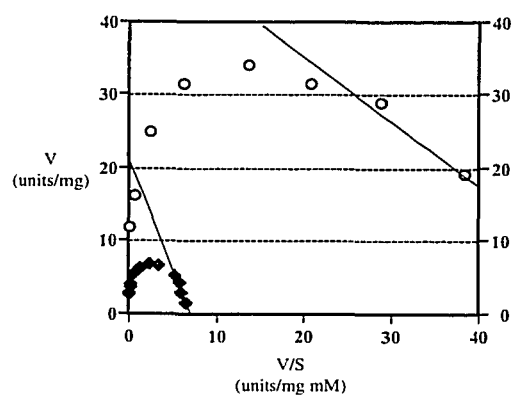
Figure 6.21. Eadie-Hofstee plot of the reaction of wild type (O) and W139L
($\blacklozenge$) β-glucosidases with pNPGlc at pH 4.5, 25°C. The enzyme
concentration used was 1.4 μg/mL. The lines represent the
portions of the plot that were utilized to determine the V_{max} and
K_m values. Only a small part of the wild type data is shown for
comparison. One unit is equivalent to one μmol of product
produced per min.

Figure 6.22. The effect of pH on the W139L β-glucosidase with pNPGlc.
Assays with W139L β-glucosidase were only done between pH 4
to pH 7 (25°C). They were done between pH 3 and 7 for wild
type. The enzyme concentration used was 1.4 µg/mL. Kinetic
values were obtained from the hydrolytic portion of the resultant
Eadie-Hofstee plots. A) V_{max}, B) K_m, C) V_{max}/K_m. wild type β-
glucosidase (O); W139L (▲). One unit is equivalent to 1 µmol
of product per min.

A)

B)

C)
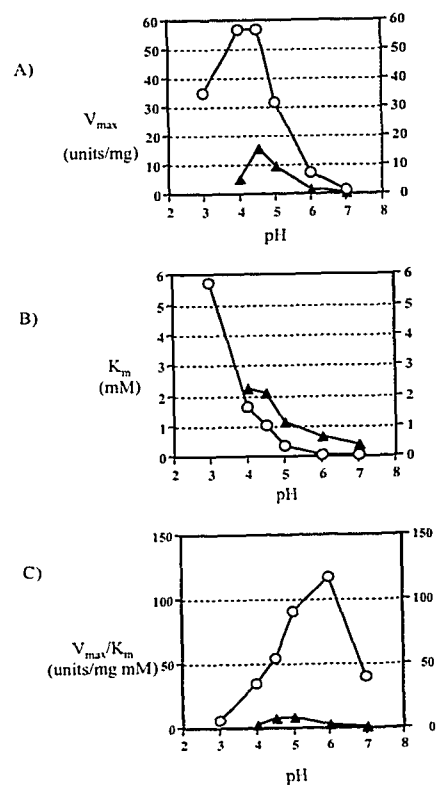
Table 6.7. Kinetic constants for the wild type (WT) and the enzyme with a substitution for Trp-139.

Substrate	Wild Type		W139L	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
pNPGlc	57	1.0	22	3
Cellobiose	78	2.9	8	0.7
Gentiobiose	46	2.2	13	0.4

[a] V_{max} values are reported as μmol/mg/min.

[b] K_m values are reported as mM.

6.5 Tryptophan 262

Trp-262 was thought to be important since alignment of family 3 glycosyl hydrolase sequence pinpointed it as part of the active site motif (SDW) for these enzymes. It is very highly conserved. Asp-261 is the catalytic nucleophile (Dan *et al.* 2000). A Trp is also present at the equivalent position in the structure of the barley exohydrolase family 3 glycosylase (Varghese *et al.* 1999).

W262L-β-Glucosidase was created by Dr. Shoseyov's group. W262A, W262C, W262F, and W262P β-glucosidases were created as part of this study. The plasmids resulting from the mutagenesis were sequenced to determine that the substitutions did occur and that there were not any substitutions at other positions. The plasmids were transformed into *Pichia* and the proteins were expressed. Integration of the β-glucosidase gene into the *P. pastoris* genome was verified through PCR analysis of the genomic DNA. The PCR product was isolated and then sequenced (refer to Section 3.15 for details). The three *P. pastoris* strains with the substituted genes were able to grow normally and expressed the substituted enzymes (the amount expressed was similar to wild type expression). The proteins were purified from the media. All of the enzymes with substitutions for Trp 262 eluted from the gel filtration columns in the same volume of buffer as the wild type enzyme. SDS-PAGE showed that the enzymes were pure, and that the sizes of the subunits were the same size as wild type. The substituted enzymes were utilized for kinetic studies only if a single band (>95% pure) was present on an SDS-PAGE gel. Each of the substituted enzymes except W262P β-glucosidase was studied in detail. W262P β-glucosidase had such low activity that a study of its properties was not pursued.

6.5.1 Fluorescence studies

The emission spectra of the substituted enzymes were collected at the excitation

wavelengths: 275, 285, and 295 (Figure 6.23, 6.24, 6.25, and 6.26). The enzymes

substituted with Ala, Phe and Leu gave spectra that were similar to the spectra of wild type

β-glucosidase at all three excitation wavelengths (Figures 6.23, 6.25, and 6.26,

respectively). The spectra of W262C β-glucosidase appeared to be slightly red-shifted. It

had peak maxima (Figure 6.25) that were at slightly lower wavelengths than was seen for

wild type and the other substituted β-glucosidases. This red shift was least noticeable when

the excitation was 295 nm.

Excitation spectra were also obtained for these enzymes (part (D) of Figures 6.23,

6.24, 6.25, and 6.26). All of these enzymes displayed characteristic excitation spectra with

the same peak at 283 nm that was observed for wild type and with a similar shoulder at 288

nm. The excitation spectrum for W262C β-glucosidase (Figure 6.24 (D)) was again slightly

red-shifted compared to the spectrum of wild type or to those of the other enzymes with

substitutions for Trp 262.

6.5.2 Kinetic characterization

6.5.2.1 pNPGlc and oNPGlc

The kinetics for these enzymes with substitutions for Trp-262 were very different

from those for the wild type enzyme. In contrast to Eadie-Hofstee plots of the wild type

enzyme, the enzymes with substitutions for Trp-262 were linear (Figure 6.27). Also, there

were no hints of downward curvatures of the lines, even at very high concentrations of

substrate. The V_{max} values of all of the substituted enzymes were much smaller (at least 10
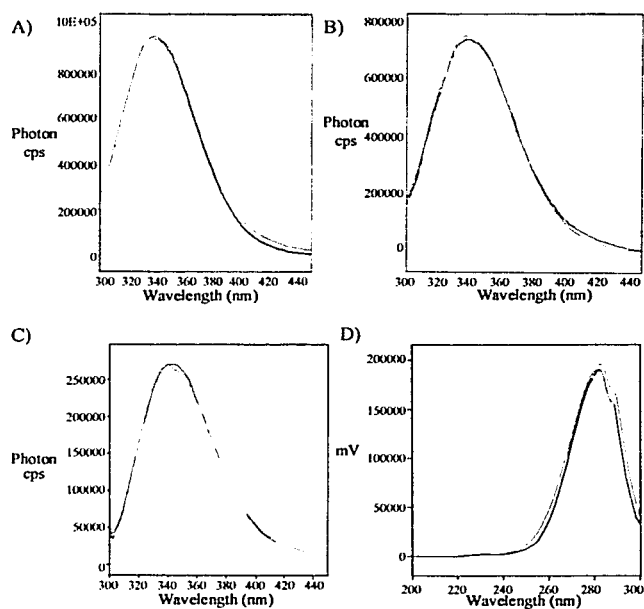
Figure 6.23. Fluorescence emission spectra of W262A β-glucosidase (blue) at
A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation
and D) excitation spectra obtained at 339 nm emission. Intensities
for emission spectra were in photon counts per second (cps).
Intensity for the excitation spectra was determined in millivolts
(mV). All spectra were collected at 25°C. Spectra for W262A β-
glucosidase (blue) and for the wild type β-glucosidase (pink) are
shown. A_{280} readings were obtained with the concentrated
enzymes. The enzymes were diluted to give an A_{280} of 0.029.

$10E+05$

800000

600000

Photon
cps
400000

200000

0

300 320 340 360 380 400 420 440
Wavelength (nm)

700000

600000

500000

Photon
cps
400000

300000

200000

100000

0

300 320 340 360 380 400 420 440
Wavelength (nm)

250000

200000

Photon
cps
150000

100000

50000

0

300 320 340 360 380 400 420 440
Wavelength (nm)

300000

mV
200000

100000

0

200 220 240 260 280 300
Wavelength (nm)
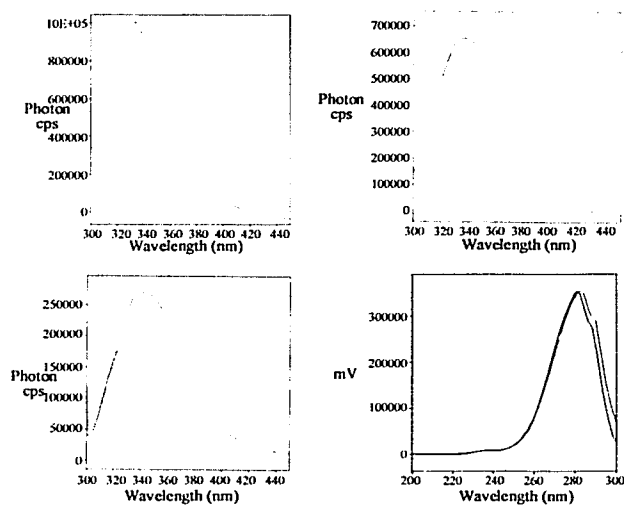
Figure 6.24. Fluorescence emission spectra of W262C β-glucosidase (blue) at:
A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation
and D) excitation spectra obtained at 339 nm emission.
Fluorescence intensities for emission spectra were in photon
counts per second (cps). Intensity for the excitation spectra was
determined in millivolts (mV). All spectra were collected at 25°C.
Spectra for W262C β-glucosidase (blue) and for the wild type β-
glucosidase (pink) are shown. A_{280} readings were obtained with the
concentrated enzymes. The enzymes were then diluted to give an
A_{280} of 0.029.
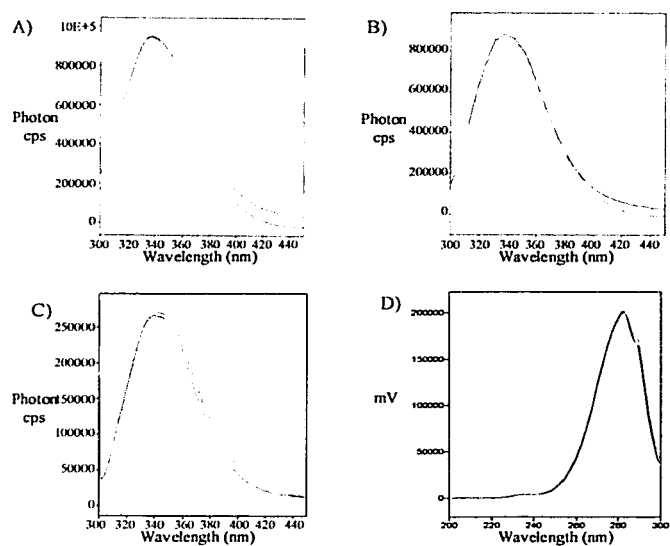
Figure 6.25. Fluorescence emission spectra of W262F β-glucosidase (blue) at: A)
275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D)
excitation spectra obtained at 339 nm emission. Fluorescence
intensities for emission spectra were in photon counts per second
(cps). Intensity for the excitation spectra was determined in millivolts
(mV). All spectra were collected at 25°C. Spectra for W262F β-
glucosidase (blue) and for the wild type β-glucosidase (pink) are
shown. A_{280} readings were obtained with the concentrated enzymes.
The enzymes were then diluted to give an A_{280} of 0.029.
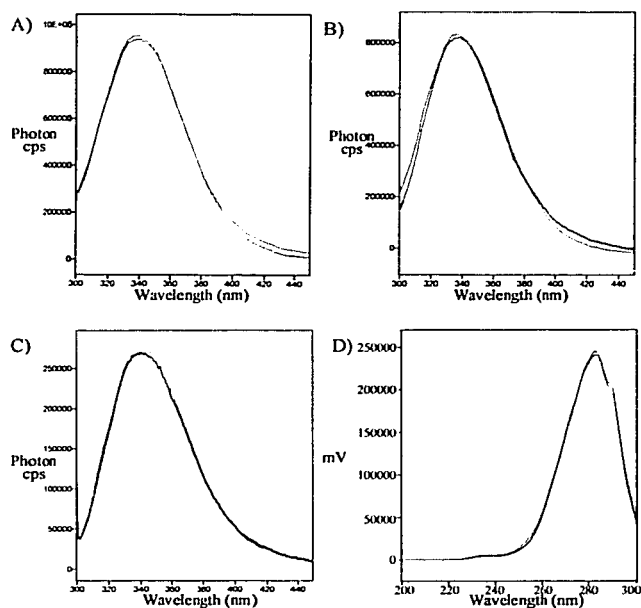
Figure 6.26. Fluorescence emission spectra of W262L β-glucosidase (blue) at:
A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation
and D) excitation spectra obtained at 339 nm emission. Intensities
for emission spectra were in photon counts per second (cps).
Intensity for the excitation spectra was determined in millivolts
(mV). All spectra were collected at 25°C. Spectra for W262L β-
glucosidase (blue) and for the wild type β-glucosidase (pink) are
shown. A_{280} readings were obtained with the concentrated
enzymes. The enzymes were then diluted to give an A_{280} of 0.029.
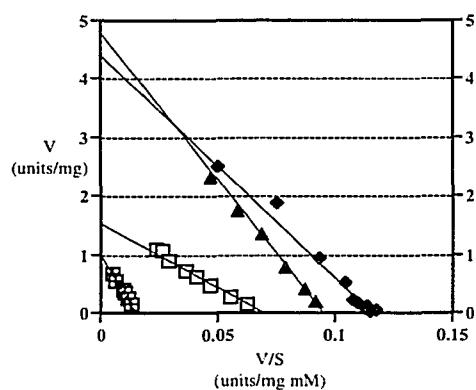
Figure 6.27. Eadie-Hofstee plots of the β-glucosidases with substitutions for Trp-262 with pNPGlc, pH 4.5, 25°C. The wild type enzyme is not shown so that the substituted enzymes may be seen more clearly. (◆) W262A, (□) W262C, (▲) W262F, and (⊞) W262L. An enzyme concentration of 1.4 μg/mL was used for all of the above assays. One unit is equivalent to 1 μmol of product produced per min.

times smaller) than the V_{max} values for the wild type enzyme. W262F β-Glucosidase had

the largest V_{max} values while W262L β-glucosidase had the smallest. The other very

significant difference was that the K_m values were much higher than those of the wild type

β-glucosidase (Table 6.8).

Table 6.8. Kinetic constants obtained with β-glucosidases with substitutions for Trp-262.

Substrate	Wild Type		W262A		W262C		W262F		W262L	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
pNPGlc	57	1.0	4.4	40	1.4	27	4.7	50	1.0	73
oNPGlc	60	3.0	1.4	84	1.1	40	6.0	65	0.4	21

[a] V_{max} values are reported in μmol/mg/min.
[b] K_m values are reported in mM.

6.5.2.1.1 pH profiles

The enzymes with substitutions for Trp 262 were also assayed for kinetic changes

in response to changes in pH (Figure 6.28). In general, the V_{max} values obtained for the

substituted enzymes followed similar trends as were observed for the wild type enzyme

except, of course, that they were much smaller. The pH optimum for the V_{max} of each

enzyme was at about pH 4 to 4.5. The K_m profiles for W262A and W262F β-glucosidase

were somewhat similar to that of wild type except, of course, they were much larger. The

K_m was high at lower pHs and then decreased as the pH approached 7. W262C and

W262L β-glucosidase were different. The K_m values for these substituted enzymes

changed very little with pH. The V_{max}/K_m values varied with pH in a manner somewhat

similar to the wild type enzyme except that the values were much smaller and that the

Figure 6.28. pH profiles of the W262 substituted enzymes with pNPGlc
(25°C). The data for the enzymes are shown in parts (A), (B), and
(C), while the corresponding wild type plots areshown in (D), (E),
and (F), respectively. Plots (A) and (D) show the change in V_{max}
with pH. Plots (B) and (E) show the change in K_m with pH. Plots
(C) and (F) show the change in the V_{max}/K_m with pH. The
enzymes are represented by (◆) W262A, (□) W262C, (▲)
W262F, (⊞) W262L and (○) for wild type. All kinetic values
were obtained from Eadie-Hofstee plots. All of the reactions were
done using 1.4 µg/mL enzyme at 25°C.
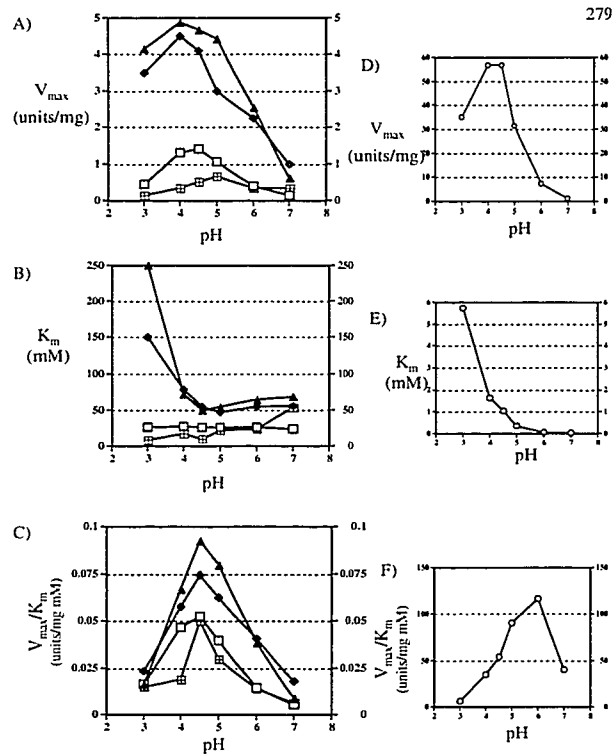
maxima were at about pH 4.5 rather than 6.0 as it is for wild type β-glucosidase. The

V_{max}/K_m values are proportional to k_c/Ks and the different maxima may represent

differences in their values as a function of pH. Studies to determine if this is the case

would require rapid flow techniques and that was beyond the scope of these studies.

6.5.2.2 Natural substrates

The β-glucosidases with substitutions for Trp-262 were also assayed with the

natural substrates (Table 6.9). As was the case with the aryl-glucosides, all of the lines on

the Eadie-Hofstee plots were linear with each substituted enzyme with each substrate.

Table 6.9. A comparison of kinetic values obtained from reaction of wild type and W262 substituted β-glucosidases with natural substrates.

Substrate	Wild type		W262A		W262C		W262F		W262L	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
Cellobiose	78	2.9	3.1	22	0.3	28	3.2	40	1.5	127
Cellotriose	94	0.7	0.5	2.9	0.7	2.6	1.9	2.3	0.3	2.7
Cellotetraose	66	0.6	0.6	1.3	0.6	2.2	1.7	1.3	0.12	8.6
Cellopentaose	13	0.6	0.4	3.9	0.3	6.1	0.8	2.8	0.5	6.0
Gentiobiose	46	2.2	1.3	42	1.6	167	2.2	82	1.6	155

[a] V_{max} values are reported in μmol/mg/min.
[b] K_m values are reported in mM.

The Eadie-Hofstee plots for the substituted enzymes with cellotetraose are shown in

Figure 6.29. The V_{max} values with these substrates were again low. The greatest

differences in K_m values for the glucosyl substrates relative to wild type were for

cellobiose and gentiobiose. The K_m values for gentiobiose were especially large. The K_m

A)

B)
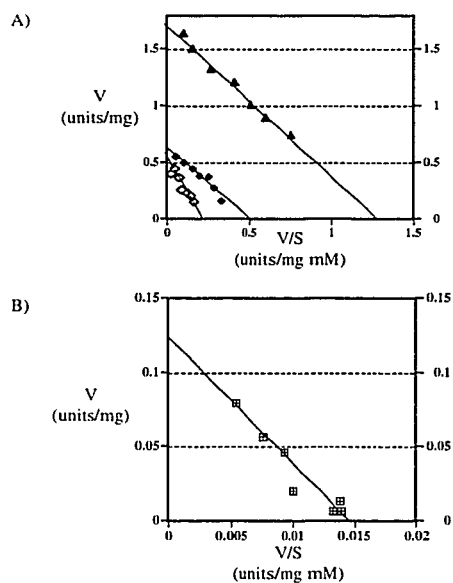
Figure 6.29. Eadie-Hofstee plots illustrating the reactions of the substituted
β-glucosidases with cellotetraose. A) The plots for W281A (■),
W281C (◇), and W281F (o). B) The plot for W281L (△). All
reactions were done at 25°C, pH 4.5 with 8 μg/mL enzyme. One
unit is equivalent to the production of 1 μmol per min. of
product.

values for the longer substrates were much lower and were closer to the wild type enzyme values. This is significant towards an understanding of what may be occurring with these substituted enzymes. In general, the highest K_m values were observed with W262L β-glucosidase and the lowest were seen with W262A β-glucosidase.

6.5.2.3 Competitive inhibitor studies

Relative to wild type, the competitive inhibition of the enzymes with substitutions for Trp 262 by glucose, cellobiose, gentiobiose, and gluconolactone was very poor (Table 6.10). Cellotriose was the best inhibitor, but it was still less effective with the Trp-262 substituted β-glucosidases than it was for wild type.

Table 6.10 K_i values[a] determined for the W262 substituted β-glucosidases.

Inhibitor	Wild type	W262A	W262C	W262F	W262L
Glucose	2.7	470	220	220	550
Mannose	220	6900	6400	2100	2200
Allose	290	295	280	515	390
Galactose	280	940	1100	1900	860
Xylose	670	920	480	2400	740
Gluconolactone	0.4	46	5.4	18	210
Cellobiose	3.4	83	170	93	160
Gentiobiose	4.8	191	290	200	125
Maltose	270	315	190	195	260
Cellotriose	0.4	2.9	6.8	2.1	ND

[a] K_i values are reported in mM.
ND not determined

The inhibition by mannose and galactose was affected to a lesser extent. Maltose, allose,

and xylose inhibited the substituted enzymes to about the same extent as they inhibited the

wild type enzyme.

6.5.3. Analysis of substrate breakdown by TLC

The substituted β-glucosidases were reacted with 50 mM cellobiose and the

products of the reaction were analyzed by TLC (Figure 6.30). In every case, the

trisaccharide product, gentiobiosyl-glucose, was detected first (usually within the first 10

min. of the reaction). Glucose could only be observed 20 to 60 minutes later. However,

there should be a glucose produced every time that a trisaccharide is synthesized (see

mechanism, Figure 4.20). When charring a TLC plate, a spot produced by a trisaccharide

should be 3 times as intense as one by the same concentration of glucose. So, in

comparison, it would be hard to see the glucose. These data show that the only reaction

occurring for each of these enzymes is trisaccharide production and one glucose is

produced for every trisaccharide. Larger oligosaccharide products were observed with the

β-glucosidases substituted with Ala, Cys and Phe. This only happened if the trisaccharide

had attained a high enough concentration. W262L β-Glucosidase would also probably

have produced larger sugars if the reaction had been monitored for a longer time. The

formation of products with more glucose did not occur nearly as readily with the wild type

enzyme. TLC was also completed for all of the substituted β-glucosidases with 10 mM

cellobiose. The results for W262F β-glucosidase are shown in Figure 6.31 (A). Even at

this low concentration of cellobiose the substituted enzyme produced roughly equal

amounts of trisaccharide and glucose (when differences in amount of intensity are taken

Figure 6.30.Thin layer chromatographic analysis of the reaction of the β-glucosidase with substitutions for Trp-262 with 50 mM cellobiose at 25°C. A) W262A, B) W262C, C) W262F and D) W262L. Enzyme was added (8 μg/mL final concentration) to the reaction mixture at zero time and an aliquot was immediately spotted onto the plate. Spotting was repeated at the times indicated in minutes on the bottom of the plate. The chromatography solvent used was n-butanol:ethyl acetate:2-propanol:acetic acid :water in the ratio of 1:3:2:1:1. The plates were developed twice, and the plates were dried between solvent elutions. Products were detected by dipping the dried plates in 2% sulphuric acid in methanol and charring. The standards used on the plates were cellobiose (C), trisaccharide (T), gentiobiose (Ge) and glucose (G). In (A) and (B) the concentrations of standards used was 10 mM cellobiose and 5 mM for glucose, gentiobiose and trisaccharide. In (C) and (D) the concentrations of the standards were all 10 mM.

