1-1-2004

Discovery of the chemical function of glycosidases: design, synthesis, and evaluation of mass-differentiated carbohydrate libraries and their application for bacterial identification

Yang Yu
Iowa State University

Discovery of the chemical function of glycosidases: Design, synthesis, and evaluation of mass-differentiated carbohydrate libraries and their application for bacterial identification

by

Yang Yu

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Organic Chemistry

Program of Study Committee:
Nicola L. Pohl, Major Professor
Richard Larock
George Kraus
Peter Reilly

Iowa State University

Ames, Iowa

2004

This is to certify that the master's thesis of

Yang Yu

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

Table of Contents

List of Abbreviations

^{13}C NMR	carbon nuclear magnetic resonance
^{1}H NMR	proton nuclear magnetic resonance
3-APBA	3-aminophenylboronic acid
4-MeU	4-methylumbelliferone
Bn	benzyl
DNA	deoxyribonucleic acid
DMAP	4-dimethylaminopyridine
DMSO	dimethyl sulfoxide
E. coli	*Escherichia coli*
EPO	human erythropoietin
ESI	electrospray ionization
Fuc*p*	fucopyranoside
Gal*p*	galactopyranoside
GalNAc	*N*-acetyl galactosamine
Glc*p*	glucopyranoside
GlcNAc	*N*-acetyl glucosamine
GTA	*N*-acetylgalactosaminyltransferase
GTB	galactosyltransferase
LC-MS	liquid chromatography-mass spectroscopy
MALDI	matrix-assisted laser desorption/ionization

Man*p*	mannopyranoside
ManNAc	*N*-acetyl mannosamine
MW	molecular weight
NESP	novel erythropoiesis stimulating protein
rHuEPO	recombinant human erythropoietin
RNA	ribonucleic acid
SIM	selected ion-monitoring
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy, free radical
TLC	thin layer chromatography
UDP	uridine 5'-diphosphate
UDP-GalNAc	uridine 5'-diphospho-*N*-acetylgalactose
UV	ultraviolet
Xyl*p*	xylopyranoside

Chapter 1

Discovery of the Chemical Function of Glycosidases: Design, Synthesis, and Evaluation of Mass-Differentiated Carbohydrate Libraries

The following work was published in *Org. Lett.* 2004, 6 (12), 2031-2033.

Yu, Y.; Ko, K-S.; Zea, C. J.; Pohl, N. L. *

I. Introduction

1. Importance of carbohydrates

Carbohydrate chemistry is one of the oldest research fields in chemistry. It developed slowly in the past one hundred years until recent decades when scientists found carbohydrates played important roles in cell-cell interaction (Ashwell 1982, Kornfeld 1987, Drickamer 1988, Bevilacqua 1992), immunological recognition (Daniels 2001, Moody 2001), pathogen attack (Feinberg 2001, Rudd 2001), and even protein folding and placement (Helenius 2001, Wells 2001). Glycobiology and glycomics-the study of the complete carbohydrate complement of an organism (Lesney 2002) have become a new research frontier in biochemistry and chemistry (Disney 2003, Love 2004, Seeberger 2004, Xu 2004.)

Compared with nucleic acids and proteins, carbohydrates are more complex in structure. One monosaccharide usually has three or four chiral hydroxyl groups to bind with other neighboring sugar residues. The linkage between two sugars can be α- or β- glycosidic bonds isomers. The carbohydrate chains can have various branches. This characteristic separates carbohydrates from the other biological macromolecules with primarily linear structures. The various and complex structures of carbohydrates mean that oligosaccharide chains contain almost unlimited information and provide numerous recognition signals between different molecules (Robyt 1998).

Cell surface carbohydrates are major components of the outer surface of mammalian cells and these carbohydrates are very often characteristic of cell types. Carbohydrate structures change dramatically during mammalian development (Fukuda 1985). Specific sets of

carbohydrates are expressed at different stages of differentiation and in many instances these carbohydrates are recognized by specific antibodies, thus providing differentiation antigens (Feizi 1985). In mature organisms, expression of distinct carbohydrates is eventually restricted to specific cell types, providing cell-type-specific carbohydrates (Sawada 1994). These carbohydrate structures are made by enzymes called glycosyltransferases that transfer sugars to a range of nucleophiles, such as other sugars, proteins, lipids. Another class of enzymes called glycosidases can remove or break down carbohydrate structures.