285

A)

C 0 5 10 20 30 45 60 90 120 180 240 300 360 420 480 T Ge G
(min)

B)

C 0 5 10 20 30 45 60 90 120 180 240 300 360 420 480 T Ge G
(min)

C)

C 0 5 10 20 3u 45 60 90 120 180 240 300 360 420 480 T G Ge
(min)

D)

C 0 5 10 20 30 45 60 90 120 150 180 240 300 360 420 480 T G Ge
(min)
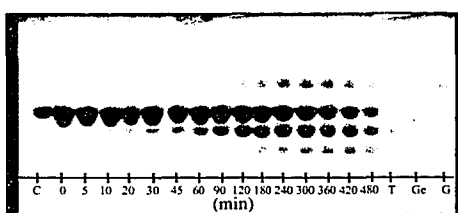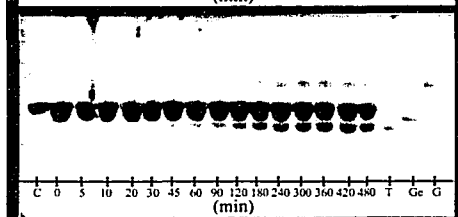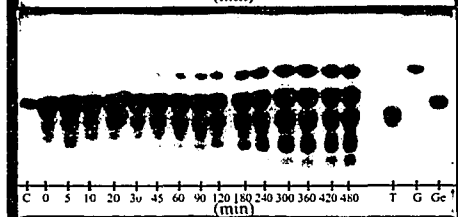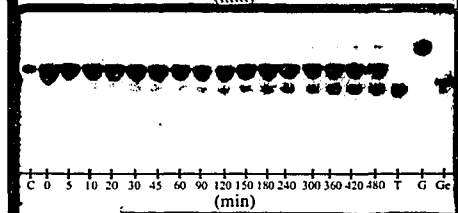
Figure 6.31.Thin layer chromatographic analysis of the W262F β-glucosidase

reaction. The substituted enzyme was reacted with 10 mM

cellobiose (A), 10 mM pNPGlc (B), and 50 mM pNPGlc (C) at

25°C over 8 hrs. Enzyme was added (8 μg/mL final concentration

for cellobiose and 1.4 μg/mL final concentration for pNPGlc) to

the reaction mixture at zero time and an aliquot was immediately

spotted onto the plate. Spotting was repeated at the times indicated

in minutes on the bottom of the plate. The chromatography solvent

used was n-butanol:ethyl acetate:2-propanol:acetic acid :water in

the ratio of 1:3:2:1:1. The plate shown in (A) was developed

twice. The plate was dried between developments. The plates

shown in (B) and (C) were only developed once. Products were

detected by dipping the dried plates in 2% sulphuric acid in

methanol and charring. The standards used on the plates were 10

mM cellobiose (C), 10 or 50 mM pNPGlc (P), 10 mM

trisaccharide (T), 10 mM gentiobiose (Ge) and 5 mM glucose

(G)in (A) and 10 mM glucose in (B) and (C).

287
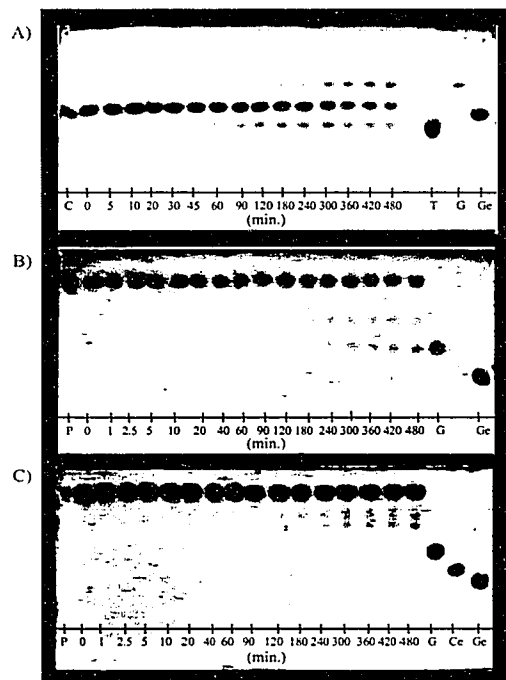
into account). Thus even at 10 mM cellobiose, only transglucosidic reactions are occurring. No gentiobiose was produced for any enzyme (Figures 6.30 A,B,C, D and 6.31 A). This is probably because of the small amount of glucose that was present (even after 8 hrs.).

TLC was also done with all of the β-glucosidases with substitutions for Trp 262 with pNPGlc (10 and 50 mM) to analyze the resultant products. The reaction of the enzymes with pNPGlc is a better way of showing the amount of hydrolysis that occurs relative to transglucosylation because, with pNPGlc, glucose is only a product of hydrolysis. When only transglucosylation occurs only nitrophenol (and the newly synthesized larger sugar) should be released. The TLC plates for W262F β-glucosidase are shown in Figure 6.31 (B) and (C) for 10 and 50 mM pNPGlc, respectively. At 10 mM pNPGlc, quite a lot of pNPGe (it migrates below pNPGlc but above glucose) was produced. The identity of this is known because it interacts with the fluor on the TLC plates and fluoresces under UV light. Additionally, its R_f value is equal to the R_f value known for pNPGe (see Section 4.6.1). In the presence of 10 mM pNPGlc a small amount of glucose was produced, which suggests that a small amount of hydrolysis did occur, although the main reaction was obviously transglucosylation. In the presence of 50 mM pNPGlc there was no glucose on the TLC plate. On the 50 mM pNPGlc plate there was only pNPGe and an additional band which is probably nitrophenyl trisaccharide. There was no detectable glucose. Therefore, very little hydrolysis seems to have occurred. The TLC plate results obtained with the other β-glucosidases with substituted for Trp 262 with

10 and 50 mM pNPGlc were very similar to those for W262F-β-glucosidase (data not shown).

6.5.4 Gas-Liquid Chromatography

Because the substitutions for Trp-262 greatly reduced the ability of the enzymes to react with cellobiose, it was attempted to match the glucose production rates between the enzymes with substitutions for W262 rather than compare them directly with the wild type enzyme. Aliquots from the reaction mixtures were also collected over a much longer period of time (~48 hrs). GLC was done at cellobiose concentrations of 5 and 50 mM.

Glucose was produced at a fairly constant rate by all of the substituted enzymes with 5 mM cellobiose (Figure 6.32(A) and (B)). This was especially true for the first 120 min. (Figure 6.32 (B)). The substituted enzymes also showed that gentiobiose was produced after a delay (Figure 6.32 (C) and (D)). However, with GLC unlike with TLC, it could be detected. The amounts of gentiobiose produced were similar for W262A, W262C and W262F β-glucosidase but W262L β-glucosidase did not produce any measurable gentiobiose. A very significant amount of the products, especially at early times (See Figure 6.32 (E) and (F)), was the trisaccharide. In the case of the wild type enzyme, the amount of trisaccharide produced with 5 mM cellobiose was hard to detect and was usually less than the error. Considering that one glucose is produced for every trisaccharide, the transglucosidic reaction even at this low cellobiose concentration made up over was over 50% of the product at very early times for at least one of the substituted enzymes (Figure 6.33). Trisaccharide could not even be detected with wild type. The maximum amount of the product was present just before the 1000 min. time point of

Figure 6.32. Results of GLC analysis of the β-glucosidases with substitutions
for Trp-262 with 5 mM cellobiose. The products that resulted
from the reaction were: glucose (A) and (B), gentiobiose (C) and
(D), and gentiobiosyl-glucose (E) and (F). The plots in parts (A),
(C) and (E) show the change in the concentration of product over
the entire reaction time. The plots in parts (B), (D), and (F) show
the first 120 min of product accumulation only. W262A (◆),
W262C (□), W262F (▲) and W262L (⊞). The substituted
enzymes concentrations were chosen so that they all had
essentially the same rates of glucose production. All reactions
were done at 25°C, pH 4.5.

A)

B)

C)

D)

E)

F)

292
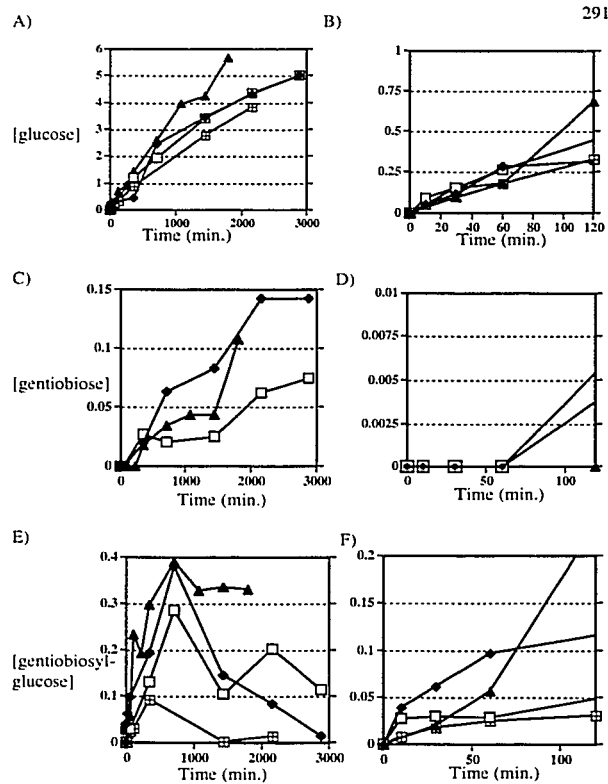
Figure 6.33. A comparison of the ratio of trisaccharide production to glucose
production by the Trp-262 substituted enzymes with 5 mM
cellobiose. Both the trisaccharide produced and the glucose
produced at early times were near to the GLC detection limit, and
thus there is some error. However, it is obvious that a significant
proportion of the product was trisaccharide when compared to the
fact that no trisaccharide is produced by wild type at this
cellobiose concentration. The complete reaction time is shown in
(A) and the first 120 min. are shown in (B). W262A (◆), W262C
(□), W262F (▲), W262L (▣) and wild type (O), which is only
shown in (B). The amounts of both trisaccharide and glucose are
in mM. The substituted enzymes concentrations were chosen so
that they all had the same rate of glucose production, the wild
type concentration was 2 μg/mL. All reactions were done at 25°C,
pH 4.5. The amounts of products here were quite small and thus
subject to some error. The error is even greater when one obtains
a ratio. However, it is obvious that the ratios are much larger (B)
for the substituted enzymes than for the wild type enzyme.

A)

Trisaccharide (mM)
Glucose (mM)

Time (min.)

B)

Trisaccharide (mM)
Glucose (mM)

Time (min.)
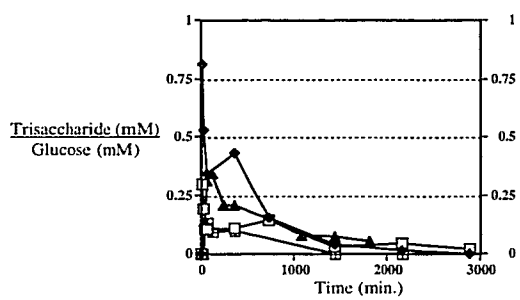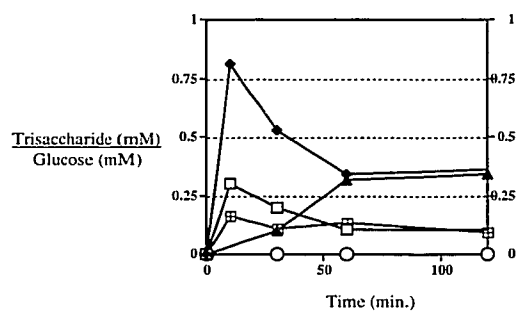
reaction and then the trisaccharide was either degraded or was used as an acceptor to produce the larger oligosaccharide (this was not studied).

When the substituted enzymes were reacted with 50 mM cellobiose the results shown in Figure 6.34 were observed. The glucose production rates of the W262F β-glucosidase were slightly higher than the other enzymes and so this must be noted when examining the data. All of the enzymes produced small but detectable amounts of gentiobiose (note that TLC was not sensitive enough to detect the small amount of gentiobiose produced). In particular, W262A β-glucosidase produced a significant amount of gentiobiose (Figure 6.34 (C)). Again, however, the production occurred after a time lag (Figure 6.34 (D) and (E)) and only after a significant amount of glucose was present. A very large portion of the reaction was trisaccharide production (Figure 6.34 (E) and (F)). At early times for each substituted enzyme except W262F β-glucosidase, the amount of trisaccharide was roughly equal to the amount of glucose produced. For W262F-β-glucosidase it was just a little lower. When it is realized that one glucose is produced every time that a trisaccharide is produced, this means that for each enzyme except W262F β-glucosidase no hydrolysis occurs when the substrate concentration is high. The only reaction that occurs is transglucosylation. Trisaccharide production did plateau in each case. This was probably because the trisaccharide was converted to higher oligosaccharides or because it was degraded back to glucose (Figure 6.34 (E)). The ratio of trisaccharide produced is compared to the amount of glucose produced in Figure 6.35. This clearly shows that initially only transglucosylation is occurring for the W262A,

Figure 6.34. Results of GLC analysis of the enzymes with substitutions for
Trp-262 with 50 mM cellobiose. The products that resulted from
the reaction were glucose (A) and (B), gentiobiose (C) and (D),
and gentiobiosyl-glucose (E) and (F). In this case, at short
reaction times, the amounts of gentiobiose were just a little
greater than the detection limits and thus there was some error.
The plots in parts (A), (C) and (E) show the change in the
concentration of product over the entire reaction time. The plots
in parts (B), (D), and (F) show the first 120 min of product
accumulation only. W262A (♦), W262C (□), W262F (▲) and
W262L (⊞). The substituted enzyme concentrations were chosen
so that they all had essentially the same rates of glucose
production. All reactions were done at 25°C, pH 4.5.
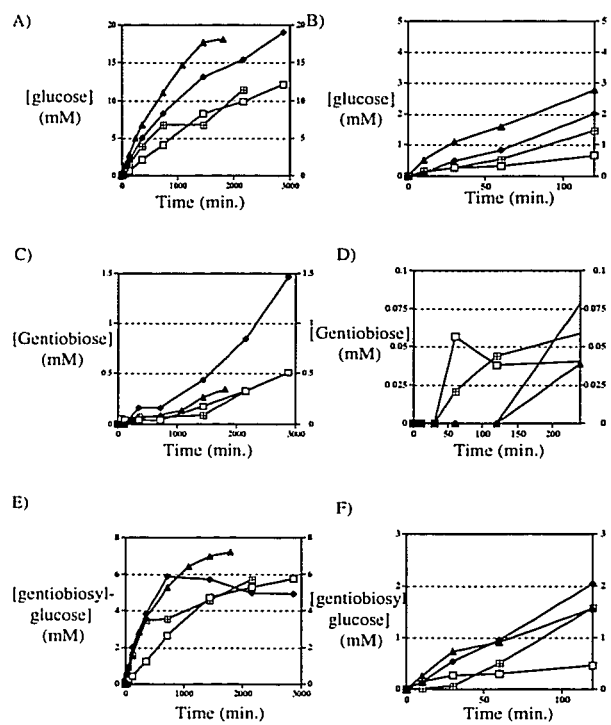
Figure 6.35. A comparison of the ratios of trisaccharide production to glucose
production by the enzymes having substitutions for Trp-262 with
50 mM cellobiose. The complete reaction time is shown in (A)
and the first 240 min. are shown in (B). W262A (◆), W262C
(□), W262F (▲), W262L (⊞) and wild type (○), which is only
shown in (B). The amounts of both trisaccharide and glucose are
in mM. The concentrations of the substituted enzymes were
chosen so that they all had essentially the same rates of glucose
production. The wild type concentration was 2 μg/mL. All
reactions were done at 25°C, pH 4.5. The amounts of products
here were quite small (near the detection limit) at early times and
thus were subject to some error. The error is even greater when
utilizing ratios. However, it is noticeable that the ratios are larger
(B) for the substituted enzymes than for the wild type enzyme.

A)

$\dfrac{\text{[Trisaccharide]}}{\text{[Glucose]}}$

Time (min.)

B)

$\dfrac{\text{[Trisaccharide]}}{\text{[Glucose]}}$

Time (min.)
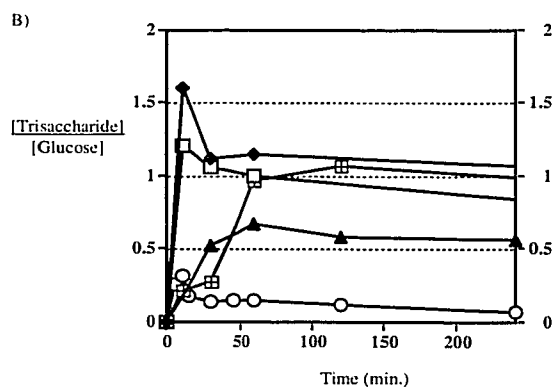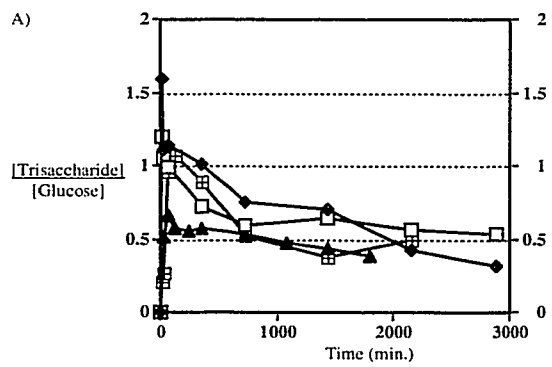
W262C and W262L β-glucosidases. The W262F β-glucosidase has much more transglucolytic activity than wild type but the rate is less than the other W262 substituted enzymes, which may result from the similarity of Phe to Trp. Comparison with the wild type ratio of trisaccharide production to glucose production, which was at about 0.25 or lower, illustrates the significant transglucosylation ability of these enzymes (Figure 6.35(B)).

6.6 Leucine 426

Leu 426 was substituted with an Ala, a Gly and a Val using degenerate primers and the Quik-Change mutagenesis kit. The plasmids with the desired mutagenesis were detected by sequencing. Sequencing was also done to show that there were not any substitutions at other positions. The plasmids were transformed into *Pichia*. Integration of the β-glucosidase gene into the *P. pastoris* genome was verified through PCR analysis of the genomic DNA, and isolation of the PCR product followed by sequencing showed that the desired substitutions were present. The three *P. pastoris* organisms with the substituted genes were able to grow normally and expressed the expected amount of enzyme (similar to wild type expression). The proteins were purified from the media. All of the enzymes with substitutions for Leu 426 eluted from the size exclusion column in the same volume as elution of the wild type enzyme, showing that there were no changes in the size and that the enzymes were dimers. SDS-PAGE was used to show that the enzymes were pure and had the same subunit size as did wild type.

6.6.1 Fluorescence studies

The emission spectra at excitation wavelengths of 275, 285 and 295 nm had peak maxima at the same wavelength as the wild type enzyme in all cases (Figures 6.36, 6.37, and 6.38). The excitation spectra, at an emission wavelength of 339 nm, were also similar to the wild type excitation spectrum (Figures 6.36, 6.37, and 6.38 part (D)). Each had a main peak at 282 nm with a shoulder at 288 nm.

A)

B)

C)

D)
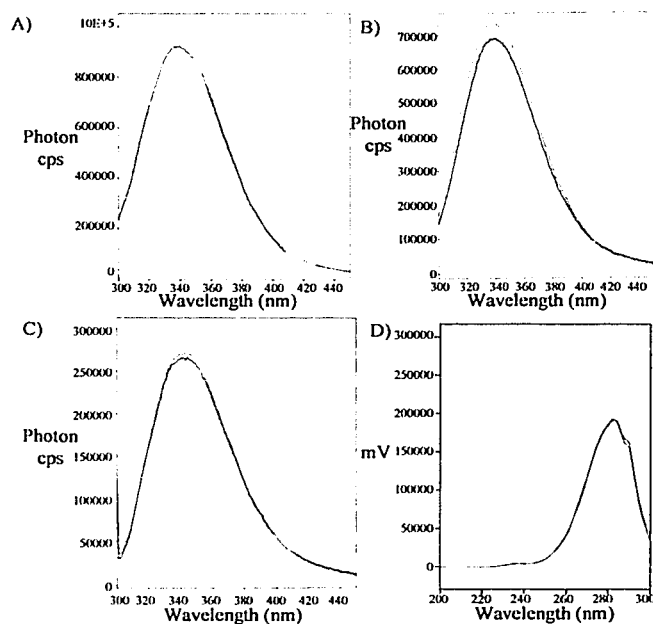
Figure 6.36. Fluorescence emission spectra produced by L426A β-glucosidase (blue) at A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were in photon counts per second (cps). Intensity for the excitation spectra was determined in millivolts (mV). All spectra were collected at 25°C. Spectra for L426A (blue) and for the wild type (pink) are displayed. A_{280} measurements of the concentrated enzymes were obtained and the enzymes were then diluted to a concentration that would give an A_{280} of 0.029.
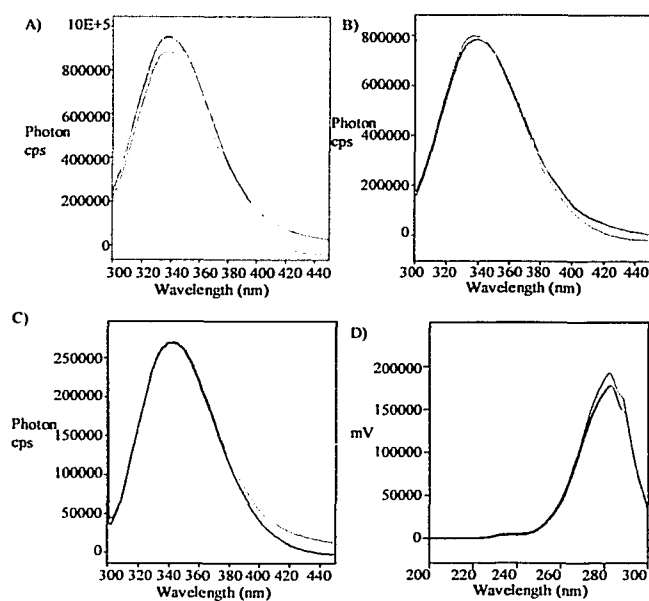
Figure 6.37. Fluorescence emission spectra produced by L426G β-glucosidase
(blue) at A) 275 nm excitation, B) 285 nm excitation, C) 295 nm
excitation and D) excitation spectra obtained at 339 nm emission.
Fluorescence intensities for emission spectra were in photon
counts per second (cps). Intensity for the excitation spectra was
determined in millivolts (mV). All spectra were collected at 25°C.
Spectra for L426G (blue) and for the wild type (pink) are
displayed. A_{280} measurements of the concentrated enzymes were
obtained and the enzymes were then diluted to a concentration
that would give an A_{280} of 0.029.

A)

Wavelength (nm)

B)

Wavelength (nm)

C)

Wavelength (nm)

D)

Wavelength (nm)
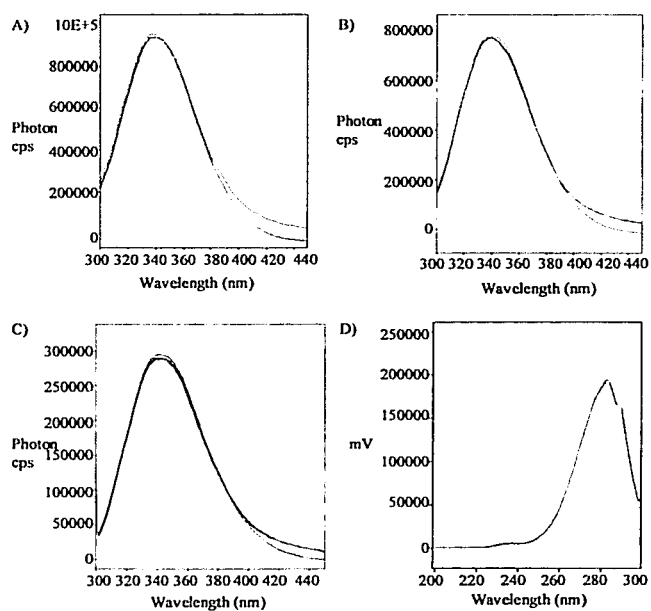
Figure 6.38. Fluorescence emission spectra produced by L426V β-glucosidase (blue) at A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were in photon counts per second (cps). Intensity for the excitation spectra was determined in millivolts (mV). All spectra were collected at 25°C. Spectra for L426V (blue) and for the wild type (pink) are displayed. A_{280} measurements of the concentrated enzymes were obtained and the enzymes were then diluted to a concentration that would give an A_{280} of 0.029.

6.6.2 Kinetic characterization

6.6.2.1 pNPGlc and oNPGlc

β-Glucosidases with substitutions for Leu 426 were assayed with pNPGlc and

oNPGlc. There were some differences from wild type but these were not dramatic (Table

6.11). L426G β-Glucosidase had V_{max} and K_m values similar to those with wild type with

both substrates. L426A and L426V β-glucosidases had V_{max} values with pNPGlc that were

about 2 times as large as with oNPGlc. This is different from wild type β-glucosidase,

where the V_{max} values for these two substrates were nearly the same. In addition, the K_m

values of the substituted β-glucosidases were similar to that of wild type for pNPGlc,

while they differed for oNPGlc (especially L426V β-glucosidase). All of the enzymes

with substitutions for Leu 426 had curved Eadie-Hofstee plots similar to wild type with

both substrates (Figure 6.39). This suggests that they have similar transglucosidic activity.

Table 6.11. Kinetic constants with pNPGlc and oNPGlc for wild type β-glucosidase and
the β-glucosidases with substitutions for Leu 426.

Substrate	Wild Type		L426A		L426G		L426V	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
pNPGlc	57	1.0	93	1.4	62	0.9	44	1.3
oNPGlc	60	3.0	47	1.9	49	2.7	20	1.2

[a] V_{max} values were recorded in units of μmol/mg/min.
[b] K_m values are in units of mM.

6.6.2.1.1 Effect of pH

Except for the magnitude, the pH profiles (Figure 6.40) for the substituted enzymes

were similar to those for the wild type enzyme. The Eadie-Hofstee plots curved
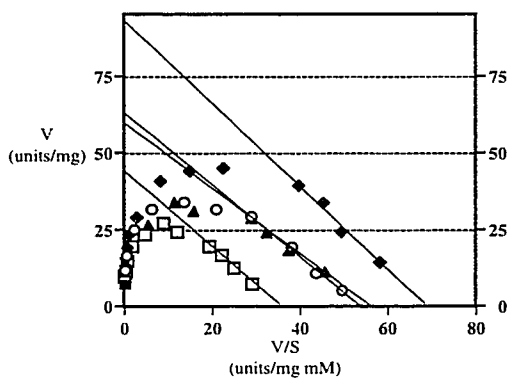
Figure 6.39. Eadie-Hofstee plots of the reaction of the L426 substituted
enzymes and the wild type β-glucosidase with pNPGlc at pH 4.5,
25°C. Data was acquired by the assay methods described in
Section 3.6.2. Wild type (O), L426A (◆), L426G (☐) and
L426V (▲). The lines illustrate the portion of the curves used to
determine the kinetic values in Table 6.11. All assays were
completed with an enzyme concentration of 1.4 μg/mL. One unit
is equivalent to the production of one μmol of product per min.

306

Figure 6.40. pH Profiles obtained with pNPGlc for β-glucosidases with
substitutions for Leu-426. (A) V_{max}, (B) K_m and (C) V_{max}/K_m. The
enzymes are represented by: (O) wild type, (◆) L426A, (□)
L426G and (▲) L426V. Eadie-Hofstee plots (at each pH) were
then used to determine the V_{max}, K_m and V_{max}/K_m values for each
plot. Assays were completed according to Section 3.6.5. The
enzyme concentration used for each assay was 1.4 μg/mL, and all
assays were completed at 25°C. One unit is equivalent to the μmol
of product produced per minute.

307
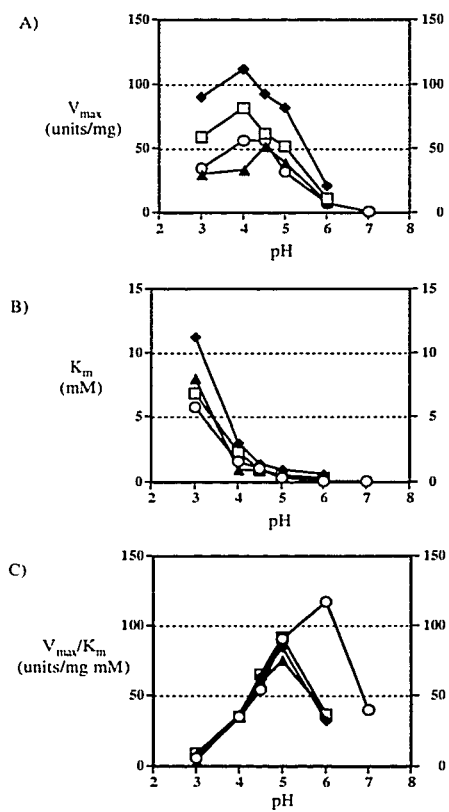
downwards at low pH and upwards at high pH as was the case with wild type enzyme. The V_{max}(trans) values were not determined but they appeared to be the same at each pH.

6.6.2.2 Reaction with natural substrates

The kinetic results of these assays are summarized in Table 6.12. The V_{max} values of most of the substituted enzymes were lower than the values obtained with the wild type enzyme. But each enzyme was still quite reactive. The K_m value with L426A β-glucosidase (with cellobiose) was higher than with wild type enzyme while both the L426G and L426V β-glucosidases had lower K_m values. The substituted enzymes' interactions with the larger cellooligosaccharides differed from the wild type enzyme. For wild type enzyme, the K_m (for the cellooligosaccharides) decreased significantly with increasing oligosaccharide size, with a large decrease between cellobiose and cellotriose. The K_m values increased in going from cellobiose to cellotriose and then decreased with increasing numbers of glucose for each of the substituted β-glucosidases. The relative changes to the V_{max} values for the substituted enzymes with the oligosaccharides of increasing length were somewhat similar to the changes with the wild type enzyme. The kinetic values (V_{max} and K_m) with gentiobiose were similar to those of the wild type enzyme except for L426V β-glucosidase. Its V_{max} value was approximately 8 times smaller than that of wild type β-glucosidase. The Eadie-Hofstee graphs had typical curved plots in the presence of all of the cellooligosaccharides with these enzymes. The curvatures were at substrate concentrations similar to the occurrences with wild type enzyme (see cellobiose example - Figure 6.41).
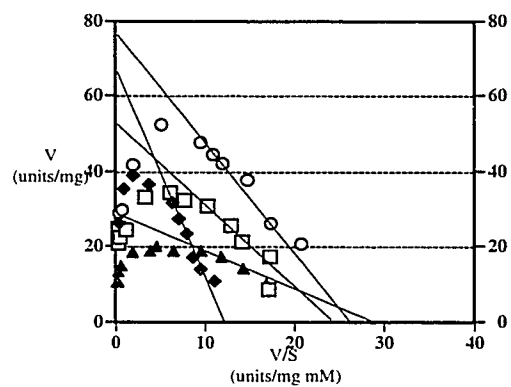
Figure 6.41. Eadie-Hofstee plot showing the decrease in hydrolytic rate (when transglycosylation occurs) in the presence of cellobiose with the enzymes with substitutions for Leu-426. Wild type (O), L426A, (◆), L426G (□) and L426V (▲). All reactions were done at 25°C, pH 4.5 with 8 µg/mL enzyme. One unit is equivalent to the production of 1 µmol per min. of product.

Table 6.12. Kinetic constant values obtained with natural glucosyl substrates.

Substrate	Wild Type		L426A		L426G		L426V	
	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]
Cellobiose	78	2.9	69	6.0	52	2.1	30	1.1
Cellotriose	94	0.7	66	4.8	37	2.5	72	4.9
Cellotetraose	66	0.6	69	3.6	33	1.4	61	2.7
Cellopentaose	13	0.6	21	1.2	10	0.2	17	1.7
Gentiobiose	46	2.2	30	2.3	42	3.3	6.0	1.4

[a] V_{max} values are reported in units of $\mu mol/mg/mL$.
[b] K_m values are in units of mM.

6.6.2.3 Inhibitor studies

The K_i values for the inhibition of the substituted enzymes by glucose and gluconolactone were similar to the K_i values of wild type β-glucosidase (Table 6.13). The substituted enzymes seemed to be inhibited a little better by the other inhibitors. This was especially true for L426V β-glucosidase. However, no consistent trend can be seen from these inhibition data.

Table 6.13. Ki values⁴ obtained with the Leu 426 substituted enzymes and a variety of inhibitors.

Inhibitor	Wild Type	L426A	L426G	L426V
Glucose	2.7	2.8	3.4	2.0
Mannose	220	240	155	98
Allose	290	362	140	140
Galactose	280	190	130	130
Xylose	670	450	600	220
Gluconolactone	0.4	0.3	0.3	0.2
Cellobiose	3.4	1.2	1.9	1.2
Gentiobiose	4.8	1.5	2.1	0.9
Maltose	270	110	110	73

⁴ All K_i values are in units of mM.

6.6.3 Analysis of substrate breakdown by TLC

The TLC plates of the reaction of the substituted β-glucosidases with cellobiose are shown in Figure 6.42. In each case both trisaccharide and glucose were produced at relatively early times. It appears that trisaccharide is present at higher amounts but it must be remembered that there are 3 glucose in the trisaccharide and thus the intensity will be 3 times as great when charred. Gentiobiose was produced later, when enough glucose had been produced to act as an acceptor for gentiobiose production. Gentiobiose was not seen in the case of L426V β-glucosidase but its presence was probably masked because the cellobiose and trisaccharide were not well separated.

Figure 6.42.Thin layer chromatography analysis of the reactions of the β-
glucosidases with substitutions for Leu-426 with 50 mM
cellobiose at 25°C. A) L426A, B) L426G, and C) L426V. Enzyme
was added (8 μg/mL final concentration) to the reaction mixture at
zero time and an aliquot was immediately spotted onto the plate.
Spotting was repeated at the times indicated in minutes on the
bottom of the plate. The chromatography solvent used was n-
butanol:ethyl acetate:2-propanol:acetic acid :water in the ratio of
1:3:2:1:1. The plates were developed twice. They were dried in
between solvent exposures. Products were detected by dipping the
dried plates in 2% sulphuric acid in methanol and charring. The
standards used on the plates were cellobiose (C), trisaccharide (T),
gentiobiose (Ge) and glucose (G). The concentrations of the
standards was generally 10 mM. The glucose standards in (A) and
(B) were 50 mM and in (B) the concentration of cellobiose and
gentiobiose standards was 5 mM. The plate in (B) was scratched
and this is apparent by the line interrupting the glucose spots.

313

A)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 150 180 Ge G
(min)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 150 Ge G
(min)

C)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 150 T Ge G
(min)
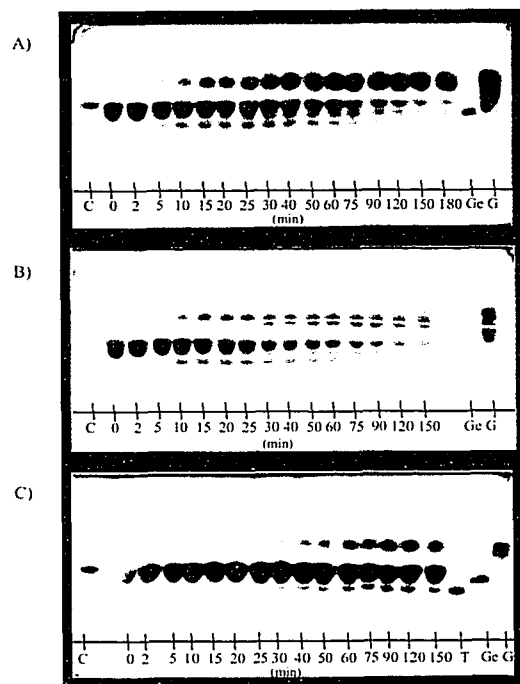
6.6.4 Gas-liquid chromatography

The reaction of the substituted enzymes with 5 mM cellobiose resulted in the normal glucose production profiles (Figure 6.43). As expected, gentiobiose or gentiobiosyl-glucose could not be detected at this low cellobiose concentration (data not shown). The reaction of the substituted enzymes with 50 mM cellobiose shows that the glucose production by the enzymes with substitutions for Leu 426 matched the glucose production of the wild type enzyme (Figure 6.44(B)) at early times. However, the glucose produced by the wild type enzyme began to plateau after about 25 mM glucose was produced and had plateaued when 50 mM glucose was produced. The substituted enzymes, however, continued to produce glucose and only began to plateau when about 50 mM glucose was produced. The production of glucose by the substituted enzymes only plateaued when higher amounts of glucose were produced. Examination of the first 2 hr of gentiobiose production suggests that the substituted β-glucosidases produced more gentiobiose than did wild type (Figure 6.44 (D)). Since more glucose is produced, this would be expected (gentiobiose production depends on the presence of glucose). All of the substituted enzymes produced gentiobiosyl-glucose at relative rates similar to that with wild type β-glucosidase (Figure 6.44 (E) and (F)). Similar amounts of gentiobiosyl-glucose were made and the presence of the trisaccharide was transient, as was the case with wild type β-glucosidase.
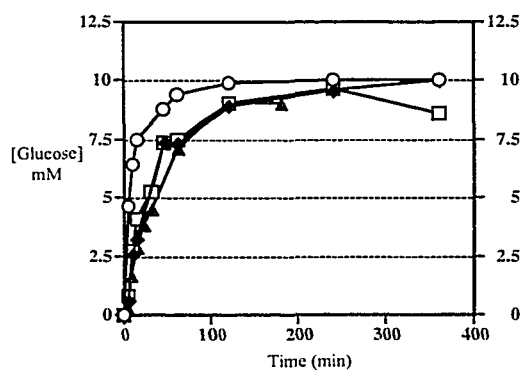
Figure 6.43. Analysis of products from the reaction of the wild type β-glucosidase and the β-glucosidase with substitutions for Leu-426 with 5 mM cellobiose over 6 hrs using GLC. Wild type (O), L426A (◆), L426G (□) and L426V (▲). The reactions were carried out at 25°C, at pH 4.5. The amount of substituted enzyme used was chosen so as to produce equivalent rates of glucose as 2 μg/mL of the wild type enzyme did. The experiments were completed as described in Section 3.8.1.

Figure 6.44. Results of GLC analysis of the enzymes with substitutions for
Leu-426 with 50 mM cellobiose. The products that resulted from
the reaction were: glucose (A) and (B), gentiobiose (C) and (D),
and gentiobiosyl-glucose (E) and (F). The plots in parts (A), (C)
and (E) show the change in the concentration of product over the
entire reaction time. The plots in parts (B), (D), and (F) show the
first 120 min of product accumulation only. Wild type (O),
L426A (◆), L426G (☐) and L426V (▲). The substituted
enzyme concentrations were chosen so that they had an
equivalent rate of glucose production as with 2 μg/mL wild type.
All reactions were done at 25°C, pH 4.5.

317

18
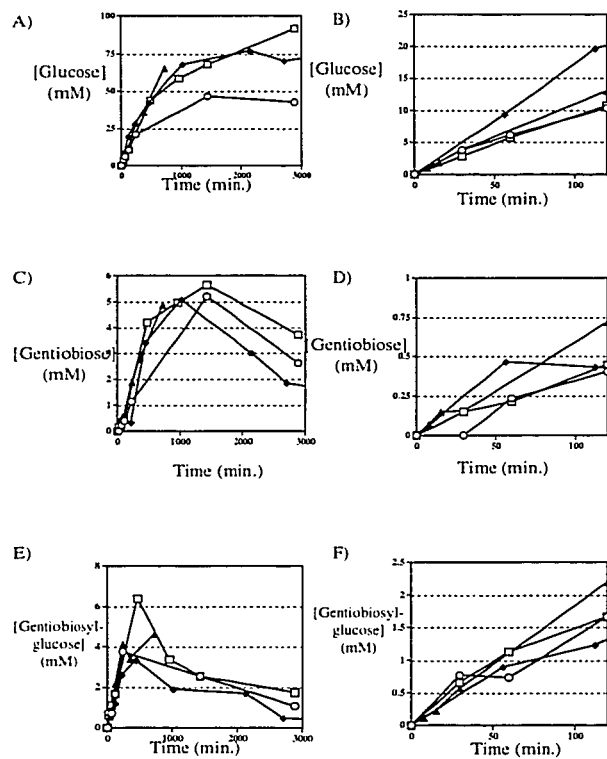
6.7 Tryptophan 430

Trp 430 was substituted with an Ala, a Phe and a Leu. All amino acid substitutions were accomplished with the Quik-Change mutagenesis kit. Selection of the plasmids that were mutated was done by sequencing and sequencing was also done to show that the substitutions did occur and that there were not any changes at other positions. The plasmids were transformed into *Pichia* and the proteins were expressed. Integration of the β-glucosidase gene with the desired mutations into the *P. pastoris* genome was verified through PCR analysis of the genomic DNA and isolation of the PCR product followed by sequencing. Purification of the substituted enzymes showed that they eluted from gel filtration columns in the same volume as elution of the wild type enzyme. This indicates that the enzymes were present as dimers of the same size as wild type. The three *P. pastoris* organisms with the substituted genes were able to grow normally and expressed the expected amount of enzyme (it was similar to wild type expression). The proteins were purified from the media. SDS-PAGE showed that the enzymes studied were pure and had the same subunit size. Only enzyme fractions that were >95% pure were used for the studies described below.

6.7.1 Fluorescence studies

Small shifts were observed in all of the emission spectra of the enzymes with substitutions for Trp 430 (Figure 6.45, 6.46 and 6.47). The shifts were so small that they are probably not significant. Excitation spectra were also obtained for these substituted enzymes (Figure 6.45, 6.46 and 6.47, part (D)). The spectra observed with W430A and
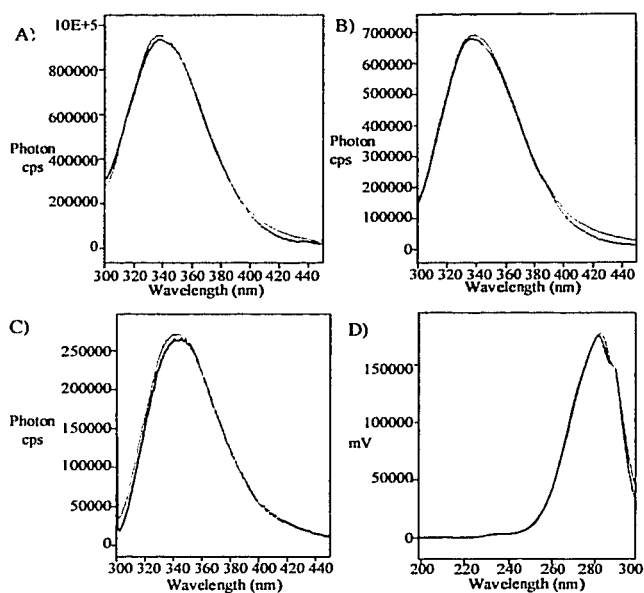
Figure 6.45. Fluorescence emission spectra of W430A at A) 275 nm
excitation, B) 285 nm excitation, C) 295 nm excitation and D)
excitation spectra obtained at 339 nm emission. Fluorescence
intensities for emission spectra were determined by photon counts
per second (cps). Intensity for the excitation spectra was
determined by millivolts (mV). Spectra for W430A (blue) and for
the wild type (pink) enzyme are displayed. A_{280} readings were
obtained with concentrated enzyme and the enzymes were then
diluted the amount that would give an A_{280} of 0.029.

A)

B)

C)

D)
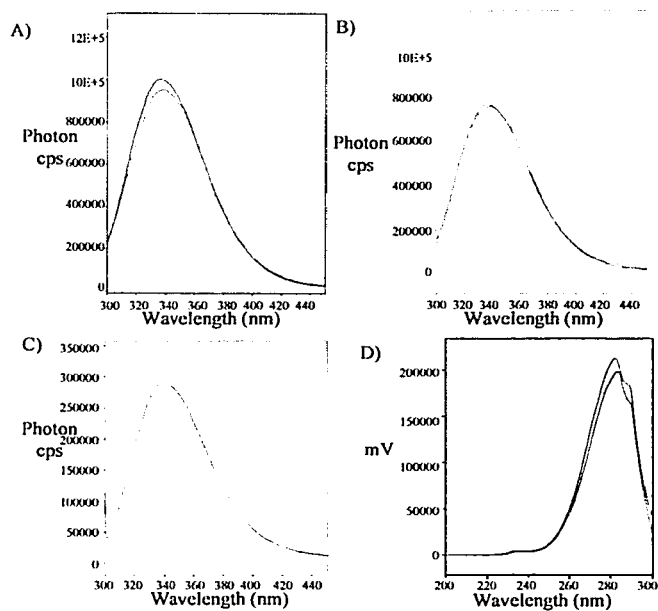
Figure 6.46. Fluorescence emission spectra of W430F at A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were determined by photon counts per second (cps). Intensity for the excitation spectra was determined by millivolts (mV). Spectra for W430F (blue) and for the wild type (pink) enzyme are displayed. A_{280} readings were obtained with concentrated enzyme and the enzymes were then diluted the amount that would give an A_{280} of 0.029.
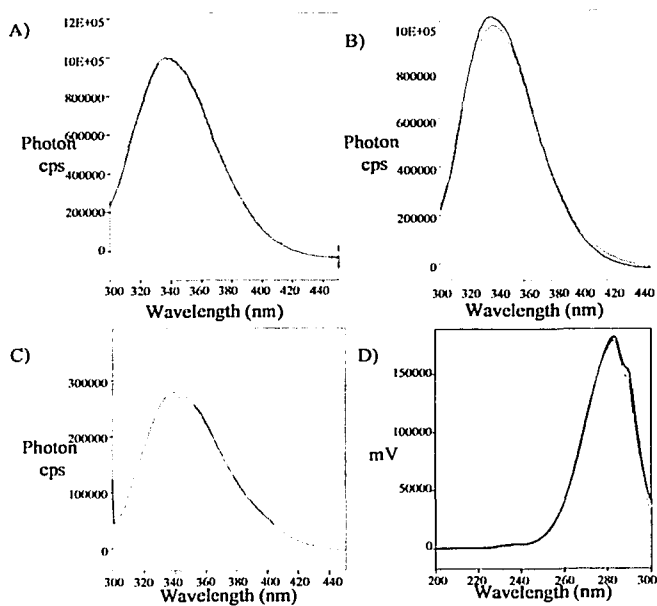
Figure 6.47. Fluorescence emission spectra of W430L at A) 275 nm excitation,
B) 285 nm excitation, C) 295 nm excitation and D) excitation
spectra obtained at 339 nm emission. Fluorescence intensities for
emission spectra were determined by photon counts per second
(cps). Intensity for the excitation spectra was determined by
millivolts (mV). Spectra for W430L (blue) and for the wild type
(pink) enzyme are displayed. A_{280} readings were obtained with
concentrated enzyme and the enzymes were then diluted the
amount that would give an A_{280} of 0.029.