A simple example of the cell-surface-specific carbohydrates is human histo-blood group ABO antigens (Larsen 1990). The human blood groups ABO antigens are defined carbohydrate determinants found on the surface of the red blood cells and are largely responsible for the failure of mismatched blood transfusions. These ABO carbohydrate antigens occur on other cell types and are important in cell development, cell differentiation, and oncogenesis (Yamamoto 1990).

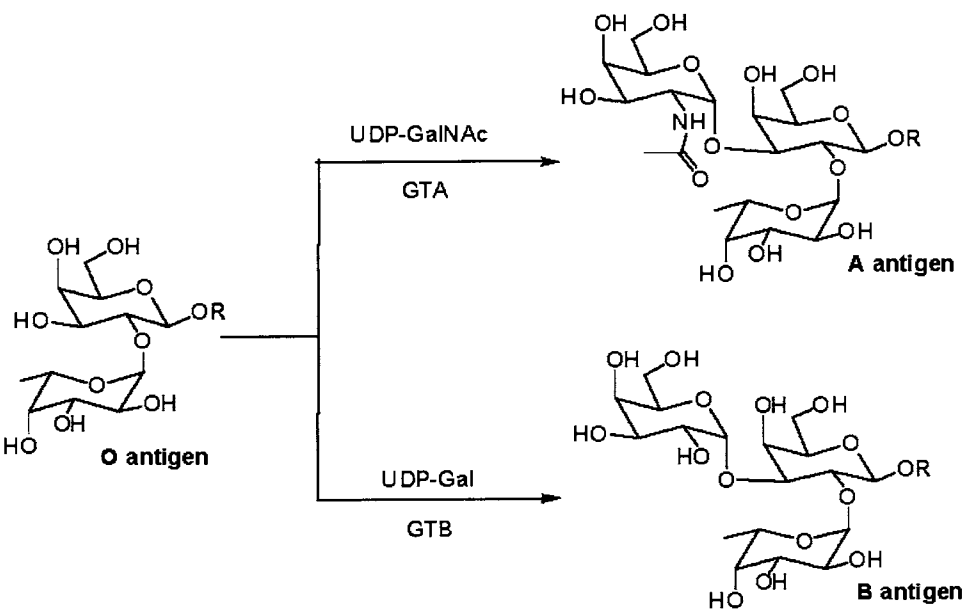


Figure 1. Human ABO blood group biosynthesis pathway. R = carbohydrate residues attached to a glycoprotein or glycolipid.

Blood group A individuals express α(1-3) *N*-acetylgalactosaminyltransferase (GTA), which catalyzes the transfer of GalNAc from donor UDP-GalNAc to the (O) H–precursor structure Fucα(1-2)Galβ-OR to give the A antigen GalNAcα(1-3)[Fucα(1-2)]Galβ-OR. Blood group B individuals express α(1-3)galactosyltransferase (GTB), which use the same (O) H-structure but catalyze the transfer of Gal from the UDP-Gal to make the B antigen Galα(1-3)[Fucα(1-2)]Galβ-OR (Figure 1) (Seto 1997). Blood group O individuals do not express either enzyme, and AB individuals express both. Alteration of only four of 354 amino acids entirely changes the carbohydrate recognition properties of the two glycosyltransferases responsible for the A versus B blood group antigen determinants.

Hence, analysis of primary sequence information of putative carbohydrate-recognizing enzymes alone does not reveal functional or substrate information. This substrate information is crucial to categorize these enzymes; glycosidases and glycosyltransferases are named according to the sugar whose glycosidic linkage is cleaved or joined. We need some other methods to validate the function of the many putative glycosidases and glycosyltransferases uncovered in genome sequencing projects.

The discovery of the function of glycosidases is of particular interest for several reasons. Appropriate sugar attachment and processing mediated by glycosidases appear critical for correct protein folding as well as for the ultimate location of proteins. Attached sugars are known to protect certain proteins from a variety of proteases. A well-known example of a glycoprotein application in the pharmaceutical industry is the development and characterization of a novel erythropoiesis stimulating protein (NESP). Studies on human erythropoietin (EPO) (Krantz 1991, Narhi 1991) found that human EPO is a 30400 Da heavily glycosylated protein hormone. 40% of the mass of this molecule is carbohydrate. The addition of carbohydrate or glycosylation is a post-translation process that leads to the addition of sugars to asparagine (*N*-linked) or Ser/threonine (*O*-linked) amino acids in the polypeptide. The glycosidases participate in this process and affect the folding and quality control of glycoproteins (Egrie 1986). Research demonstrated that there is a direct relationship between sialic acid-containing carbohydrate content of the molecule and its

serum half-life and in vivo biological activity, but an inverse relationship with its receptor-binding affinity. NESP was engineered to contain five *N*-linked carbohydrate chains, two more than recombinant human erythropoietin (rHuEPO) (Figure 2).