W430L β-glucosidase were similar to the wild type spectrum. There was a small shift in the excitation spectrum of W430F β-glucosidase compared to the wild type spectrum.

6.7.2 Kinetic characterization

6.7.2.1 pNPGlc and oNPGlc

The kinetic constants obtained for the substituted enzymes with pNPGlc and oNPGlc at pH 4.5 are summarized in Table 6.14. The V_{max} and K_m values of W430F β-glucosidase were similar to those of the wild type enzyme. W430A and W430L β-glucosidase had lower V_{max} values than those observed for the wild type enzyme. The K_m values of W430A and W430F β-glucosidase were similar to those of wild type and only the substitution by Leu resulted in notable changes to the K_m. For both substrates the K_m values obtained were smaller than the values observed for the wild type enzyme and the other W430-substituted enzymes. All of the enzymes with substitutions for Trp 430 had the characteristic downward bends (Figure 6.48) that were found with wild type when there was transglucosidic activity.

Table 6.14. Kinetic constants with pNPGlc and oNPGlc for wild type β-glucosidase and the β-glucosidases with substitutions for Trp 430.

Substrate	Wild Type		W430A		W430F		W430A	
	V_{max}^a	K_m^b	V_{max}^a	K_m^b	V_{max}^a	K_m^b	V_{max}^a	K_m^b
pNPGlc	57	1.0	15	0.9	59	1.7	24	0.5
oNPGlc	60	3.0	43	2.6	61	3.5	29	1.1

[a] V_{max} values were recorded in units of μmol/mg/min.
[b] K_m values are in units of mM.
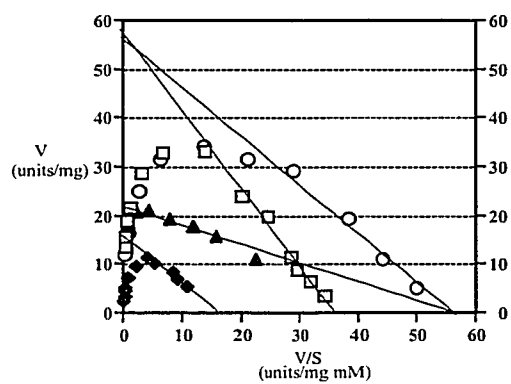
Figure 6.48 Eadie-Hofstee plots with pNPGlc generated with the wild type β-
glucosidase and the β-glucosidase with substitutions for Trp-430.
Wild type (O), W430A (◆), W430F (□) and W430L (▲). The
lines illustrate the portion of the curves used to determine the
kinetic values in Table 6.14. All assays were completed with an
enzyme concentration of 1.4 μg/mL at 25°C, pH 4.5. One unit is
equivalent to the production of one μmol of product per min.

324

6.7.2.1.1 pH effects

The change from downward slope at high substrate concentration to upward slopes was not observed for the enzymes with substitutions for Phe and Leu. Both of these substitutions resulted in a downward curving plot at each pH. The enzyme substituted with Ala showed the expected upward curving plot at pH 6 and higher. It is also interesting that W430F β-glucosidase had a V_m at pH 7 that was 5 times larger than that of the wild type enzyme.

The pH profiles (Figure 6.49) were not dramatically different from those for the wild type enzyme, despite the changes observed above. The profiles for W430L β-glucosidase were the most different.

6.7.2.2 Reaction with natural substrates

The Eadie-Hofstee plots for the reactions of the β-glucosidases with substitutions for Trp 430 with cellobiose are shown (Figure 6.50). In general the V_{max} values (Table 6.15) observed were lower than those seen with the wild type enzyme. The values tended to decrease with increasing substrate length. This trend was observed for all of the substituted enzymes except W430F β-glucosidase. The K_m values also decreased with increasing substrate length, although the values were generally higher than the K_m values of the wild type enzyme. Overall, no consistent trends were noted.

325

Figure 6.49. pH Profiles obtained with the β-glucosidases with substitutions
for Trp-430 with pNPGlc as the substrate. (A) V_{max} ;(B) K_m :(C)
V_{max}/K_m . Wild type (○), W430A (◆), W430F(◻) and W430L
(▲). Eadie-Hofstee plots (at each pH) were used to determine the
V_{max} and K_m values for each pH. Assays were completed
according to Section 3.6.5. The enzyme concentration used for
each assay was 1.4 µg/mL, and all assays were completed at 25°C.
One unit is equivalent to the µmol of product produced per
minute.

A)

B)

C)

A)

B)
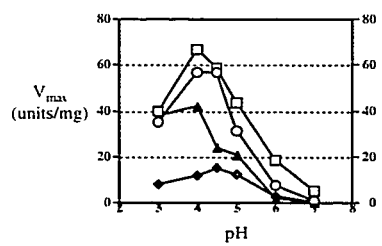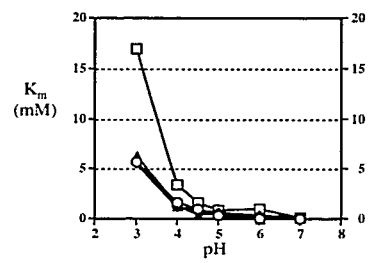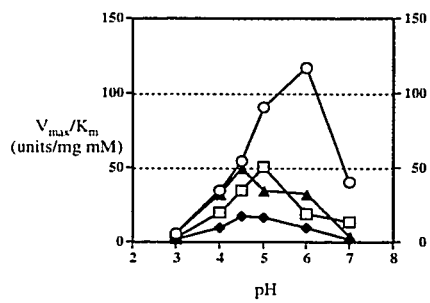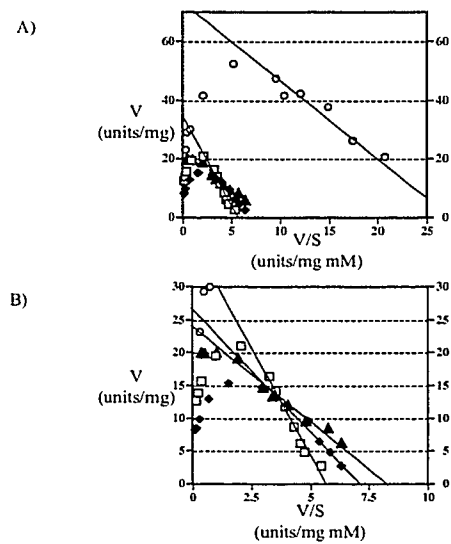
Figure 6.50. Eadie-Hofstee plots of the reaction of the wild type β-glucosidase
and the enzymes with substitutions for Trp-430 with cellobiose at
pH 4.5, 25°C. The graph in (A) is shown on a smaller scale in (B).
The assay was completed as described in Section 3.6.7. Wild type
(O); W430A (◆); W430F (□); W430L (▲). The enzyme
concentration used for all of the above assays was 8 μg/mL. One
unit is equivalent to one μmol of product produced per min. The
lines indicate the portions of the plot that were used to determine
the kinetic values in Table 6.15.

Table 6.15. Kinetic constant values obtained with natural glucoside substrates and the
W-430 substituted enzymes.

Substrate	Wild type		W430A		W430F		W430L	
	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]	V_{max} [a]	K_m [b]
Cellobiose	78	2.9	24	2.9	35	5.8	24	2.9
Cellotriose	94	0.7	12	0.6	63	4.6	22	1.8
Cellotetraose	66	0.6	10	1.5	41	1.6	16	1.0
Cellopentaose	13	0.6	5	1.7	20	1.0	10	1.27
Gentiobiose	46	2.2	15	1.8	50	5.7	8	1.8

[a] All V_{max} values are reported as μmol/mg/min.
[b] All K_m values are reported in mM.

6.7.2.3 Inhibitor studies

The competitive K_i values are listed in Table 6.16. In most cases the K_i values were

lower for the substituted enzymes than were the wild type enzyme values. The K_i values for

W430A β-glucosidase were especially low. The K_i value for cellobiose with W430F β-

glucosidase was much higher than expected. However, no consistent trends were noted.

Table 6.16. Inhibition constants (K_i) [a] for the β-glucosidases with substitutions for Trp 430
with a variety of sugars.

Inhibitor	Wild Type	W430A	W430F	W430L
Glucose	2.7	1.0	2.7	1.9
Mannose	220	72	170	135
Allose	290	124	230	390
Galactose	280	77	190	650

Xylose	670	89	430	505
Gluconolactone	0.4	0.7	0.2	0.2
Cellobiose	3.4	0.5	16.5	1.8
Gentiobiose	4.8	0.5	6.7	2.1
Maltose	270	170	86	435

ª All K_i values are reported in mM.

6.7.3 Analysis of substrate breakdown by TLC

The degradation of cellobiose by the substituted β-glucosidases occurred in a similar way as it did with the wild type enzyme. At low cellobiose concentrations (7 mM) the reactions were mainly hydrolytic (data not shown). At higher concentrations of cellobiose (50 mM) all of the enzymes produced oligosaccharide products (Figure 6.51). However, the concentrations of trisaccharide (gentiobiosyl-glucose) that were produced were much higher for the W430A and W430F β-glucosidases than for W430L ß-glucosidase. Oligosaccharides larger than trisaccharides were also synthesized in greater amounts by W430A and W430F β-glucosidase. There was no evidence of this for W430L β-glucosidase.

6.7.4 Gas-liquid chromatography

The β-glucosidases with substitutions for Trp 430 were investigated by GLC by following the degradation of 5 and 50 mM cellobiose at an enzyme concentration that produced approximately the same amount of glucose as did the wild type (2 μg/mL). W430L β-Glucosidase was further investigated to determine the cellobiose concentration required for the transglucosidic reactions to occur because the results obtained with that

330

Figure 6.51 Thin layer chromatography analysis of the reaction of the β-
glucosidase with substitutions for Trp-430 with 50 mM
cellobiose. A) W430A B) W430F and C) W430L. Enzyme was
added (8 μg/mL final concentration) to the reaction mixture at
zero time and an aliquot was immediately spotted onto the plate.
Spotting was repeated at the times indicated in minutes on the
bottom of the plate. The chromatography solvent used was n-
butanol:ethyl acetate:2-propanol:acetic acid :water in the ratio of
1:3:2:1:1. The plates were eluted twice (with drying in between).
Products were detected by dipping the dried plates in 2%
sulphuric acid in methanol and charring. The standards used on
the plates were cellobiose (C), trisaccharide (T), gentiobiose (Ge)
and glucose (G). The concentrations of the standards were all 10
mM except the cellobiose in (C) which was 5 mM and the
glucose in (C) which was 40 mM.

331

A)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 105 120 150 180 210 240 G Ge T
(min.)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90105120 150180 210 240 G Ge T
(min.)

C)

C 0 2 5 10 15 20 30 45 60 75 90 120150 180 240 Ge G
(min.)

332
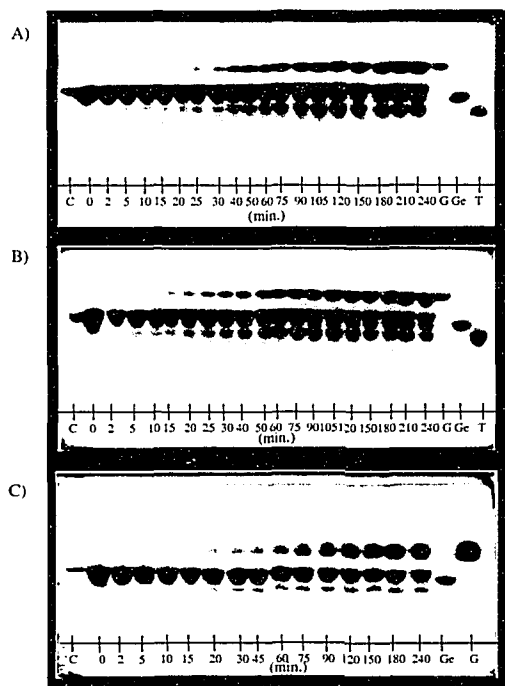
enzyme suggested that it was not as adept at transglucosidation as the other substituted enzymes.

The glucose production rate patterns for the β-glucosidases with substitutions for Trp 430 were similar to that of the wild type enzyme in the presence of 5 mM cellobiose (Figure 6.52). The amounts of gentiobiose and gentiobiosyl-glucose that were produced were barely above background levels and are not shown.

The normal delays in gentiobiose production occurred in the presence of 50 mM cellobiose (Figure 6.53). W430A and W430F β-glucosidases produced gentiobiose and trisaccharide at the same rates as did the wild type enzyme. The rates of gentiobiose and trisaccharide production by W430L β-glucosidase were, however, slower. The slow transglucosidic rate observed with W430L β-glucosidase was the only significant effect of the substitutions for Trp 430. Since the rates of gentiobiose and gentiobiosyl-glucose production by W430L β-glucosidase were relatively slow, preliminary studies of the initial rates of glucose and trisaccharide production were monitored at 6 concentrations of cellobiose. The preliminary results are plotted on a Michaelis-Menten graph in Figure 6.54(B). The plot illustrates that the rate of trisaccharide production only takes a large increase when the cellobiose concentration is at 50 mM. The increase occurs at about 15 - 30 mM with wild type β-glucosidase. The preliminary data also shows that when trisaccharide production begins to increase, there is a decline in hydrolysis. At 80 mM cellobiose the rate of trisaccharide production is quite a lot smaller than the rate of glucose production. For wild type, the rates were essentially equal at 80 mM cellobiose. These
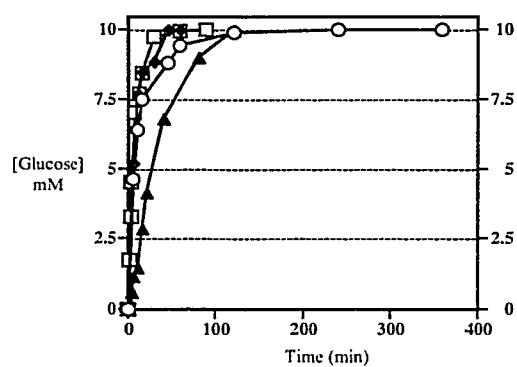
Figure 6.52. Analysis of glucose production by the reaction with wild type β-
glucosidase and the enzymes with substitutions for Trp-430 with
5 mM cellobiose over 6 hrs. Wild type (O); W430A (◆);
W430F (□); W430L (▲). The experiments were completed as
described in Section 3.8.1.

334

Figure 6.53. Analysis of the products of the reaction with wild type β-
glucosidase and the enzymes with substitutions for Trp-430 with
50 mM cellobiose over 6 hrs. A) glucose, B) gentiobiose, and C)
gentiobiosyl-glucose. The symbols used to represent the
different enzymes in all the graphs were (o) wild type; (◆)
W430A; (□) W430F; (▲) W430L. All reactions were done at
25°C, pH 4.5. The concentration of the β-glucosidases was
chosen so that they produced an equivalent amount of glucose as
was produced by the wild type enzyme when its concentration
was 2 μg/mL. The experiments were completed as described in
Section 3.8.1.

A) [Glucose] mM

B) [Gentiobiose] mM

C) [Gentiobiosyl-Glucose] mM
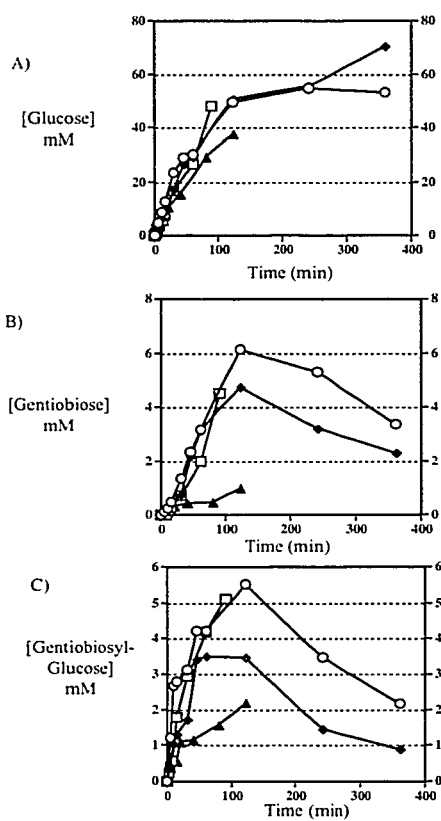
Figure 6.54. Michaelis-Menten plots for the wild type and β-glucosidases with
substitutions for W430L. The rates of glucose production (O)
and gentiobiosyl-glucose production (▲) are shown. All rates
were determined from individual plots of sugar production over
time at each cellobiose concentration. (A) wild type; (B) W430L.
The enzyme concentrations of the substituted β-glucosidases
were chosen because they had the same glucose production rate
as 10 μg/mL wild type enzyme. All reactions were done at 25°C,
pH 4.5. Experimental details are given in Section 3.8.1.

A)
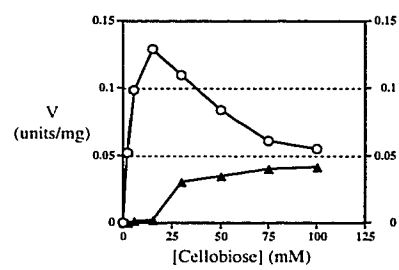
B)
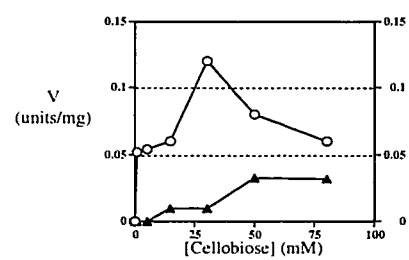
preliminary studies show that the effect of substitution of Trp 430 with Leu is on both the

rate of the transglucosidic reaction and on the binding of cellobiose to the acceptor site.

6.8 Tryptophan 551

Trp 551 was substituted by Phe and Ala. The plasmids with the desired mutated genes were detected by sequencing and the plasmids with the substituted codons were transformed into *P. pastoris*. PCR analysis of the yeast genome confirmed the integration of the β-glucosidase gene, and sequencing of the PCR product verified that the Trp 551 mutations were present. The yeast was able to grow normally and the substituted proteins were expressed and secreted into the media. The proteins were purified. The β-glucosidases with substitutions for Trp 551 eluted from the gel filtration columns in the same volume of buffer as was required to elute the wild type enzyme. This showed that the enzymes were the same size as wild type and that the enzymes were present as dimers. Analysis by SDS-PAGE showed that W551A and W551F β-glucosidase were pure and that the subunit sizes were the same as the wild type monomer weight. Only fractions that were >95% pure were used for the studies described below.

6.8.1 Fluorescence studies

Emission spectra were obtained for the enzymes with substitutions for Trp 551. The spectra of both enzymes were very similar to the spectra for the wild type enzyme at all three excitation wavelengths tested (275, 285, and 295). The spectra obtained are shown in Figure 6.55 for W551A β-glucosidase and in Figure 6.56 for the W551F β-glucosidase. The excitation spectra obtained for both enzymes with substitutions for Trp 551 were also similar to the excitation spectrum of the wild type enzyme (Figures 6.55(D) and 6.56(D)).

A)

Photon cps

300 320 340 360 380 400 420 440
Wavelength (nm)

B)

Photon cps

300 320 340 360 380 400 420 440
Wavelength (nm)

C)

Photon cps

300 320 340 360 380 400 420 440
Wavelength (nm)

D)

mV

200 220 240 260 280 300
Wavelength (nm)
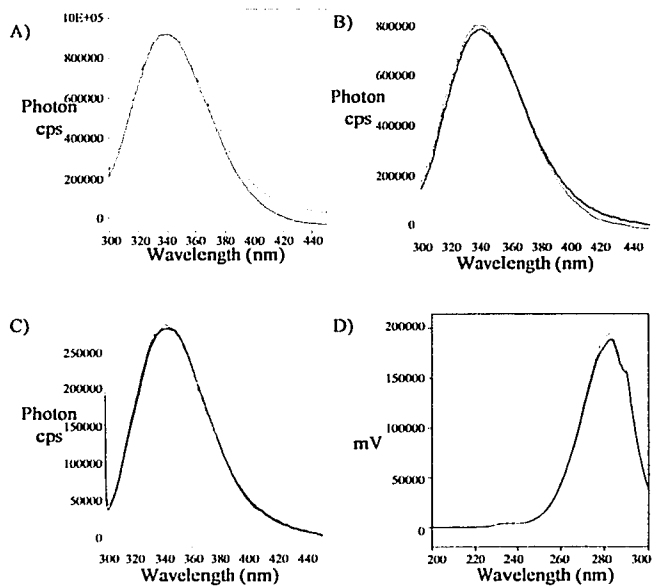
Figure 6.55. Fluorescence emission spectra of W551A β-glucosidase. A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were determined by photon counts per second (cps). Intensity for the excitation spectra was determined by millivolts (mV). All spectra were collected at 25°C. Spectra for W551A (blue) and for the wild type (pink) are displayed. A_{280} values of the concentrated enzymes were obtained and the enzymes were then diluted so that the A_{280} was 0.029.
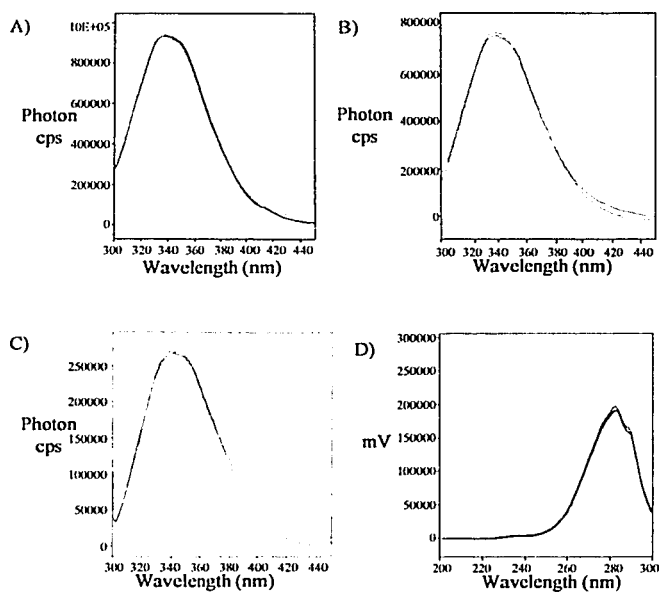
Figure 6.56. Fluorescence emission spectra of W551F β-glucosidase. A) 275 nm excitation, B) 285 nm excitation, C) 295 nm excitation and D) excitation spectra obtained at 339 nm emission. Fluorescence intensities for emission spectra were determined by photon counts per second (cps). Intensity for the excitation spectra was determined by millivolts (mV). All spectra were collected at 25°C. Spectra for W551F (blue) and for the wild type (pink) are displayed. A_{280} values of the concentrated enzymes were obtained and the enzymes were then diluted so that the A_{280} was 0.029.

6.8.2 Kinetic characterization

6.8.2.1 pNPGlc and oNPGlc

The K_m values of the β-glucosidases with substitutions for Trp 551 were similar to

those of the wild type enzyme (Table 6.17). The V_{max} values were, however, 2-3 times

smaller. The Eadie-Hofstee plots of the substituted enzymes had similar downward

curvatures as did the wild type enzyme, which suggested that they catalyzed

transglucosidic reactions (Figure 6.57).

Table 6.17. Kinetic constants for the wild type and the enzymes with substitutions for Trp
551 with pNPGlc and oNPGlc.

Substrate	Wild Type		W551A		W551F	
	$V_{max}{}^a$	$K_m{}^b$	$V_{max}{}^a$	$K_m{}^b$	$V_{max}{}^a$	$K_m{}^b$
pNPGlc	57	1.0	18	0.9	32	0.9
oNPGlc	60	3.0	25	2.1	30	2.9

[a] V_{max} values are reported in units of μmol/mg/min.
[b] K_m values are reported in mM.

6.8.2.1.1 pH profiles

The effect of pH was investigated from pH 3 to pH 7 with pNPGlc (Figure 6.58).

The V_{max} pH profile for W551A β-glucosidase was similar to the V_{max} profile for wild type

but the magnitude was lower. W551F β-Glucosidase had its maximum V_{max} value at pH 5

rather than at 4 - 4.5. Both substitutions resulted in K_m pH profiles that were similar to

that obtained with the wild type enzyme. As a result of the above differences, the V_{max}/K_m

profiles were slightly different for the substituted enzymes. The Eadie-Hofstee plots of the

enzymes curved downwards at low pH values and upwards at high pH values (similar to
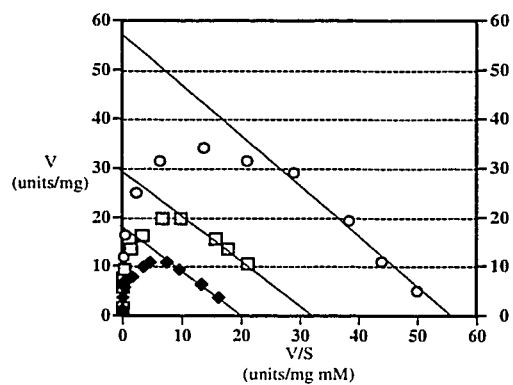
Figure 6.57. Eadie-Hofstee plot of the reaction of the β-glucosidases with
substitutions for Trp-551 with pNPGlc, pH 4.5, 25°C. Data was
acquired by the assay methods described in section 3.6.2. (O) wild
type; (◆) W551A ; and (□) W551F. The lines illustrate the
portion of the curves used to determine the kinetic values in Table
6.17. All assays were completed with an enzyme concentration of
1.4 μg/mL. One unit is equivalent to the production of one μmol of
product per min.

Figure 6.58. The effect of pH on β-glucosidase hydrolysis reaction (pNPGlc) with W551A (◆); W551F (□); wild type (○). The plots represent A) V_{max}. B) K_m. C) V_{max}/K_m. Eadie-Hofstee plots (at each pH) were used to determine the V_{max} and K_m values for each plot. Assays were completed according to Section 3.6.5. The enzyme concentration used for each assay was 1.4 μg/mL, and all assays were completed at 25°C. One unit is equivalent to the μmol of product produced per minute.

A)

B)

C)
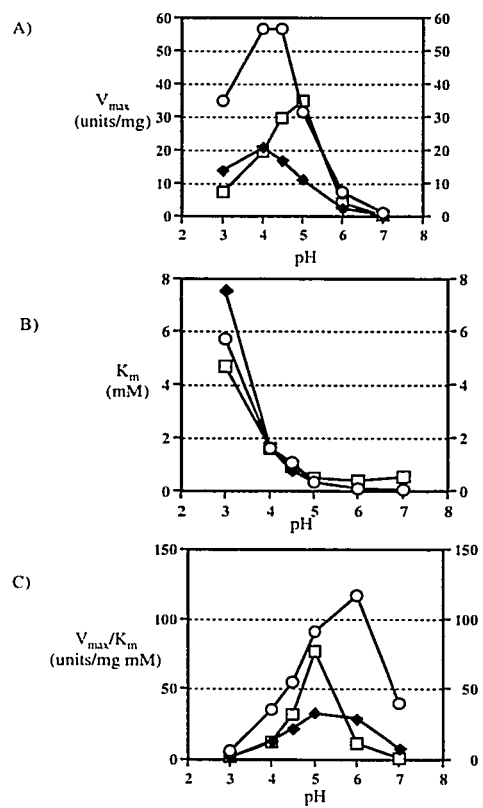
wild type enzyme). The V_{max}(trans) values, although not determined, seemed to be
independent of pH (from 3-7).

6.8.2.2 Reaction with natural substrates

The V_{max} values for the substituted enzymes were usually lower than for wild type.
However, they followed the same trend for the different substrates as was the case for wild
type. In general, similar trends of the K_m values as with the wild type enzyme were also
seen except that W551A β-glucosidase had a lower K_m when gentiobiose was the
substrate, while the K_m values for the cellooligosaccharides were all higher than those of
wild type β-glucosidase. However, no consistent overall trend was observed.

Table 6.18. Kinetic constants obtained with natural glucosyl substrates and the β-
glucosidases with substitutions for Trp 551.

Substrate	Wild type		W551A		W551F	
	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]	V_{max}[a]	K_m[b]
Cellobiose	78	2.9	12	3.2	34	2.7
Cellotriose	94	0.7	34	1.8	65	3.1
Cellotetraose	66	0.6	26	1.2	40	0.9
Cellopentaose	13	0.6	10	1.1	29	2.2
Gentiobiose	46	2.2	12	1.0	36	3.5

[a] All V_{max} values are reported in μmol/mg/min..
[b] The K_m values are in units of mM.

The downward curves that were normally observed on Eadie-Hofstee plots of the β-
glucosidases were only seen with W551F β-glucosidase (Figure 6.59). Eadie-Hofstee plots
of W551A β-glucosidase did not have any curvature at concentrations up to 20 mM
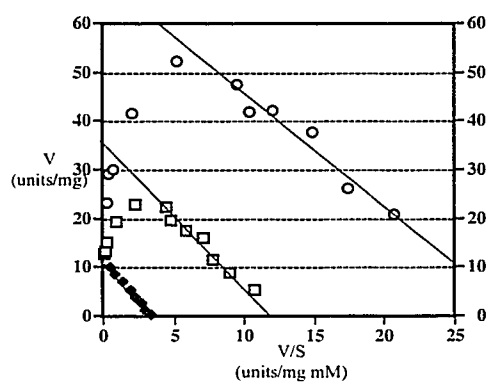
Figure 6.59. Comparison of Eadie-Hofstee plots resulting from the wild type

β-glucosidase and the W551 substituted enzymes reacting with

cellobiose. Reaction conditions are detailed in Section 3.6.7. (O)

wild type; (◆) W551A; and (□) W551F. All reactions were

done at 25°C, pH 4.5 with 8 µg/mL enzyme. One unit is

equivalent to the production of 1 µmol per min. of product.

cellobiose. However, some downward curvature did occur with longer

cellooligosaccharides.

6.8.2.3 Inhibitor studies

The K_i values obtained for the enzymes with substitutions for Trp 551 were

somewhat different from the values obtained for wild type (Table 6.19). W551A β-

Glucosidase was inhibited to a greater extent by all of the sugars except glucose and

allose. W551F β-Glucosidase was inhibited to a greater extent by all of the sugars except

mannose and maltose. Glucose, gluconolactone, cellobiose and gentiobiose were generally

good inhibitors, while maltose and xylose and the epimers of glucose were poor inhibitors.

Overall, however, there were not any consistent trends observed.

Table 6.19. K_i[a] values for β-glucosidase inhibitors with the W551 substituted enzymes.

Inhibitor	Wild Type	W551A	W551F
Glucose	2.7	4.8	1.4
Mannose	220	130	450
Allose	290	300	80
Galactose	280	135	85
Xylose	670	280	425
Gluconolactone	0.4	0.3	0.3
Cellobiose	3.4	1.5	0.8
Gentiobiose	4.8	1.1	2.0
Maltose	270	95	285

[a] K_i values are reported in mM.

6.8.3 Analysis of substrate breakdown by TLC

Analysis of cellobiose breakdown by TLC was only done at 50 mM cellobiose (Figure 6.60). Considering the lack of a curvature in the Eadie-Hofstee plot of W551A β-glucosidase, it was surprising that both of the substituted enzymes produced trisaccharide at more or less the same relative rate as did wild type enzyme. Gentiobiose was produced after the typical lags by both of the substituted enzymes.

6.8.4 Gas-liquid chromatography

The reactions of the β-glucosidases with substitutions for Trp 551 were analyzed by GLC in presence of 5 and 50 mM cellobiose. The enzyme concentrations used were chosen such that they produced glucose (Figure 6.61) at initial rates similar to the rates of wild type enzyme (2 μg/mL). The production of gentiobiose and gentiobiosyl glucose at 5 mM cellobiose was hardly above background. In the presence of 50 mM cellobiose (Figure 6.62) the amounts of glucose produced by the substituted β-glucosidases did not seem to plateau as was the case with the wild type enzyme. Gentiobiose was produced, as expected, after a lag during which glucose accumulated (Figure 6.62 (D)). Relative to the glucose produced, the rate of production of gentiobiosyl-glucose was a little slower for the substituted β-glucosidases than for wild type (Figure 6.62 (E) and (F)).

Figure 6.60 Thin layer chromatographic analysis of the reactions of the β-
             glucosidase with substitutions for Trp-551 with 50 mM
             cellobiose. A) W551A and B) W551F. Enzyme was added (8
             μg/mL final concentration) to the reaction mixture at zero time
             and an aliquot was immediately spotted onto the plate. Spotting
             was repeated at the times indicated in minutes on the bottom of
             the plate. The chromatography solvent used was n-butanol:ethyl
             acetate:2-propanol:acetic acid :water in the ratio of 1:3:2:1:1. The
             plates were developed twice (after drying in between). Products
             were detected by dipping the dried plates in 2% sulphuric acid in
             methanol and charring. The standards used on the plates were 5
             mM cellobiose (C), 10 mM gentiobiose (Ge) and 25 mM glucose
             (G).

A)

C 0 2 5 10 20 30 40 50 60 75 90 120 150 180 Ge G
(min.)

B)

C 0 2 5 10 15 20 25 30 40 50 60 75 90 120 150 180 Ge G
(min.)
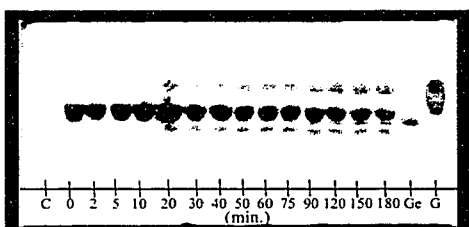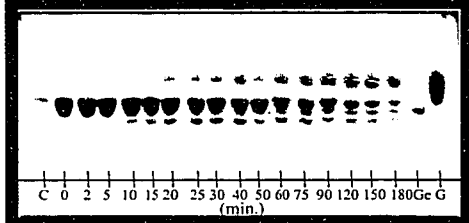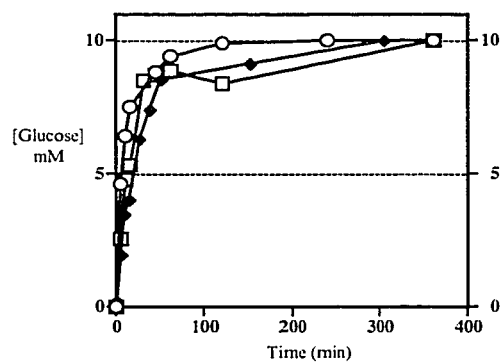
Figure 6.61. Analysis of products from the reaction of the wild type β-glucosidase and of enzymes with substitutions for Trp-551with 5 mM cellobiose over 6 hrs. (O) Wild type; (◆) W551A; and (□) W551F. The experiments were completed as described in Section 3.8.1.

353

Figure 6.62. Results of GLC analysis of the enzymes with substitutions for
Trp-551 with 50 mM cellobiose. The products that resulted from
the reaction were: glucose (A) and (B), gentiobiose (C) and (D),
and gentiobiosyl-glucose (E) and (F). The plots in parts (A), (C)
and (E) show the change in the concentration of product over the
entire reaction time (6 hr). The plots in parts (B), (D), and (F)
show the first 60 min. of product accumulation only. The
enzymes are represented by (o) wild type; (▲) W551A; and
(□) W551F. The substituted enzymes concentrations were
chosen so that they had an equivalent rate of glucose production
as 2 µg/mL wild type. All reactions were done at 25°C, pH 4.5.

A)

[Glucose]
mM

B)

[Glucose]
mM

C)

[Gentiobiose]
mM

D)

[Gentiobiose]
mM

E)

[Gentiobiosyl-
Glucose]
mM

F)

[Gentiobiosyl-
Glucose]
mM

Time (min)
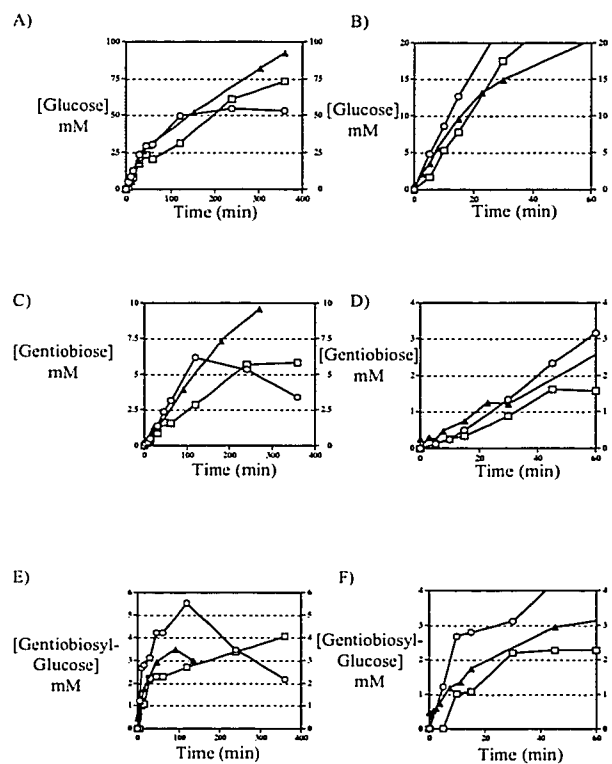
7. Discussion

7.1 Overview

Cellulases have enormous biotechnological potential. The breakdown of cellulose, a renewable resource, is important for many reasons. *A. niger* is known as an excellent producer of cellulases (Birk *et al.*, 1997) and β-glucosidase, the enzyme described herein, is one of the components of cellulases.

This project began with an investigation of a commercial cellulase mixture from *A. niger*. Several unique enzymes for cellulose breakdown were initially isolated from the mixture. One of these was identified as a β-glucosidase. The first experiments completed with the purified β-glucosidase were done with 3,4-dinitrophenyl-cellobiose (DNPC). The kinetics were quite unusual. The rate increased as a function of time and the enzyme seemed to be inhibited at high DNPC concentrations. A preliminary TLC analysis of the reaction products with cellobiose was also done. A disaccharide with a R_f similar to that of gentiobiose (β-D-glucopyranosyl-(1-6)-D-glucopyranose) was formed. This finding was interesting since β-galactosidase (an enzyme from *E. coli* that is also being studied in this laboratory) produces allolactose. Allolactose is similar to gentiobiose except that the non-reducing moiety is a galactose rather than a glucose. Allolactose is the natural inducer of the *lac* operon in *E. coli* and it was thought possible that the β-glucosidase from *A. niger* produced gentiobiose for similar reasons. Also, if gentiobiose were to be synthesized by the same mechanism as allolactose, it would be of interest to compare the two enzymes. There were also other reasons that the study of this enzyme was undertaken. A review of the literature indicated that the β-glucosidase from *A. niger* had

not been characterized extensively. In addition, each of the enzymes reported in the literature seemed to have quite different properties (as discussed in the Introduction - Section 1.2.3.2). Finally, there was no report in the literature of a gene for the *A. niger* β-glucosidase that had been cloned or sequenced.

A thorough physical and kinetic characterization of the β-glucosidase from the Sigma mixture was done. The intention was to clone the gene after the characterization was complete and express it so that substitutions of conserved amino acids could provide mechanistic insights.

Following the characterization of the β-glucosidase from the commercial mixture it was discovered that a group from Israel (Dr. O. Shoseyov, Hebrew University of Jerusalem) had already cloned and sequenced an *A. niger* β-glucosidase. The group had shown that the enzyme was a family 3 glycosidase. Comparison of the amino acid sequence of a CNBr peptide of the commercial β-glucosidase to the sequence encoded by the gene, SDS-PAGE, native PAGE, and preliminary kinetic studies showed that the two enzymes were identical. This led to a collaboration with the group from Israel. Analysis of family 3 glycosyl hydrolase sequences showed which residues were conserved and likely important in the β-glucosidase. Since Trp often plays a role in substrate binding in glycosyl hydrolases it was determined that enzymes with substitutions for Trp be examined. There are 21 Trp in the protein. Alignment of the available family 3 sequences showed which Trp were most conserved, and therefore which could be of importance. In addition, a Leu was present at the same position in this enzyme that is occupied by a Trp in some other family 3 glycosidases and it was decided to also study enzymes with

substitutions for that Leu. One structure of a family 3 glycosidase (barley β-D-glucan

exo-hydrolase) was also available and the structure was examined to help determine

which residues should be studied. The objective was to find substituted enzymes that had

large differences of their hydrolytic or transglucosidic activities compared to the wild

type enzyme. This information would allow the roles of the specific substituted amino

acids to be elucidated. Additionally, I wanted to acquire substitutions that could be

exploited for biotechnological applications. β-Glucosidases with substitutions for the

following amino acids were studied: Trp 21, 49, 139, 262, 430, 551 and Leu 426.

Substitutions for Trp 806 and 813 were also created but were only studied in a

preliminary way.

The sequences of the substituted β-glucosidase enzymes were checked in two

ways to ensure that the substitutions were correct and that no other mutations had

occurred to the gene sequences during mutagenesis. Firstly, following mutagenesis the

plasmids were isolated from *E. coli* and the β-glucosidase gene was sequenced in its

entirety. The gene was then transformed into *Pichia pastoris*, where it was integrated into

the genome. In order to verify that integration occurred, the *Pichia* genomic DNA was

isolated and primers specific to the *AOX* region (where the gene was inserted) of the

genome were used to amplify the integrated β-glucosidase gene. A PCR fragment of the

correct size was obtained which showed that the gene was inserted correctly into the

genome. The PCR fragment was ligated into a plasmid and sequencing was repeated to

ensure that the correct substitution was present.

7.2 Wild type β-glucosidase

Analyses of the commercial and the recombinant β-glucosidases by SDS and native PAGE, size exclusion chromatography and DLS were done. Native PAGE showed that the molecular weight is 223 kDa. Size exclusion chromatography showed that the molecular weight is 210 kDa. DLS indicated a molecular weight of 188 kDa. SDS-PAGE gave a subunit molecular weight of 125 kDa while the sequence indicated a subunit molecular weight of 91.4 kDa. Taken together, this data shows that the enzyme is most likely present as a dimer.

NMR spectra of the purified commercial β-glucosidase showed that oligosaccharides were present. Reaction of the purified enzyme with N-glycosidase F and subsequent SDS-PAGE showed that the molecular weight was decreased and that the carbohydrate part made up a significant portion of the molecular weight. Analysis of the amino acid sequence showed that 14 possible N-linked glycosylation motifs (Asn-X-Ser or Asn-X-Thr) are present. It is unlikely that all of these would be glycosylated. It has been shown that the motifs are probably not glycosylated if the X in the motifs is Leu, Glu, Asp, Trp or Pro (Shakim-Eshleman et al.,1996). Examination of the 14 possible motifs using this stricter criterion indicated that only 10 were likely to be glycosylated. Removal of the carbohydrate chains and profiling showed that there were 7 unique oligosaccharides. This suggests that only 7 of the 10 sites that had the favored glycosylation motifs were glycosylated or that some of the oligosaccharides are identical. Sequencing of the isolated oligosaccharides using the Glyko FACE[a] imaging system indicated that mannose was the most prevalent sugar in these chains and that the

oligosaccharides differed mostly in length; not in sugar composition. This was also observed for other glycosylated proteins from *A. niger* (Panchal and Wodzinski, 1998) and *A. fumigatus* (Rudick and Elbein, 1973). The recombinant enzyme was not analyzed for the oligosaccharide content but the molecular weights before reaction with N-glycosidase F, were the same for the two enzymes. In addition, the recombinant enzyme would probably not be active if it were not glycosylated.

The only significant difference in properties that was noted between the commercial and the recombinant β-glucosidases was that the recombinant β-glucosidase had larger V_{max} values. This is likely because the recombinant enzyme was always purified just prior to experimentation. The enzyme mixture from Sigma could have been months old. Additionally, in order for Sigma to provide a cellulase powder it is likely that the enzymes were lyophilized. This was shown to cause a decrease in β-glucosidase activity. In the light of the identity of the enzymes, the discussion below pertains to both forms of the enzyme.

The competitive inhibition studies with a number of sugars showed that the β-glucosidase is specific for glucosyl substrates. Mannose, allose, galactose and xylose were all very poor competitive inhibitors. This indicates that the orientation of the hydroxyls at the C2, C3 and C4 positions and the presence of a primary -CH_2OH at the C6 position are very important. In addition, each β-glucosyl substrate was a good inhibitor. Good inhibition by gluconolactone showed that it is a transition state analog inhibitor. The form of gluconolactone that resembles the transition state the most (D-glucono-(1-5)-lactone) is present in very small amounts in solution. D-Glucono-(1-4)-

lactone (which is not likely to be a transition state analog) predominates (Parke *et al.*, 1997).

The TLC reactions with cellotriose, cellotetraose and cellopentaose demonstrated that the enzyme preferentially removes glucose from the non-reducing end. Products were released and subsequently used as substrates to be shortened by another glucose (see Section 4.6.1, Figure 4.12, Figure 4.13). This shows that the enzyme only cleaves between the first and second sugars at the non-reducing end. Further evidence of the exo-hydrolase nature was that β-D-glucopyranosyl-(1-6)-β-D-glucopyranosyl-(1-4)-D-glucopyranose only resulted in glucose and cellobiose formation with no trace of gentiobiose (see Section 4.6.1). This can only happen if the enzyme is an exo-hydrolase. If it were an endo-hydrolase, gentiobiose would have been formed. The results obtained with DNPC (Figure 4.14 and Figure 4.23) provide further evidence that the enzyme is an exo-hydrolase. With this substrate, the evidence suggests that the two glucose components of DNPC bind to the first two glucose subsites (-1, +1) while the DNP portion is positioned at the third subsite (+2). This results in cleavage of the glucose-glucose bond but not the glucose-DNP bond. Color formation only occurs after the first bond is broken and when enough DNPGlc is formed to compete significantly with DNP-cellobiose for binding at the first two glucose subsites. Thus the rate of DNP production increases with time. The reason that the higher concentrations of DNPC result in slower rates of hydrolysis than lower concentrations of DNPC is that the DNPC at high concentrations binds to the -1 and +1 sites and blocks any DNPGlc that was formed from binding and being hydrolyzed.

A. niger β-glucosidase has at least 5 subsites (-1 to +4) for glucose. The K_m values

of cellooligosaccharides of increasing length decreased progressively and the V_{max} as well

as the V_{max}/K_m changed in each case (see Table 4.1 and 5.1). If there were not specific

subsites for the glucoses the values of these constants would stop changing when the

number of glucoses of the cellooligosaccharides was greater than the number of subsites.

It was found that the decreases in K_m were smaller between cellotriose and cellotetraose,

and between cellotetraose and cellopentaose than between cellobiose and cellotriose. This

implies that the enzyme binds very tightly at the first three glucose subsites (-1, +1, +2),

and that the remaining two (or more) sites (+3 and +4) bind glucose more weakly.