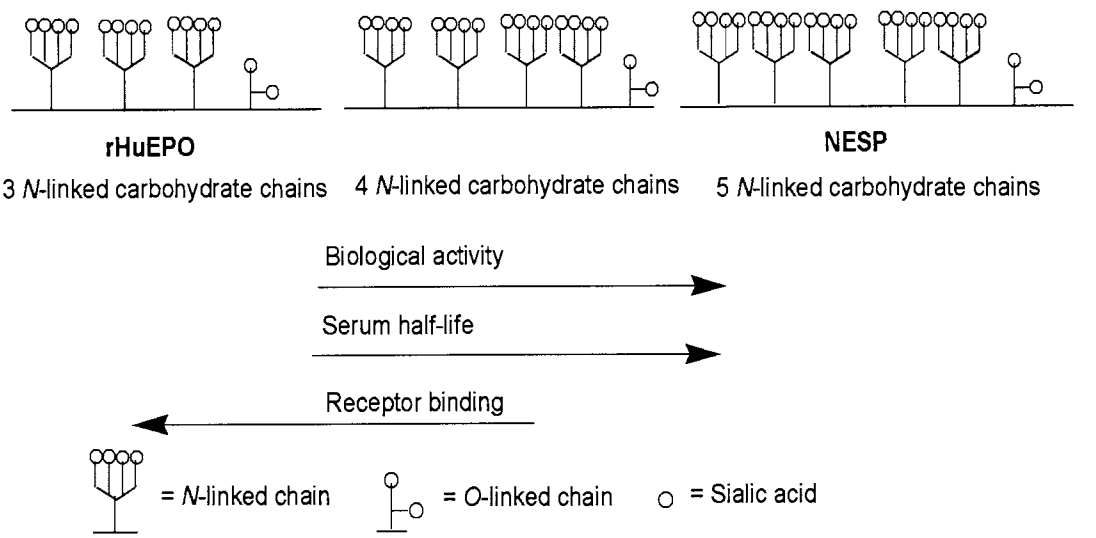


Figure 2. The development of new anti-anaemia drug erythropoiesis stimulating protein (NESP) and its characterization.

Its increased carbohydrate content prolongs NESP's serum half-life from 3 days to more than one week with greater in vivo potency. Glycosidases played an important role in the identification of the structure of human erythropoietin and the determination the relationship between the carbohydrate content and biological activities of NESP.

3. Glycosidases

From structure, function to biosynthesis and catabolism, glycosidases can be used to study almost all aspects of complex carbohydrates and glycoconjugates. Glycosidases can be divided into exo-glycosidases and endo-glycosidases. Exo-glycosidases cleave the monosaccharide from the end of the sugar chain, whereas endo-glycosidases cleave the sugar chain between specific carbohydrate sequences. In addition to cleaving the glycosidic

linkage, glycosidases can also catalyze the synthesis of oligosaccharide by transglycosylation reactions (Perugino 2004).

The glycosidic bond is the most stable linkage within naturally occurring biopolymers, with half-lives for spontaneous hydrolysis of cellulose and starch being in the range of five million years (Wolfenden 1998). The glycosidases accelerate this process for biological utility. The rate constant of glycosidases is up to 1000/S, making them among the most efficient catalysts. Structural information on glycosidases has steeply increased in recent years (Ly 1999). There are more than 2000 glycoside hydrolases available, according to the predicted amino acid sequences from genome sequencing projects and they are divided into over 70 different families on the basis of sequence similarities (Henrissat 1996). Although all of these enzymes catalyze the same reactions, their three-dimensional structures are quite diverse and they are divided into 30 different families (Davies 1995). The families which adopt similar folds have been attributed to the same clans (White 1997).

The inverting or retaining stereochemical outcomes result from distinctive mechanisms of glycosidic bond cleavage that are intrinsic to the geometry and stereochemistry of the enzyme active site (Gebler 1992). Thus an enzyme will either cleave with an inverting mechanism or a retaining mechanism, but not both. Enzymes that are related in sequence and structure generally all catalyze bond cleavage by a similar mechanism.

Both inverting and retaining mechanisms employ a pair of catalytic residues: a proton donor and a nucleophile/base. In both classes of mechanism the position of the proton donor is within hydrogen bonding distance from the glycosidic oxygen (McCarter 1994). In retaining enzymes the second catalytic residue is a nucleophilic carboxylate group that acts in a double displacement reaction, either as an ion pair stabilizing a transient oxonium ion transition state (lysozyme) or a nucleophile/leaving group for a transient covalent intermediate (β-glucosidases). In inverting enzymes a water molecule is used as a nucleophile in a single-step reaction following activation by a carboxylate acting as a catalytic base (Wang 1994).

In the inverting mechanism, the carboxyl group in one of the catalytic amino acids works as a general acid and the other acts as a base and the space between these two catalytic

residues is 8 Å ~ 10.5 Å, to accommodate the substrate *and* a water molecule to bind between them (Figure 3). Reaction occurs via a direct displacement of the leaving group by water, involving an oxocarbenium ion-like transition state (Sinnott 1990, White 1996).

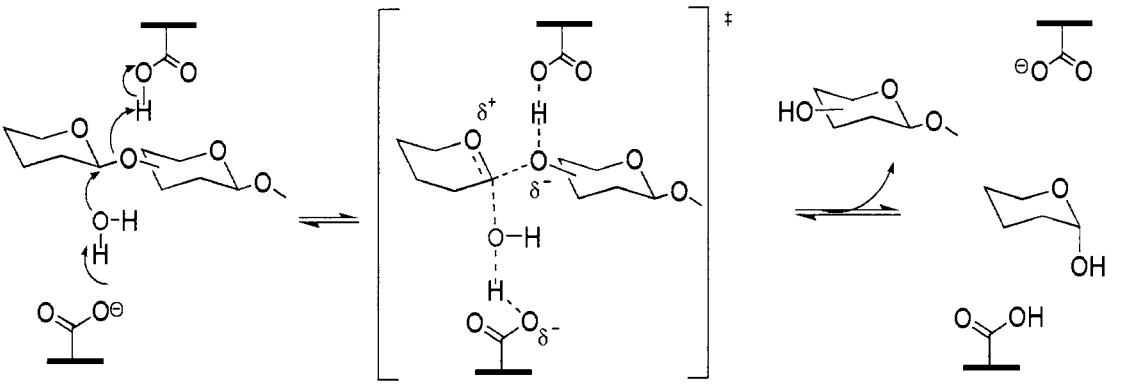


Figure 3. The inverting mechanism of glycosidases.

In the retaining mechanism, the space between the two catalytic carboxyl groups is only 5.5 Å. A double-displacement mechanism occurs, involving a covalent glycosyl-enzyme intermediate. In the first step, one of the carboxyl groups acts as a general acid catalyst to protonate the glycosidic oxygen concomitantly with bond cleavage. The other serves as a nucleophile to attack the glycosidic bond and form a covalent glycosyl-enzyme intermediate via an oxocarbenium ion transition state. In the second step, the carboxylic group which functions as a general acid in the first step works as a general base and deprotonates the incoming water molecule, which attacks at the anomeric center and releases the sugar, via oxocarbenium ion-like transition states(Figure 4).

The biggest barrier on the study of glycobiology is the lack of necessary technique and instrumentation to perform rapid and accurate analyses (Hirichs 1990, Imberty 1988). Lectins, which are the earliest protein tools to study carbohydrates and glycoproteins, can specifically recognize and bind with unique glycan patterns, just like antigen-antibody interaction. Lectins were first purified from plants and later research discovered they also

existed in mammals (Drickamer 1988, Baranski 1991). The lectins are widely used in the glycoprotein analysis until today.