Therefore, the enzyme is an exo-hydrolase that is adept at hydrolysis of

cellooligosaccharides of different sizes. Yazaki *et al.* (1997) observed that an *A. niger* β-

glucosidase with properties that appear to be different from those of the enzyme

described here had 6 glucose subsites.

pNPGlc and oNPGlc had different K_m values. The K_m variances between pNPGlc

and oNPGlc are probably a result of the different nitrophenols binding with unique

affinities to the second glucose subsite (+1). This suggests that the second glucose subsite

(+1) has hydrophobic interactions. This is a common phenomenon for glycosidases and it

has been suggested that hydrophobic interactions also have a substantial role in the

binding at subsite +1 for a β-glucosidase studied by Ohnishi *et al.* (1998). The V_{max}/K_m

value for pNPGlc was significantly higher than oNPGlc. This suggests that the enzyme

reaction is more efficient with pNPGlc. One of the reasons for doing the substitution

studies was to determine if one of the conserved Trp studied was at this binding site.

Unfortunately, there was not any definite evidence that any of the residues that were substituted in this study were at the +1 site.

Unusual kinetic profiles were obtained for the reactions with this enzyme. There were pH dependent rate increases or decreases at high substrate concentration. The evidence for this has been discussed in the results section but the evidence and reasoning will be briefly reiterated here. TLC, NMR, and GLC evidence as well as isolation and analysis of the unknown sugars produced by the β-glucosidase indicated that the enzyme has two possible reaction pathways. At low concentrations of substrate, the reaction is mainly hydrolytic while at high concentrations the reaction is mainly transglucosidic. The initial common first step for each pathway is the glycolytic cleavage step with the rate constant, k_2 (see mechanism Figure 7.1). The enzyme cleaves the glycosidic bond in this step. The reducing glucose (or nitrophenol) that occupied the +1 site is rapidly released while the non-reducing portion of the substrate remains bound to the enzyme (it is probably attached covalently (Dan *et al.*, 2000)). The hydrolytic reaction pathway occurs when this glucosyl enzyme reacts with water (k_3) to form free glucose. The transglucosidic pathway occurs when a second substrate molecule binds at the position vacated by the aglycone (dissociation constant is K_i'') and the C6 OH group of the non-reducing glucose (of the substrate acting as an acceptor) reacts (k_4) to form the transglucosidic product (Figure 7.2). Most molecules with hydroxyl groups that were tested were able to act as acceptors. The action was not just restricted to the substrates. Also, with both cellobiose and gentiobiose, sugars larger than trisaccharides are also formed. This indicates that once a trisaccharide is formed in high enough amounts, it can

363
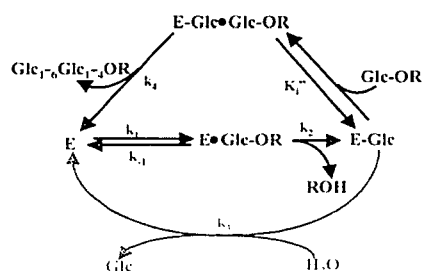
Figure 7.1. Proposed mechanism for the hydrolytic and transglucosidic
reactions of β-glucosidase. The enzyme is represented by E,
substrate is shown as Glc-OR, where the R group can be another
sugar or it can be a nitrophenyl group. $Glc_{1-6}Glc_{1-4}OR$ is the
transglucosidic product. Glc is glucose. Solid dots represent an
enzyme complex. The hyphen between the E and the Glc indicates
that this is probably a covalent intermediate. The common parts of
the mechanism are shown in green, the hydrolytic reaction is pink
and the transglucosidic reaction is blue. Production of gentiobiose
occurs as shown by the transglucosidic portion of the mechanism,
but in that case the acceptor is just glucose.

Figure 7.2. Representation of the β-glucosidase active site during transglucosylation. The non-reducing end glucose is held in the -1 glucose subsite(possibly covalently) while a disaccharide (in this case sophorose) diffuses into the +1 and +2 glucose subsites. The reducing end glucose is diffusing out of the binding site. Possible hydrogen bonds to the sugars in the subsites are represented by dotted lines. Other interactions are shown by dashed lines.

364

365

act as the acceptor of another glucose to produce a tetrasaccharide and so forth. It will be assumed that the subsites on the enzyme for binding the glucoses of the substrate are probably the same subsites to which the acceptors bind. If the acceptor is a monosaccharide it probably binds to the same subsite (+1) that the reducing glucose (or the nitrophenol) was bound to before it was cleaved and released from the enzyme. If the acceptor is a disaccharide or a nitrophenol substrate it probably binds to the +1 and +2 subsites.

The unusual rate decreases at high substrate concentrations at lower pH values and the rate increases at high substrate concentrations at high pH values occur because the hydrolytic reaction is pH dependent over the pH range studied while the transglucosidic reaction is pH independent. At low pH, k_4 is smaller than k_3 and transglucosylation occurs at a slower rate than hydrolysis. Since transglucosylation is slower than hydrolysis, the reaction appears inhibited at high substrate concentration. At high pH, k_3 is smaller while k_4 has the same value regardless of pH and this results in transglucosylation occurring at a faster rate than hydrolysis. It should be noted that at high pH, k_2 has to be larger than k_1 and k_4 has to be larger than k_3 or the results obtained would not be possible. Most of the other kinetic studies (Watanabe $et\ al.$, (1992), McCleary & Harrington (1988), Himmel $et\ al.$,(1993) Yan $et\ al.$, (1998), and Rashid & Siddiqui, (1999)) in the literature on β-glucosidases similar to this one have only reported rates at low substrate concentrations and at those concentrations the 'inhibition' and 'activation' would not have been noted.

The data show that k_2 is at least partially rate-limiting at pH 4.5.

For this enzyme: $k_{cat} = k_2k_3/(k_2 + k_3)$.

The portion of the reaction mechanism with k_3 as the rate constant is common for all substrates. Since the V_{max} values were different (see Tables 4.1 and 5.1) for the different substrates and because the step with k_2 is not common (since the leaving groups are either different in structure or if they are glucoses they are attached by different hydroxyl groups), k_2 is at least partially rate determining. To confirm this, stopped-flow studies would have to be done. At pH 7, k_2 is no longer rate determining. As already stated, the increases in rate at high substrate concentration (at high pH) can only happen if k_2 is relatively large compared to k_3 and if k_4 is larger than k_3.

The pH studies with oNPGlc, pNPGlc and cellobiose showed that the K_m values of each increased with decreasing pH. It is possible that one or more carboxyl groups in the binding site become protonated at the lower pH values and reduce the enzyme's ability to bind substrate at lower pH. However, the situation is not that clear. For glycosidases:

$$K_m = K_s k_3/(k_2 + k_3)$$

So, changes of any of K_s, k_3 or k_2 with pH could affect the pH profiles for K_m. To determine which caused the large increases in K_m at lower pH, one would have to do stopped flow studies. The enzyme had maximum V_{max} values at pH 4 - 4.5. The V_{max}/K_m values are equivalent to k_2/K_s. Similar profiles were seen for the different substrates. Again, however, no attempt was made to determine how changes in k_2, k_3 or K_s affected the profiles obtained.

Since both the hydrolytic and transglucosidic reactions probably occur by similar reaction mechanisms, it is difficult to hypothesize what would render transglucosidation pH independent. It is possible that a conformational change is involved. Possibly there is a group at the active site that is ionizable and is important to keep the active site in the correct conformation for activity. As the pH is raised and the group is deprotonated, it could lose the correct charge and the subsequent conformation change would cause the activity to decrease. Since sugar acceptors are much larger than water, it is conceivable that the acceptors could use their binding interactions to hold the enzyme in the correct conformation for reaction regardless of whether the group is protonated or not. Effects of pH have also been found to be different for the hydrolytic and transglycolytic reactions of β-galactosidase from *E. coli* (Huber *et al.*, 1983).

As mentioned above, TLC analyses of the reaction with cellobiose showed that a disaccharide was formed that had the same R_f as gentiobiose. Purification of the unknown and analysis by NMR showed that the disaccharide was indeed gentiobiose. Since β-galactosidase forms allolactose (Huber at al., 1976) in a direct transglycosidic addition (without prior release of glucose) it was expected that the gentiobiose would also be formed through direct transglycosylis with the β(1-4) bond breaking and a β(1-6) bond forming. However, all of the data obtained for each enzyme studied showed that there was always a delay before gentiobiose formed and that gentiobiose was only synthesized when the glucose concentration had built up to high concentrations. The gentiobiose is, therefore, produced in the same manner as the larger oligosaccharides, via indirect transglycosylation with glucose as the acceptor. The difference is probably that the +1

site in this β-glucosidase has much poorer affinity than the comparable site of β-galactosidase.

The V_{max}(trans) when the substrates themselves are acceptors depended on the make-up of the substrate (Table 4.1). In particular, the V_{max}(trans) was higher for cellobiose than for pNPGlc or oNPGlc. When pNPGlc and oNPGlc are the acceptors, the nitrophenol group should bind at the +2 position. This could change the positioning of the C6 hydroxyl of the glucose at the +1 position that reacts to form the transglucosidic product. The data obtained with the cellooligosaccharide substrates showed that the +2 subsite has a high affinity for glucose. It may also be very important for the proper positioning of the acceptor. Thus, there would be different V_{max}(trans) for each substrate.

There must be some purpose for the transglucosidic activity of the β-glucosidase or it would be a very inefficient way of hydrolyzing cellobiose. The enzyme may have transglucosidic activity in order for it to produce an inducer molecule. Gentiobiose may be an inducer similar to allolactose for the lac operon in *E. coli* (Huber *et al.*, 1976). It has been suggested that sophorose is an inducer for the *Trichoderma reesei* cellulases. It is produced by that enzyme's transglycosidic action (Carle-Urioste *et al.*, 1997). The trisaccharide produced from cellobiose could also be an inducer. It is, however, also possible that the enzyme may only produce these larger oligosaccharides as a way of storing glucose for future use.

None of the metals studied had any effect on the enzyme activity. EDTA also had no effect on the enzyme reaction. Metals also had no effect on a 200 kDa β-glucosidase from *A. niger* characterized by Galas and Romanowska (1997) that has different

properties, suggesting that it may be common for metals not to be important in *A. niger* β-glucosidases.

7.3 Tryptophan 21

Fourteen of 17 sequences in subfamily 4 (glycosyl hydrolase family 3) have a Trp at a position equivalent to position 21 of the β-glucosidase from *A. niger*. Experimentally, however, when this residue was substituted, the results obtained did not differ greatly from those of the wild type enzyme.

SDS-PAGE and size exclusion chromatography indicated that the enzymes with substitutions for Trp 21 were the same size as the wild type enzyme and were dimers. The fluorescence spectra were not different from wild type for any of the substitutions, which indicated there were no major structural changes. It is possible that minor conformational changes did occur but were masked in the spectra because of the large number of Trp in the protein. The inhibition studies indicated that the substituted enzymes were still specific for glucosyl substrates. There were some differences in the V_{max} and K_m values. The reasonably small K_m changes suggest that Trp 21 does not have a direct impact on substrate binding. The differences in V_{max} were larger and may result from small conformational changes that effect k_2. Small conformation effects would have to be transferred to the active site and upon substitution affect the positioning of the active site catalytic residues to cause the V_{max} changes. It is unfortunate that an equivalent Trp was not found in the barley enzyme that could provide some clue as to the position of Trp-21 in the *Aspergillus niger* β-glucosidase. The transglucosidic reactions did not seem

to be affected by these substitutions. The same results at high and low pH and high and low substrate concentrations were observed as occurred with wild type enzyme.

7.4 Tryptophan 49

There is less conservation of Trp 49 than of Trp 21 in branch #4 of the family 3 glycosyl hydrolases. Only 10 of the 17 sequences contained a Trp at this position. In the dissimilar sequences a variety of amino acids were observed (Lys, Phe, Ile, Gly, Ser, Ala). In the barley β-glucan exohydrolase structure Trp-70 is equivalent to Trp-49 (Figure 7.3). It appears to be buried some distance from the active site and if it is present at the same position in the *A. niger* β-glucosidase the affect on the conformation at the active site would have to be transmitted through the structure.

The fluorescence spectra obtained with the Trp-49 substituted enzymes indicated that there were no major structural differences between the substituted β-glucosidases and the wild type enzyme. Gel exclusion chromatography and SDS-PAGE indicated that the enzymes are the same size as the wild type and are dimers. The inhibition studies showed that the substitutions did not affect the β-glucosidase preference for glucosyl substrates. The pH profiles that were obtained were similar to wild type except for the magnitude of the V_{max} values. Assays with pNPGlc, oNPGlc and cellobiose showed that the V_{max} and K_m values were decreased. The major significant kinetic effect was that the larger (longer) cellooligosaccharides had distinctly higher K_m values than the shorter cellooligosaccharides (see Table 6.5). This is opposite to what occurs with wild type β-glucosidase. With wild type the K_m values decreased with increased substrate length (probably due to cooperative affects). The data showing that the K_m values for cellotriose,
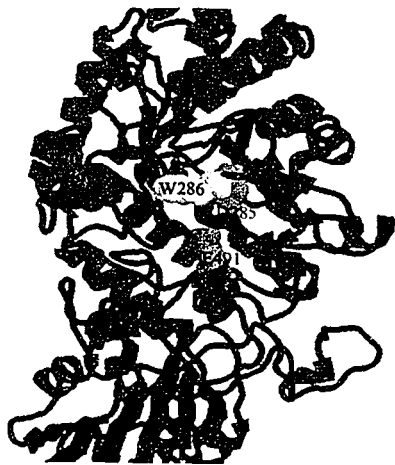
Figure 7.3 Barley β-D-glucan exohydrolase structure illustrating the positioning
of Trp-49 (W70-barley)(red) relative to Trp-262 (W286-barley)
(yellow). The catalytic nucleophile (D285-*H.vulgare*, D261-*A. niger*)
and putative catalytic acid/base (E491- *H.vulgare*, no equivalent in *A.
niger*) are shown in cyan. The structure was obtained from the RCSB
protein data bank (http://www.rcsb.org/pdb/) using the code 1EX1.
The structure was reported by Varghese *et al.*, (1999).

cellotetraose and cellopentaose are relatively large compared to that for cellobiose suggests that substitutions for Trp 49 affect the +2 subsite and maybe some higher glucose subsites. They may decrease the cooperative effects normally involved in binding longer substrates. Since the affinity of the enzymes having substitutions for Trp-49 decreased with substrate length, it is probable that the affinity for acceptors binding to these sites is also decreased. This would explain the need for (see Section 6.3.4) higher substrate concentrations for the transglucosidic reactions with these enzymes. The maximum rate of the transglucosidic reactions also appears to be small for these enzymes (see Figure 6.20). As stated earlier (Section 7.2), the binding at the +2 site can affect how the acceptor is oriented for the transglucosidic reaction. Wrong positioning can cause the V_{max}(trans) values to be low. Over and above this, the low binding affinity for acceptors would also allow hydrolysis to occur more readily since the enzyme would not be forced into the transglucosidic pathway. It is likely that the acceptor site in the wild type enzyme has even higher affinity for acceptors in these subsites after glycosidic cleavage (k_2) has occurred (Huber *et al.*, 1984) and that the effect of loss of acceptor binding is even more significant than indicated by the increased Km values of the longer cellooligosaccharides.

Trp-49 was substituted with 5 different amino acids and each gave similar results. (see Figure 6.20). There were some differences but they were not large. This strongly suggests that Trp 49 does not interact directly with the substrate. If Trp 49 were directly involved the different substitutions would have very different effects. Since Trp is a large residue, substitution should result in at least a small cavity and if there is a collapse of the cavity that causes conformation changes these could be transmitted to the active site. The

Trp at the matching position in barley β-D-glucan-exohydrolase sits buried under the active site with 1 or 2 peptide backbone chains between. The affects of the substitutions could be transmitted through the backbone chains.

The enzymes with substitutions for Trp 49 have biotechnological implications. For cellulose degradation to be complete, a β-glucosidase that does not form oligosaccharides would undoubtedly be more efficient. The list of biotechnological applications for efficient cellulose breakdown to glucose is large. Some examples are increasing the nutritive value of animal feeds by converting cellulose to glucose, production of glues and for the manufacture of raw material (glucose) for the fermentation industry (Mandels, 1985). It would be more efficient to utilize a β-glucosidase that does not have acceptor activity for the formation of the glucose.

7.5 Tryptophan 139

The Trp 139 is quite conserved within the subfamily 4 sequences of the family 3 glycosyl hydrolases. Of the 17 sequences, 14 have a Trp that matches the 139 position. Two of these others have a Phe, and one has a Gly. W139L β-Glucosidase had less affinity for pNPGlc then did the wild type enzyme. Changes in V_{max} and K_m indicate that Trp-139 may have some type of interaction with the substrates at the +1 subsite. However, since the results are only preliminary, the role of Trp-139 cannot be predicted with any accuracy. The equivalent residue (Tyr-160) in barley β-D-glucan exo-hydrolase was close to the active site and thus, there would be some interactions with the glucose subsites.

7.6 Tryptophan 262

Trp 262 is strictly conserved in all of the sequences of family 3 glycosyl hydrolases. The results reported here verify that the residue is important and show that substitution converts the enzyme into one having mainly transglucosidic activity

The fluorescence spectra were very similar to the spectra of wild type with the exception of those obtained with W262C β-glucosidase. The emission and excitation spectra of that enzyme were red-shifted. It is probable that the enzyme structure is different. There are 21 Trp and unless a structural change is quite large, there would not be a noticeable shift. Despite this difference the overall results with W262C β-glucosidase were quite similar to those of the other β-glucosidases with substitutions for Trp 262. SDS-PAGE and size exclusion chromatography indicated that all of the enzymes with substitutions for Trp-262 were the same size as the wild type enzyme and were dimers.

Very large K_i values were obtained for the inhibition of the enzymes that had substitutions for Trp-262 (Table 6.10). Thus substrate affinity is very poor. Analysis of the values on the table show that the greatest increase of K_i values over wild type occurred with glucose and gluconolactone. Increases of the K_i values of up to 200 fold were noted. The K_i values of the epimers were not increased very much except for mannose that had increases in the K_i values of up to 30 fold. The K_i values of xylose and maltose were also not increased very much. The inhibition by disaccharides having β-glucosyl bonds (cellobiose and gentiobiose) was affected less by the substitutions (20 to 80 fold) than was inhibition by glucose. Inhibition by cellotriose was affected even less

(8-17 fold). These data indicate that substitutions for Trp-262 cause decreases of binding at the -1 site. In other words the substitutions cause the non-reducing glucose to bind poorly. This places Trp-262 at the non-reducing glucose subsite (-1). This is the same position that the equivalent Trp is present in the barley β-D-glucan exo-hydrolase (Figure 7.4).

The K_m data (Tables 6.8 and 6.9) confirm this positioning. Large K_m values were obtained for each nitrophenyl substrate, for cellobiose, and for gentiobiose relative to the K_m values of the wild type enzyme. The relative affects on the K_m were much smaller for cellotriose and larger cellooligosaccharides. The K_m values with these substrates were still larger than they were for wild type but by a much smaller factor.

The major result of the poor binding (and probably also poor interaction with nucleophile) with these substituted enzymes was that transglucosidation was the major reaction pathway. Only W262F β-glucosidase had any significant hydrolytic activity. In general the V_{max} values for these enzymes were quite small.

Despite the very high K_m values and low V_{max} values, the pH trends in most cases were similar to wild type, except that the V_{max} values were much smaller. Since these enzymes have mainly transglucosidic activity it could be expected that the reactions would be pH independent just as the transglucosidic activity of the wild type enzyme is pH independent. However since the transglucosidation rate (k_4) of these enzymes is fast, the rate of the glycosidic cleavage step (k_2) is slow in comparison. Therefore, k_2 is rate-determining and the V_{max} pH profiles are similar to those of wild type enzyme because in that case k_2 was also rate-determining. The k_2 values of these substituted enzymes are
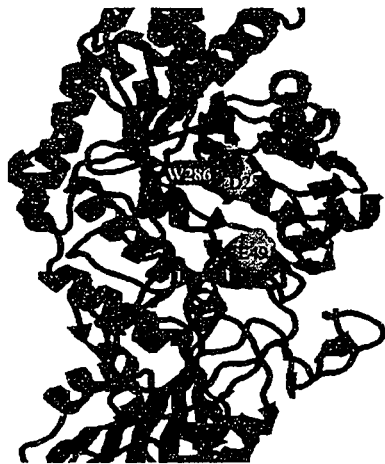
Figure 7.4 Barley β-D-glucan exohydrolase structure illustrating the positioning
    Trp-262 (W286-barley) (red). The catalytic nucleophile (D285-
    H.vulgare, D261-*A. niger*) and putative catalytic acid/base (E491-
    H.vulgare, no equivalent in *A. niger*) are shown in cyan. The
    structure was obtained from the RCSB protein data bank
    (http://www.rcsb.org/pdb/) using the code 1EX1. The structure was
    reported by Varghese *et al.*, (1999).

probably small and rate determining because of the effects of the substitutions on the

adjacent catalytic nucleophile, Asp-261.

The pH profiles of the effects on the K_m values of W262C and W262L β-

glucosidases were unusual (Figure 7.5). They did not change much as a function of pH.

The substitution must affect the structure so that the affinity for the substrate is not

affected by pH changes. The substitutions could move an ionizable group that normally

affects binding out of position so that the binding becomes pH independent.

Since Phe has an aromatic side-chain, one might think that substitution by a Phe

would result in an enzyme with similar properties as wild type. However, when Trp-262

was substituted by a Phe the same binding effects as those observed with the other

substitutions were found. On the other hand, the effects on the V_{max} on the transglucosidic

and reactions did indicate that the enzyme substituted with a Phe had some characteristics

of the wild type enzyme. The hydrogen bonds with Trp-262 may be more important for

binding the non-reducing glucose than are the stacking interactions. It is also possible that

Phe is not oriented into the same position as Trp is in the wild type enzyme and thus does

not allow for good stacking.

The enzyme with the Phe substitution had the highest V_{max} values of the

substituted enzymes and the proportion of the transglucosidic reaction was the largest

with that substitution. It is possible that Trp is only required to maintain a hydrophobic

environment in the active site (Kurik *et al.*, 1996). However, it would be expected that

Leu could also fulfill this role and the data do not support that it does. It is interesting that

transglucosidation was not affected as much by the substitution with Phe as was binding.

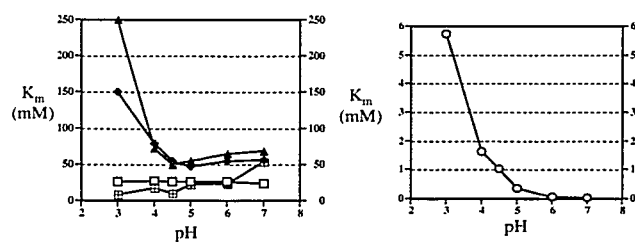Figure 7.5. pH profiles (K_m) of the W262 substituted enzymes with pNPGlc
(25°C). The data for the substituted enzymes is shown in part (A)
and the corresponding wild type plots is shown in (B). The
enzymes are represented by (◆) W262A, (□) W262C, (▲)
W262F, (⊞) W262L and (○) for wild type. All kinetic values
were obtained from Eadie-Hofstee plots. All of the reactions were
done using 1.4 µg/mL enzyme at 25°C.

Most likely, the k_3 is also not affected as much as is k_2 but stopped flow evidence would be needed to confirm this.

The evidence obtained suggests that the following occur with the enzymes having substitutions for Trp-262. Upon substitution for Trp-262, the non-reducing glucose of the substrate no longer binds well and in addition Asp-261 (the catalytic nucleophile adjacent to Trp 262) is not orientated normally. Thus, very high concentrations of substrate (because of the large K_m values) are needed to bind the substrates for the glycosidic bond cleavage step. The rate is also very slow (probably because Asp-261 is mis-orientated). However, cleavage does occur slowly (a small k_2). Since the subsites to which the substrates bind as acceptors (subsites +1 and +2) seem to have been unaffected by substitutions for Trp-262 they become essentially saturated with acceptor at the high concentrations of substrate needed for binding as substrates. Thus, only the transglucosidic reaction takes place. It is possible that the k_3 value is also decreased by the substitutions and thus, that the hydrolytic reaction is further suppressed. The evidence for this is that the enzyme substituted by a Phe still has some hydrolytic activity. When Phe is substituted it still may be able to line the residues up more favorably for the hydrolysis reaction (k_3) since Phe is also aromatic.

The predominant transglycosidic activity of the enzymes with substitutions for Trp-262 has valuable biotechnological potential. If an acceptor is added, one should be able to readily synthesize a glucopyranoside with the acceptor added in the β configuration. Karthaus et al. (1994) produced unique molecules by varying the acceptor added to an unsubstituted glycosidase from another organism. The transglycosidic ability

of some other enzymes has also been utilized for synthesis of larger sugars (Hansson *et al.*, 2001, Shoda *et al.*, 1993). Sugars larger than disaccharides are quite expensive to synthesize because they are difficult to make and purify. Chemical synthesis of oligosaccharides takes many steps since it is necessary to add protective groups. When transglycosidic enzymes are used, the specificity of the particular enzyme can be used to select for the desired reactions and products (Shoda *et al.*, 1993) without protection. The use of enzymes for oligosaccharide synthesis allows much milder conditions to be utilized (Hansson *et al.*, 2001). For most of the enzymes where the transglycosidic activity has been investigated for use in oligosaccharide synthesis, the yield of products was fairly low (20-40%) (Nikolova *et al.*, 1996, Onishi and Tanaka, 1995,). Recently, Hansson *et al.* (2001) were able to engineer a β-glucosidase from *Pyrococcus furiosus* that produced 45% transglycolytic products. The yield from three of the β-glucosidases with substitutions for Trp-262 was 100%. No hydrolysis occurred. Thus, the potential is obvious.

7.7 Leucine 426

Leu 426 is highly conserved. Of 17 aligned sequences in branch 4 of the family 3 glycosidases only one did not have a Leu and it had a Ile. Due to this high conservation it was expected that substitution of this residue would provide some interesting results. Furthermore, the equivalent of this residue in the barley β-D-glucan exo-hydrolase is a Trp that is thought to interact with a glucose (Varghese *et al.*, 1999). It was, therefore, hoped that studies with this substituted residue would reflect these interactions.

The fluorescence studies showed that the emission and excitation spectra were all very similar to the spectra of wild type. Thus there were no major structural differences as a result of the substitutions for the Leu. SDS-PAGE and size exclusion chromatography showed that the substituted ß-glucosidases were the same size as the wild type enzyme and were dimers.

Surprisingly, the kinetic results did not show large differences from the wild type enzyme. The inhibitor studies showed that the β-glucosidases with substitutions for Trp-426 were still specific for glucosyl residues and the pH profiles did not show any major differences from the wild type enzyme. Overall the V_{max} and K_m changes were not too different. Reactions with the cellooligosaccharides also showed the same general trends as wild type. The only real difference noted was that the glucose production plateaued at higher concentrations than for wild type (Figure 6.44). Studies to see if this is significant would have to be checked out at even higher cellobiose concentrations than were used in the experiment of Figure 6.44.

The results suggest that this residue is not directly involved in binding the glucose subsites in the β-glucosidase from A. niger. It is possible that the effect of these substitutions may be masked by the Trp-430 residue. These amino acids may have to function together. They are quite close together in barley β-D-glucan exo-hydrolase (Figure 7.6) and if they were only required to maintain a hydrophobic environment, the substitutions chosen would not greatly affect that role.
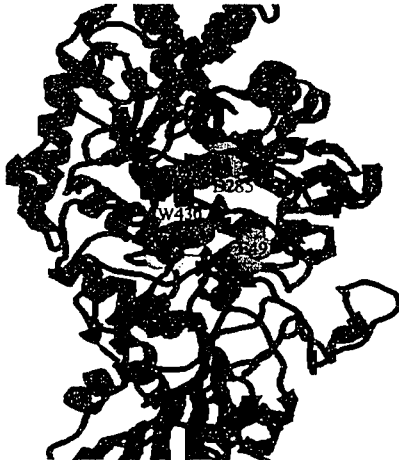
Figure 7.6 Barley β-D-glucan exohydrolase structure illustrating the positioning
Leu-426 (W430-barley) (yellow), Trp-430 (W434-barley) (magenta)
and Trp-262 (W286-barley) (red). The catalytic nucleophile (D285-
H.vulgare, D261-*A. niger*) and putative catalytic acid/base (E491-
H.vulgare, no equivalent in *A. niger*) are shown in cyan. The
structure was obtained from the RCSB protein data bank
(http://www.rcsb.org/pdb/) using the code 1EX1. The structure was
reported by Varghese *et al.*, (1999).

7.8 Tryptophan 430

Trp-430 was also expected to show interesting results. This site is less conserved than Leu 426, but 12 of the 17 sequences have a Trp at this position. One of the others has a Tyr and the remaining 4 have a Gly. Again, comparisons to the structure of barley β-D-glucan exo-hydrolase suggest that this is an important residue (Figure 7.6).

The fluorescence spectra of the β-glucosidases with substitutions for Trp 430 indicated that none of the substituted enzymes had major structural differences from the wild type β-glucosidase. Size exclusion chromatography and SDS-PAGE indicated that the enzymes were the same size as wild type and were dimers.

The inhibition studies indicated that there were no changes of the glucosyl preference of the wild type enzyme and the pH profiles were generally similar. The largest effects were on the V_{max} values and only small changes of the K_m values were observed.

W430L β-Glucosidase was somewhat different from the other substituted β-glucosidases. Its transglucosidic reaction was relatively slow. It is hard to explain why this substitution affected the transglucosidic rate but that the others did not.

It is possible that the role of this Trp is to maintain a hydrophobic environment in conjunction with Leu-426. The substitutions chosen would not greatly change the environment since no polar or charged residues were investigated. It is also possible that Trp-430 is not near the active site of the enzyme.

7.9 Tryptophan 551

The branch #4 sequence alignment showed that 10 of the 17 sequences have a Trp at a site equivalent to that of Trp-551. The other residues were 2 Leu, 3 Ala, 1 Phe and 1 Tyr.

The fluorescence spectra obtained for the W551 substituted enzymes showed that the substituted enzymes did not have any major structural difference from the wild type enzyme. SDS-PAGE and size exclusion chromatography indicated that the enzymes had the same molecular weight as the wild type enzyme and that they were dimers.

The enzymes were kinetically similar to wild type except that the pH optimum for the V_{max} of W551F β-glucosidase was slightly higher than for wild type and that W551A β-glucosidase reacted abnormally with cellobiose. If Trp-551 is located in a similar position as in the barley enzyme, it is not close to the binding or catalytic sites. The small effects of the substitutions probably result from small conformation changes that are transmitted from some distance away. There again was an effect where the glucose production plateaued at higher levels than with wild type.

7.10 Future work

Stopped flow kinetics studies would be useful to show more precisely how the enzymes are acting. This technique was not available at this institution when this work was done.

It would be very helpful to have a structure of this enzyme in order to understand the effects of the substitutions better. The β-D-glucan exo-hydrolase structure was somewhat helpful but the two enzymes only have 37% identical sequence and the A.

niger β-glucosidase is almost 200 amino acids larger. Some of the residues that seem to be located at the active site of that enzyme do not seem to play a role in the β-glucosidase studied here.

Studies with substitutions at other positions could be carried out. The catalytic nucleophile has been identified (Dan *et al.*, 2000) but the catalytic acid/base base is still unknown. Additionally there is the possibility that a third carboxylic acid could be involved in the transglucosylation reaction (Kim *et al.*, 2000). Additionally, there are several residues that are positioned in the barley β-D-glucan exo-hydrolase active site (Varghese *et al.*, 1999) and that are conserved in the *A. niger* enzyme that could be studied. The hydrophobic patch may be involved in glucose binding. Met-250 and Met-316 (*H. vulgare* numbering) of this patch are strictly conserved. The basic patch might also be of interest as these residues could also be involved in hydrogen bonding to the glucose in the -1 subsite. In addition Tyr-253 and Asp-95 are strictly conserved in branch #4 of the family 3 glycosidases. Tyr-253 may interact with the glucose in the -1 site while Asp-95 may interact at the +1 binding site. Tyr-229 (*A. niger*) which is the equivalent of Tyr-253 from the barley enzyme, may have a role similar to that of Tyr-503 in β-galactosidase (*E. coli*), which is highly conserved and found in the catalytic site (Jacobson *et al.*, 1994).

The two substitutions that were the most interesting could be investigated further. The β-glucosidases with substitutions for Trp-262 could be reacted with a variety of acceptors to try to form unique products. Further studies with the enzymes with substitutions for Trp-49 would give more information on how they function. Also, it

might be beneficial to look at the affects of substituting both Leu-426 and Trp-430 at the same time. These residues may mask the actual roles when substituted separately. The effects of some of the substitutions on the levels at which glucose plateaus might be worth following up.

One reason that *Pichia pastoris* was chosen for expression of the recombinant enzymes is that it is known to glycosylate proteins that are secreted (Fu *et al.*, 2001). Since the glycosylation pattern of the Sigma β-glucosidase has been determined, the recombinant enzyme's glycosylation patterns should also be explored.

References

Abuja, P.M.(1988) Structural and functional domains of cellobiohydrolase I from *Trichoderma reesei*. *European Biophysical Journal* 15: 339-342.

Alzari, P.M., Souchon, H. and Dominguez R. (1996) Crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum*. *Structure* 4: 265-275.

Bains, W. (1998) Biotechnology from A to Z. 2nd ed. Oxford University Press, New York. pp 190-191.

Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J., and Hughes, M.A. (1995) The crystal structure of a cyanogenic β-glucosidase from white clover, a family1 glycosyl hydrolase. *Structure* 3, 951-960.

Bause E., and Legler, G. (1980) Isolation and structure of a tryptic glycoprotein from the active site of β-glucosidase A₃ from *Aspergillus wentii*.

Bayer, E.A., Morag, E. and Lamed, R. (1994) The cellulosome-a treasure-trove for biotechnology. *Trends in Biotechnology* 12, 379-386.

Beguin, P. (1990) Molecular biology of cellulose degradation. Annual *Reviews in Microbiology* 44, 219-248.

Beguin, P., Raynaud, O., Chaveroche, M.K., Pridi, A., and Alzari, P.M. (1996) Subcloning of a DNA fragment encoding a single cohesion domain of the *Clostridium thermocellum* cellulosome-integrating protein CipA: purification, crystallization and preliminary diffraction analysis of the encoded polypeptide. *Protein Science* 5, 1192-1194.

Beguin, P., and Alzari, P.M. (1998) The cellulosome of *Clostridium thermocellum*. *Biochemical Society Transactions* 26, 178-185.

Birsan, C., Johnson, P., Joshi, M., Macleod, A., McIntosh, L., Monem, V., Nitz, M., Rose, D.R., Tull, D., Wakarchuck, W.W., Wang, Q., Warren, R.A.J., White, A., and Withers, S.G. (1998) Mechanisms of cellulases and xylanases. *Biochemical Society Transactions* 26, 156-160.

Birk, R., Ikan, A., Bravado, B., Braun, S., and Shoseyov, O. (1997) Synthesis of iso-propyl-1-thio-β-D-glucopyranoside (IPTGlc), an inducer of *Aspergillus niger* B1 β-glucosidase production. *Applied Biochemistry and Biotechnology* 66, 25-29.

Christakopoulos, P. (1994) Purification and characterization of an extracellular β-glucosidase with transglycosylation and exo-glucosidase activities from *Fusarium oxysporum*. *European Journal of Biochemistry* 224, 379-385.

Coughlan, M.P. (1985) Cellulases: production, properties and applications. *Biochemical Society Transactions* 13, 405-416.

Coutinho, J.B., Gil;kes, N.R., Warren, R.A., Kilburn, D.G., and Miller, R.G. Jr. (1992) The binding of *Cellulomonas fimi* endoglucanase C (CenC) to cellulose and sephadex is mediated by the N-terminal repeats. *Molecular Microbiology* 6, 1243-52.

Damude, H.G., Ferro, V., Withers, S.G., and Warren, R.A.J. (1996) Substrate specificity of endoglucanase A from *Cellulomonas fimi*: fundamental differences between endoglucanases and exoglucanases from family 6. *Biochemical Journal* 315, 467-472.

Dan, S., Marton, I., Dekel, M., Bravdo, B-A., He, S., Withers, S.G., and Shoseyov, O. (2000) Cloning, expression, characterization, and nucleophile identification of family 3, *Aspergillus niger* β-glucosidase. *Journal of Biological Chemistry* 275, 4973-4980.

Davies, G.J., Wilson, K.S., and Henrissat, B. (1997) Nomenclature for sugar-binding sites subsites in glycosyl hydrolases. *Biochemical Journal* 321, 557-559.

389

Deacon, J.W. (1984) Introduction to Modern Mycology. Ed. J.F.
    Wilkinson. Blackwell Scientific Publications, Oxford.

Dekker, R.F.H. (1986) Kinetic Inhibition and stability properties of a
    commercial β-D-glucosidase (cellobiase) preparation from
    Aspergillus niger and its suitability in the hydrolysis of lignocellulose.
    Biotechnology and Bioengineering **28**, 1438-1442.

Deschavanne, P.J., Viratelle, O.M., and Yon, J.M. (1978) Conformational
    adaptability of the active site of β-galactosidase. *The Journal of
    Biological Chemistry* **253**, 833-837.

Divne, C., Stahlberg, J., Reinkainin, T., Ruohonen, L., Pettersson, G.,
    Knowles, J.K., Teeri, T.T. and Jones, T.A. (1994) The three-
    dimensional crystal structure of the catalytic core of
    cellobiohydrolase I from *Trichoderma reesei*. *Science* **265**, 524-528.

Divne, C., Stahlberg, J., Teeri, T.T., and Jones, T.A. (1998) High-
    resolution crystal structures reveal how a cellulose chain is bound in
    the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei*.
    Journal of Molecular Biology **275**, 309-325.

Dominguez, R., Souchon, H., Spinelli, S., Dauter, Z., Wilson, K.S.,
    Chauvaux, S., Beguin, P and Alzari, P.M. (1995) A common protein
    fold and similar active site in two distinct families of β-glycanases.
    Nature: Structural Biology **2**, 569-576.

Drickamer, K (1997) Making a fitting choice: common aspects of sugar
    binding sites in plant and animal lectins. *Structure* **5**, 465-468.

Duff, S.J.B.(1985) Cellulose and β -glucosidase production by mixed
    culture of *Trichoderma reesei* RUTC30 and *Aspergillus phoenicis*.
    Biotechnology Letters **7**, 185-190.

Fu, J. Prade, R. and Mort, A. (2001) Expression and action pattern of
    Botryotinia fuckeliana (Botrytis cinerea) rhamnogalacturonan
    hydrolase in *Pichia pastoris*. *Carbohydrate Research* **330**, 73-81.

Galas, E., and Romanowska, I. (1997) Purification and some properties of β-glucosidase from *Aspergillus niger* IBT-90. *Acta Microbiologica Polonica* **46**, 241-252.

Gilkes, N.R. Claeyssens, M., Aebersold, R., Henrissat, B., Meinke, A., Morrison, H.D., Kilburn, D.G., Warren, R.A. and Miller RC Jr. (1991) Structural and functional relationships in two families of β-1,4-glycanases. *European Journal of Biological Chemistry* **202**: 367-377.

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller R.C. Jr., and Warren R.A. (1991) Domains in microbial β-1,4-glycanases: sequence conservation, function and enzyme families. *Microbiological Reviews* **55**, 303-315.

Glyko Inc. FACE^x N-linked sequencing manual, (1996). Novato, CA.

Glyko Inc. FACE^x N-linked profiling manual, (1994). Novato, CA.

Grace, M.E., Desnick, R.J. and Pastores, G.M. (1997) Identification and expression of acid β-glucosidase mutations causing severe type I and neurological type2 Gaucher disease in non-Jewish patients. *Journal of Clinical Investigations* **99**, 2530-2537.

Grous, W. (1985) Kinetics of cellobiose hydrolysis using cellobiase composites from *Trichoderma reesei* and *Aspergillus niger*. *Biotechnology and Bioengineering* **27**, 463-470.

Gupte, A. and Madamwar, D. (1997) Solid state fermentation of lignocellulosic waste for cellulose and β-glucosidase production by cocultivation of *Aspergillus ellipticus* and *Aspergillus fumigatus*. *Biotechnology Progress* **13**, 166-169.

Hansson T., Kaper, T., Oost, J., Vos, W.M. and Adlercreutz, P. (2001) Improved oligosaccharide synthesis by protein engineering of β-glucosidase CelB from hyperthermophilic *Pyrococcus furiosus*. *Biotechnology and Bioengineering* **73**, 203-210.

Harvey, A.J., Hrmova, M., Gori, R., Varghese, J.N. and Fincher, G.B. (2000) Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases. *Proteins: Structure, Function and genetics* 41, 257-269.

Hays, W.S., Jenison, S.A., Yamada, T., Pastuszyn, A., and Glew, R.H. (1996) Primary structure of the cytosolic β-glucosidase of guinea pig liver. *Biochemical Journal* 319, 829-837.

Henrissat, B. (1991) Classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* 280, 309-316.

Henrissat, B., and Bairoch, A. (1993) New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochemical Journal* 293, 781-788.

Henrissat,B. and Romeu, A. (1995) Families, superfamilies and subfamilies of glycosyl hydrolases. *Biochemical Journal* 311, 350-351

Himmel, M.E., Adney, W.S., Fox, J.W., Mitchell, D.J., and Baker, J.O. (1993) Isolation and characterization of two forms of β-D-glucosidase from *Aspergillus niger*. *Applied Biochemistry and Biotechnology* 39/40, 213-225.

Hoh, Y.K.(1992) Properties of β-glucosidase purified from *Aspergillus niger* mutants USDB0827 and USDB0828. *Applied Microbiology and Biotechnology* 37, 590-593.

Huber, R.E., Kurz, G., and Wallenfels, K. (1976) A quantitation of the factors which affect the hydrolase and transgalactosylase activities of β-galactosidase (*E. coli*) on lactose. *Biochemistry* 15, 1994-2001.

Huber, R.E., Gaunt, M.T., Sept, R.L., and Babiak, M.J. (1983) Differences in the effects of the pH on the hydrolytic and transgalactosylic reactions of β-galactosidase (*Escherichia coli*). *Canadian Journal of Biochemistry and Cell Biology* 61, 198-206.

Huber, R.E., Gaunt, M. T., and Hurlburt, K.L. (1984) Binding and reactivity
at the 'glucose'site of galactosyl-β-galactosidase *(Escherichia coli)*.
Archives of Biochemistry and Biophysics 234, 151-160.

Huber, R.E., and Brockbank, R.L. (1988) Binding specificity and reactivity
studies on a broad-specificity β-glycosidase from porcine kidney.
Biochemistry and Cell Biology 66, 830-838.

Huber, R.E., Gupta, M.N., and Khare, S.K. (1994) The active site and
mechanism of the β-galactosidase from *Escherichia coli*.
International Journal of Biochemistry 26, 309-318.

Invitrogen, (1997) *Pichia* easy-select manual.

Jacobson, R.H., Zhang, X., Dubose, R.F. and Matthews, B.W. (1994)
Three-dimensional structure of β-galactosidase from *E. coli*. *Nature*
369, 761-766.

Karthaus, O. (1994) Cellulase-catalyzed glycosylation reactions: simple
route towards a highly selective synthesis of oligosaccharides.
Journal of the Chemical Society Perkin Transactions 1: 1851-1857.

Kim, T., Park, C., Cho, H., Cha, S., Kim, J., Lee.,S., Moon, T., Kim, J.,
Oh. O., and Park, K. (2000) Role of the glutamate 332 residue in the
transglycosylation activity of *Thermus* maltogenic amylase.
Biochemistry 39, 6773-6780.

Kimura, I., and Shigeyuki, T. (1999) Subsite affinities of β-glucosidase
from *Aspergillus sojae* on various xylooligosaccharides. *Journal of
Bioscience and Bioengineering* 87, 572-575.

Kimura, I., Yoshioka, N., and Tajima, S. (1999) Purification and
characterization of a β-glucosidase with β-xylosidase activity from
Aspergillus sojae. *Journal of Bioscience and Bioengineering* 87, 538-
541.

Klahorst, S., Kumar, A., and Mullins, M.M. (1994) Optimizing the use of cellulase enzymes. *Textile Chemist and Colorist* 26, 13-18.

Koivula, A., Kinnari, T., Harjunpaa, V., Ruohonen, L., Teleman, A., Drakenberg, T., Rouvinen, J., Jones, T.A., and Teeri, T.T. (1998) Tryptophan 272: an essential determinant of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A. *FEBS Letters* 429, 341-346.

Koshland, D.E. (1953) *Biology Reviews* 28, 416-436

Kovar, J.(1987) Chromatofocusing of cellulolytic enzymes produced by *Trichoderma reesei, Aspergillus niger* and *Pleurotus ostreatus*. *Journal of Chromatography* 389, 322-326.

Kraulis, P.J., Clore, G.M., Nilges, M., Jones, T.A., Pettersson, G., Knowles, J. and Gronenborn, A.M. (1989) Determination of three-dimensional solution structure of the C-terminal domain of cellobiohydrolaseI from *Trichoderma reesei*. A study using nuclear magnetic resonance simulated annealing. *Biochemistry* 28, 7241-7257.

Kruger, N.J. (1994) The Bradford method for protein quantitation. *Methods in Molecular Biology* 32, 9-15.

Kumar, A., Yoon, M-Y., and Purtell, C. (1997) Optimizing the use of cellulase enzymes in finishing cellulosic fabrics.*Textile Chemist and Colorist* 29, 37-42.

Kuriki, T., Kaneko, H., Yanase, M. et al. (1996) Controlling substrate preference and transglycosylation activity of neopullanase by manipulating steric constraint and hydrophobicity in active center. *Journal of Biological Chemistry* 271, 17321-17329.

Kuriyama, K. (1995) Some properties of transglycosylation activity of sesame β-glucosidase. *Bioscience, Biotechnology and Biochemistry* 59, 1142-1143.

394

Lamed, R., Setter, E., and Bayer, E.A. (1983) Characterization of a
cellulose-binding, cellulase-containing complex in *Clostridium
thermocellum*. *Journal of Bacteriology* 156: 828-836.

Leah, R., Kigel, J., Svendsen, I. and Mundy, J. (1995) Biochemical and
molecular characterization of a barley seed β-glucosidase. *Journal of
Biological Chemistry* 270, 15789-15797.

Lee, I. (1996) Substrate-enzyme interactions in cellulase systems.
Bioresource Technology 58: 163-169.

Ly, H.D. and Withers, S.G. (1999) Mutagenesis of glycosidases. *Annual
Reviews in Biochemistry* 68, 487-522.

MacLeod, A.M., Linhorst, T., Withers S.G. and Warren R.A.J. (1994) The
acid/base catalyst in the exoglucanase/xylanase from *Cellulomonas
fimi* is glutamic acid 127: evidence from detailed kinetic studies of
mutants. *Biochemistry* 33, 6371-6376.

Macnaka, K., Kawai, G., Watanabe, K., Sunada, F., and Kumagai, I. (1994)
Functional and structural role of a tryptophan generally observed in
protein-carbohydrate interaction. *Journal of Biological Chemistry*
269, 7070-7075.

Mandels, M. (1985) Applications of cellulases. *Biochemical Society
Transactions* 13, 414-416.

Maranville, E. and Zhu, A. (2000) Assessment of amino-acid substitutions
at tryptophan 16 in α-galactosidase. *European Journal of
Biochemistry* 267, 1495-1501.

Mark, B.L., Vocadlo, D.J., Knapp, S., Triggs-Raine, B.L., Withers, S.G.
and James, M.N.G. (2001) Crystallographic evidence for substrate-
assisted catalysis in a bacterial β-hexosaminidase. *Journal of
Biological Chemistry* 276, 10330-10337,

McCleary, B.V. and Harrington, J. (1988) Purification of β-glucosidase
from *Aspergillus niger*. *Methods in Enzymology* 160, 575-583.

Mikhaylova, M., Wiederschain, G., Mikhaylov, V. and Aerts, J.M. (1996)
The enzymatic hydrolysis of 6-acylamino-4-methylumbelliferyl-β-D-
glucosides: identification of a novel human acid β-glucosidase.
Biochimmica et Biophysica Acta **1317**, 71-79.

Nagy, T., Simpson, P., Williamson, M.P., Hazlewood, G.P., Gilbert, H.J.,
and Orosz, L. (1998) All three surface tryptophans in type IIa
cellulose binding domains play a pivotal role in binding both soluble
and insoluble ligands. *FEBS Letters* **429**, 312-316.

Neelam, A. and Sexton, R. (1995) Cellulase (endo β-1,4-glucanase) and cell
wall breakdown during anther development in sweet pea (*Lathyrus
odoratus L*): isolation and characterization of partial cDNA clones.
Journal of Plant Physiology **146**, 622-628.

Nikolova, P.V., Duff, S., MacLeod, A. and Haynes, C.A. (1996)
Transglycosylation by wild type and mutants of a β-1,4-glycosidase
from *Cellulomonas fimi* (Cex) for synthesis of oligosaccharides.
Annals of the New York Academy of Science **799**, 19-25.

Notenboom V., Birsan, C., Nitz, M., Rose, D.R., Warren, R.A.J. and
Withers, S.G. (1998) Insights into transition state stabilization of the
β-1,4-glycosidase Cex by covalent intermediate accumulation in
active site mutants. *Nature Structural Biology* **5**, 812-817.

Ohnishi, N. and Tanaka T, (1995) Purification and properties of a novel
thermostable galacto-oligosaccharide-producing β-galactosidase from
Sterigmatomyces elviae CBS8119. *Applied and Environmental
Microbiology* **61**, 4026-4030.

Ohnishi, M., Okada, G., and Yazaki, T. (1998) Characterization of the
subsite structure of the β-glucosidase from *Aspergillus niger*, an
aspect of the mechanism of carbohydrate recognition. *Carbohydrate
Research* **308**, 201-205.

Okada, G. (1985) Purification and properties of a cellulase from *Aspergillus niger. Biological Chemistry* 49, 1257-1265.

Panchal, T. and Wodzinski, R.J. (1998) Comparison of glycsoylation patterns of phytase from *Aspergillus niger (A. ficuum)* NRRL 3135 and recombinant phytase. *Preparative Biochemistry and Biotechnology* 28, 201-217.

Parke, S.A., Birch, G.G., MacDougall, D.B. and Stevens, D.A. (1997) Tastes, structure and solution properties of D-glucono-1,5-lactone. *Chemical Senses* 22, 53-65.

Penttila, M., Lehtovaara, P., Nevalainen, H., Bhikhabhai, R. and Knowles, J. (1986) Homolgy between cellulase genes of *Trichoderma reesei*: complete nucleotide sequence of the endoglucanase I gene. *Gene* 45, 253-63.

Peshin, A., and Mathur, J.M.S. (1999) Purification and characterization of β-glucosidase from *Aspergillus niger* strain 322. *Letters in Applied Microbiology* 28, 401-404.

Pushalkar, S., Rao, K.K., and Menon, K. (1995) Production of β-glucosidase by *Aspergillus terreus. Current Microbiology* 30, 255-258.

Quiocho, F.A. (1986) Carbohydrate-binding proteins: tertiary structures and protein-sugar interactions. *Annual Reviews in Biochemistry* 55, 287-315.

Quiocho, F.A. (1993) Probing the atomic interactions between proteins and carbohydrates. *Biochemical Society Transactions* 21, 442-448.

Rashid, M.H. and Siddiqui, K.S. (1998) Thermodynamic and kinetic study of stability of the native and chemically modified β-glucosidase from *Aspergillus niger. Process Biochemistry* 33, 109-115.

Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis., P., Jones, T.A., Knowles, J.K. and Teeri, T.T. (1992) Investigation of

the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. *Proteins: Structure, Function, and Genetics* **14**: 475-482.

Riou, C., Salmon, J., Vallier, M., Gunata, Z., and Barre, P. (1998) Purification, characterization, and substrate specificity of a novel highly glucose-tolerant β-glucosidase from *Aspergillus orzae*. *Applied and Environmental Microbiology* **64**, 3607-3614.

Roberts, D.P., Denny, T.P., and Schell, M.A. (1988) Cloning of the egl gene of *Pseudomonas solanacearum* and analysis of its role in pathogenicity. *Journal of Bacteriology* **170**, 1445-1451.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K. and Jones, T.A. (1990) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* **249**: 380-385.

Rudick, M.J. and Elbein, A.D. (1973) Glycoprotein enzymes secreted by *Aspergillus fumigatus*. *Journal of Biological Chemistry* **248**, 6506-6513.

Rudick, M.J. and Elbein, A.D. (1975) Glycoprotein enzymes secreted by *Aspergillus fumigatus*: Purification and properties of a second β-glucosidase. *Journal of Bacteriology* **124**, 534-541.

Sanyal, A. (1988) Extracellular cellulolytic enzyme system of *Aspergillus japonicus*: 2. Purification and characterization of an inducible extracellular β-glucosidase. *Enzyme Microbial Technology* **10**, 91-99.

Schulein, M. (1998) Kinetics of fungal cellulases. *Biochemical Society Transactions* **26**, 164-167.

Shakin-Eshleman, S.H., Spitsalnik, S.L. and Kasturi, L. (1996) The amino acid at the X position of an Asn-X-Ser sequon is an important determinant of N-linked core-glycosylation efficiency. *Journal of Biological Chemistry* **271**, 6363-6366.

398

Shoda, S., Obata, K., Karthaus, O., and Kobayashi, S. (1993) Cellulase catalysed, stereoselective synthesis of oligosaccharides. *Journal of Chemical Society and Chemical Communications* **53**, 1402-1404.

Shoseyov,O. (1988) Endo- β -glucosidase from *Aspergillus niger* grown on a monoterpene glycoside-containing medium. *Biochemistry* **27**, 1973-1976.

Skory, C.D., Freer, S.N. and Bothast, R.J. (1996) Properties of an intracellular β-glucosidase purified from the cellobiose-fermenting yeast *Candida wickerhamii*. *Applied Microbiology and Biotechnology* **46**, 353-359.

Solovyeva, I.V., Ananjin, V.M., Boev, A.V., and Okunev, O.N. (1997) The controlled biosynthesis of cellobiase by *Aspergillus* fungi. *Process Biochemistry* **32**, 21-28.

Srivastava, S.K. (1984) Kinetic characterization of a crude β-D-glucosidase from *Aspergillus wentii* Pt2804. *Enzyme Microbiology and Technology* **6**, 508-512.

Stewart, J.C. and Parry, J.B. (1981) Factors influencing the production of cellulase by *Aspergillus fumigatus* (*Fresenius*). *Journal of General Microbiology* **125**, 33-39.

Stryer, L (1995) Biochemistry, 4th edition. W.H. Freeman and Company, New York.

Svensson, B., Jespersen, H., Sierks, M.R. and MacGregor, E.A. (1989) Sequence homology between putative raw-starch binding domains from different starch degrading enzymes. *Biochemical Journal* **264**, 309-311.

Teeri, T.T., Koivula, A., Linder, M., Wohlfahrt, G., Divne, C., and Jones, T.A. (1998) *Trichoderma reesei* cellobiohydrolases: why so efficient on crystalline cellulose. *Biochemical Society Transactions* **26**, 173-178.

Tomme, P., Beeumen, J., and Claeyssens, M. (1992) Modification of catalytically important carboxy residues in endoglucanase D from *Clostridium thermocellum*. *Biochemical Journal* **285**, 319-324.

Torronen, A., Kubicek, C.P., and Henrissat, B. (1993) Amino acid sequence similarities between low molecular weight endo-1,4-β-xylanases and family H cellulases revealed by clustering analysis. *FEBS Letters* **321**, 135-139.

Tull, D., Withers, S.G., Gilkes, N.R., Kilburn, D.G., Warren, R.A.J.,et al, (1991) Glutamic acid 274 is the nucleophile in the active site of a "retaining" exoglucanase from *Cellulomonas fimi*. *Journal of Biological Chemistry* **266**, 15621-15625.

Unno, T., Ide, K., Yazaki, T., Tanaka, Y., Nakakuki, T., and Okada, G. (1993) High recovery purification and some properties of a β-glucosidase from *Aspergillus niger*. *Biosciences, Biotechnology, Biochemistry* **57**, 2172-2173.

Vanderjagt, D.J., Fry, D.E., and Glew, R.H. (1994) Human glucocerebrosidase catalyses transglucosylation between glucocerebroside and retinol. *Biochemical Journal* **300**, 309-315.

Varghese, J.N., Hrmova, M., and Fincher, G.B. (1999) Three-dimensional structure of a barley β-D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure with Folding and Design* **7**, 179-190.

Wang, C., Eufemi, M., Turano, C. and Giartosio, A. (1996) Influence of the carbohydrate moiety on the stability of glycoproteins. Biochemistry **35**, 7299-7307.

Wang, Q., Tull, D., Meinke, A., Gilkes, N.R., Warren, R.A.J., Aebersold, R. and Withers, S.G. (1993) Glu280 is the nucleophile in the active site of Clostridium thermocellum CelC, a family A endo-beta-1,4-glucanase. *Journal of Biological Chemistry* **268**, 14096-14102.

Watanabe, T., Sato, T., Yoshioka, S., Koshijima, T. and Kuwahara, H. (1992) Purification and properties of *Aspergillus niger* β - glucosidase. *European Journal of Biochemistry* **209**, 651-659.

Watt, D.K., Ono, H., and Hayashi, K. (1998) *Agrobacterium tumefaciens* β-glucosidase is also an effective β-xylosidase, and has a high transglycosylation activity in the presence of alcohols. *Biochimica et Biophysica Acta* **1385**, 78-88.

Weis, W.I. and Drickamer, K. (1996) Structural basis of lectin-carbohydrate recognition. *Annual Reviews in Biochemistry* **65**, 441-473.

Wood, T.M. (1992) Fungal Cellulases. *Biochemical Society Transactions* **20**, 46-53.

Woodward, J. and Wiseman, A. (1982) Fungal and other β-D-glucosidases-their properties and applications. *Enzyme and Microbiology Technology* **4**, 73-79.

Ximenes, E.A. (1996) Production of cellulases by *Aspergillus fumigatus* and characterization of one β-glucosidase. *Current Microbiology* **32**, 119-123.

Yan, T-R., and Lin, C-L. (1997) Purification and characterization of a glucose-tolerant β-glucosidase from *Aspergillus niger* CCRC 31494. *Bioscience, Biotechnology, Biochemistry* **61**, 965-970.

Yan, T-R., Lin, Y-H., Lin, C-L. (1998) Purification and characterization of an extracellular β-glucosidase II with high hydrolysis and transglucosylation activities from *Aspergillus niger*. *Journal of Agriculture and Food Chemistry* **46**, 431-437.

Yazaki, T. (1997) Subsite structure of the β-glucosidase from *Aspergillus niger*, evaluated by steady-state kinetics with cello-oligosaccharides as substrates. *Carbohydrate Research* **298**, 51-57.

Appendix I

Dynamic Light Scattering Parameters

Table I-1 Parameters used for Dynamic Light Scattering

PARAMETER	SETTING
Hardware Configuration:	
Laser wavelength (A)	8276
Scattering angle (degrees)	90
Clean water count	6500
APD Bias Voltage:	
Normal (100%)	1401
Reduced (10%)	1261
Sensitivity Setting	100%
Count Rates:	
Threshold	6500
Medium	19500
High	100000
Photon Counts:	
Low	58500
Medium	175500
High	900000
Solvent:	
Name	aqueous buffer
Refractive Index	1.333
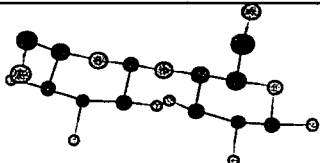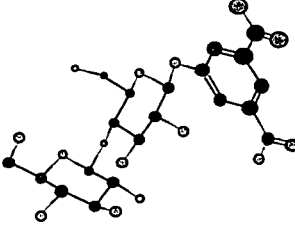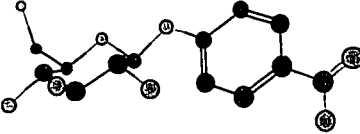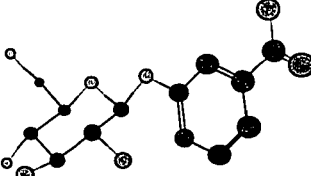
Viscosity	1.019
Viscosity Temperature coefficient	2024
Molecular weight model selected	standard curve
Standard Curve:	
Name	Globular
Rh factor	1.549
Power	2.426
Volume Shape Hydration:	
Name	bovine pancreas trypsin inhibitor (BPTI)
Specific Volume	0.718
Frictional Ratio	1.18
Correlator	20 CH(orig)
Maximum Acquisition Time (s)	10
S/N Threshold	1
pH	4.5
Concentration	2.61 (recombinant) / 5.05 (Sigma) mg/mL
Buffer	Sodium Acetate 0.15 M
Detergent	none

404

Appendix II

Appendix II: Carbohydrate Structures

All carbohydrate structures were drawn using Chem3D Pro™.
The same color scheme was used for all of the carbohydrates. Carbons
are black, oxygens are red, nitrogens are blue, and hydrogens (shown
only for monosaccharides) are pale blue. The common name (or name
used in the text) is shown first followed by the scientific nomenclature in
brackets. Cellobiose is shown for comparative purposes on each page.
The carbohydrate structures are grouped as substrates, products, and
inhibitors.
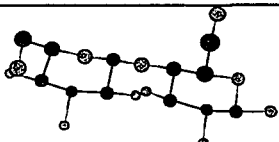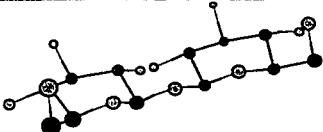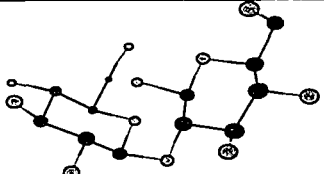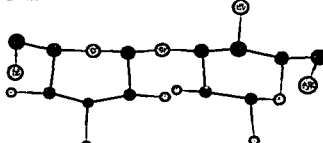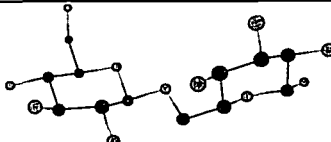
Aryl β-Glucosidase Substrates
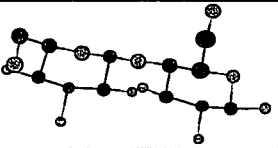
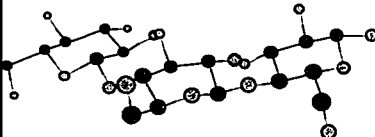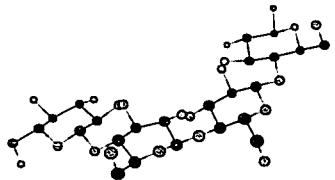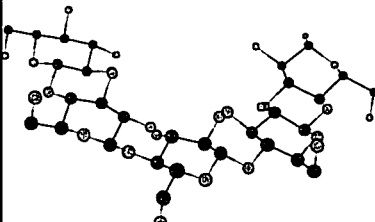| Sugar Name	Structure
Cellobiose (β-D-glucopyranose-(1→4)-β-D-glucopyranose)	
DNPC (3,4-dinitrophenyl-cellobiose)	
pNPGlc (p-nitrophenyl-β-D-glucose)	
oNPGlc (o-nitrophenyl-β-D-glucose)	

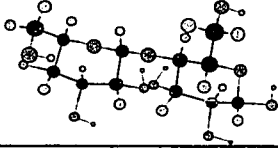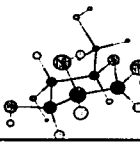| Sugar Name	Structure
Cellobiose (β-D-glucopyranose-(1-4)-β-D-glucopyranose)	
iso-trehalose (β-D-glucopyranose-(1-1)-β-D-glucopyranose)	
sophorose (β-D-glucopyranose-(1-2)-β-D-glucopyranose)	
laminaribiose (β-D-glucopyranose-(1-3)-β-D-glucopyranose)	
gentiobiose (β-D-glucopyranose-(1-6)-β-D-glucopyranose)	
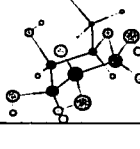
Cellooligosaccharide β-Glucosidase Substrates

Sugar Name	Structure
Cellobiose (β-D-glucopyranose-(1-4)- β-D-glucopyranose)	
cellotriose (β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose)	
cellotetraose (β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose)	
cellopentaose (β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose-(1-4)- β-D-glucopyranose)	
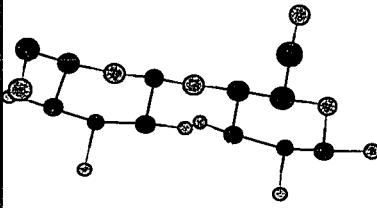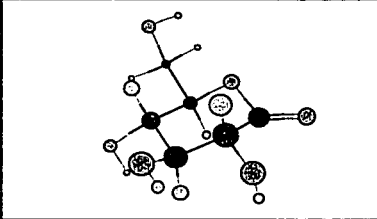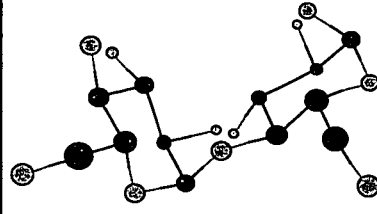
Carbohydrates Produced by the β-Glucosidase Transglycosylation Reaction with Cellobiose, pNPGlc, and Gentiobiose (in vertical order).

Sugar Name	Structure
Cellobiose (β-D-glucopyranose-(1-4)-β-D-glucopyranose)	
4-O-gentiobiosyl-glucose (β-D-glucopyranose-(1-6)-β-D-glucopyranose-(1-4)-β-D-glucopyranose)	
pNPGe (p-nitrophenyl-β-D-glucopyranose-(1-6)-β-D-glucopyranose)	
6-O-gentiobiosyl-glucose (β-D-glucopyranose-(1-6)-β-D-glucopyranose-(1-6)-β-D-glucopyranose))	

Monosaccharides used as Inhibitors of the β-Glucosidase [410]

Sugar Name	Structure
cellobiose (β-D-glucopyranose-(1-4)- β-D-glucopyranose)	
glucose (β-D-glucopyranose)	
mannose (β-D-mannopyranose)	
allose (β-D-allopyranose)	
galactose (β-D-galactopyranose)	

Other β-Glucosidase Inhibitors

Sugar Name	Structure
cellobiose (β-D-glucopyranose-(1–4)- β-D-glucopyranose)	
δ-gluconolactone	
maltose α-D-glucopyranose-(1–4)- α-D-glucopyranose	

412

Appendix III
NMR

Appendix III: NMR Values

The NMR chemical shift values for the various sugars were determined from either HMQC or TOCSY. The chemical shift values are reported in terms of ppm, according to convention. The sugars (that are larger than monosaccharides) are listed from the non-reducing end to the reducing end. Since the reducing end can be either α or β due to mutarotation there are to possibilities for each sugar. Both are shown. Each sugar of the carbohydrates is shown on a new line. Thus a trisaccharide will have chemical shift values on three lines (from non-reducing to reducing end). The linkage between monosaccharide units is indicated in brackets with and arrow, therefore (1→4) indicates that two monosaccharides are joined from the anomeric carbon (carbon #1) of the non-reducing end sugar to the carbon #4 of the second sugar. Glc indicates that the sugar is glucose. NP represents nitrophenol.

Table III-1. ^{13}C values for unknown sugars.

Compound	C-1	C-2	C-3	C-4	C-5	C-6
β-Glc-(1→6)-	103.5	74.2	76.8	70.8	77	62
β-Glc	97	75	77	70.5	75.8	69.9
β-Glc-(1→6)-	103.5	74.2	76.8	70.8	77	62
α-Glc	93	72.8	73.8	70.5	71.5	69.8
β-Glc-(1→4)-	104	74.5	77	70.8	77.5	62
β-Glc	97	75	75.5	79.8	76	61.5
β-Glc-(1→4)-	104	74.5	77	70.8	77.5	62
α-Glc	93	72.5	72.5	80	71.3	61.1
β-Glc-(1→6)-	102.2	73	75.7	69.8	76	60.8
β-Glc-(1→4)-	102.7	73	75.7	69.8	76	69
β-Glc	95.8	74	74.1	79	74.8	60
β-Glc-(1→6)-	102.2	73	75.7	69.8	76	60.8
β-Glc-(1→4)-	102.7	73	75.7	69.8	76	69
α-Glc	95.9	71.2	71.5	78.5	70.4	60
β-Glc-(1→6)-	104	74.5	76.9	71	77	62
β-Glc -NP	100.8	74	70.5	76.8	76.8	69.8
β-Glc-(1→4)-	103.8	74.5	76.8	70.7	77.2	61.8
β-Glc -NP	100.5	73.9	69	77.2	76.8	61

Table III-2. ^{1}H NMR values for unknown sugars.

Compound	H-1	H-2	H-3	H-4	H-5	H-6	
β-Glc-(1→6)-	4.49	3.3	3.49	3.38	3.45	3.72	3.91
β-Glc	4.64	3.24	3.4	3.42	3.62	4.14	3.84
β-Glc-(1→6)-	4.49	3.3	3.49	3.38	3.45	3.72	3.91
α-Glc	5.22	3.52	3.7	3.46	3.96	4.19	3.88
β-Glc-(1→4)-	4.5	3.34	3.52	3.44	3.48	3.74	3.9
β-Glc	4.65	3.28	3.63	3.63	3.6	3.82	3.85
β-Glc-(1→4)-	4.5	3.34	3.52	3.44	3.48	3.74	3.9
α-Glc	5.22	3.57	3.82	3.64	3.95	3.79	3.94
β-Glc-(1→6)-	4.55	3.33	3.53	3.42	3.5	3.55	3.92
β-Glc-(1→4)-	4.52	3.38	3.55	3.52	3.47	3.87	4.23
β-Glc	4.68	3.28	3.67	3.65	3.62	3.75	3.98
β-Glc-(1→6)-	4.55	3.33	3.53	3.42	3.5	3.55	3.92
β-Glc-(1→4)-	4.52	3.38	3.55	3.52	3.47	3.87	4.23
α-Glc	5.24	3.59	3.84	3.7	3.96	3.86	4.0
β-Glc-(1→6)-	4.48	3.32	3.47	3.4	3.39	3.39	3.9
β-Glc -NP	5.28	3.68	3.62	3.67	3.92	3.94	4.23
β-Glc-(1→4)-	4.58	3.38	3.58	3.48	3.53	3.8	3.95
β-Glc -NP	5.32	3.72	3.68	3.75	3.92	3.88	4.03

416

Appendix IV

Table IV-1. GenBank accession numbers for sequences used in sequence alignments.

Organism	GenBank Accession Number
Ajellomyces capsulatus	U20346
Aspergillus aculeatus	D64088
Aspergillus kawachi	AB003470
Aspergillus niger	AJ132386
Aspergillus wentii	P29090
Botryotinia fuckeliana	AJ130890
Coccidioides immitis (a)	U87805
Coccidiodes immitis (b)	AF022893
Cochliobolus heterostrophus	AF027687
Gaeumannomyces gaminis (a)	U35463
Gaeumannomyces gaminis (b)	U17568
Hordeum vulgare	U46003
Phanerochaete avenaria	AJ276675
Phanerochaete chrysosporium	AF036872
Pichia anomola	X02903
Saccharomyces fibuligera (a)	M22474
Saccharomyces fibuligera (b)	M22476
Septoria lycopersici	U24701
Trichoderma reesei	U09580