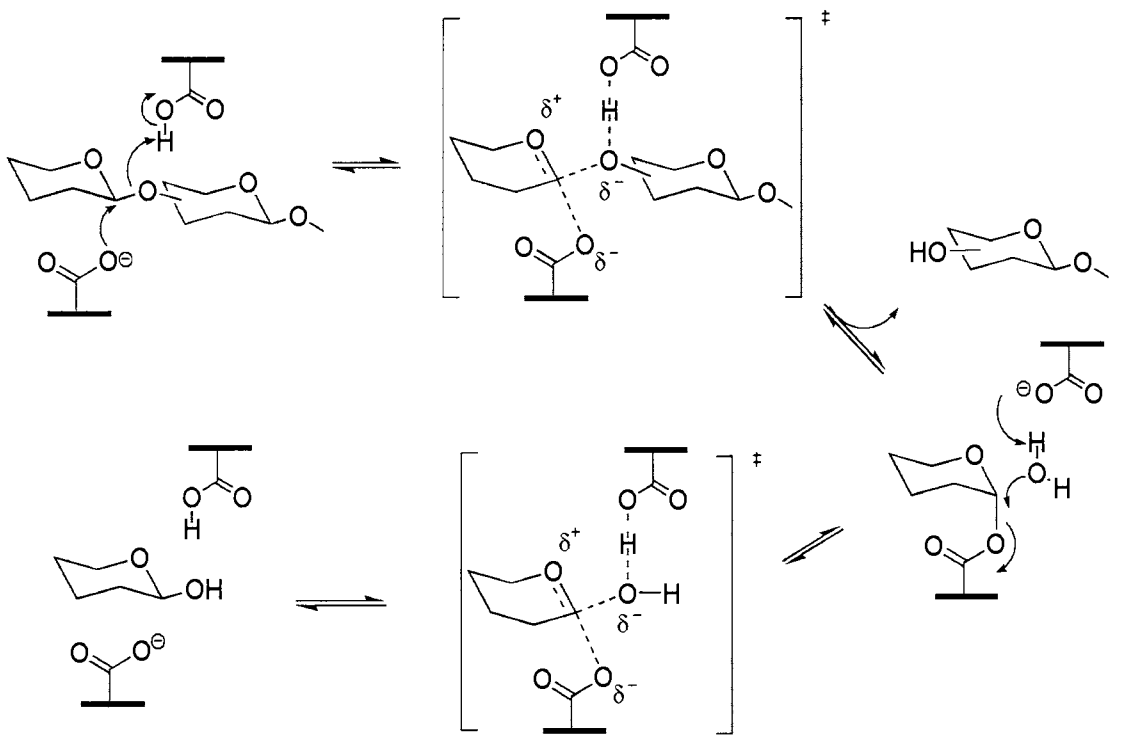


Figure 4. The retaining mechanism of glycosidases.

Glycosidases provide another powerful tool to further identify glycan structures (Dunn 1991, Dunn 1998), since some of the glycosidases cut only one terminal sugar attached in a specific way to the glycan chains and some of the glycosidases only cut between specific sugar sequences, just like DNA restriction enzymes cut the DNA sequences at specific points. This characteristic makes it possible to carry out the glycan sequence analysis and calculate the prior structures of glycoproteins.

Furthermore, glycosidases can be converted to glycosynthases, which catalyze the biosynthesis of oligosaccharides, by specifically mutated amino acid residues in the glycosidases (Mackenzie, 1998). Compared to the traditional chemical synthesis method, the main advantage of enzymatic synthesis of oligosaccharides is the region- and stereo-

selectivity that can be achieved without the need for protecting functional groups. Scientists now are disclosing more new glycosidases which can be used to cleave carbohydrate chains or to convert glycosynthases to biosynthesize oligosaccharides (Withers, 1992, Lawson 1998), which have potential applications as therapeutics (Maeda 1971, Estrada 1997). We need more glycosidases as mutation candidates to develop glycosynthases.

4. Chemical Proteomics

The completion of the human genome map caused an information explosion in human gene structure and provided sufficient recourse for scientists to identify and discover new drug targets based on genetic functions (Reiss 2001, Lander 2001, Venter 2001). Although the study of genomics is promising and achieves some exciting results, scientists notice that analysis of DNA or RNA is not enough to disclose all the functions of the gene and to understand the mechanism of diseases and other biological phenomena. Furthermore, the post-transcriptional and post-translational processing and modification of proteins, such as methylation and glycosylation, make the structure of proteins extremely different from that predicted by DNA sequence alone. We can only find the treasure behind genomics by studying the functions of the proteins encoded by gene. Proteomics is the study of the structure, function, expression, location, and regulation of proteins on a genome wide scale (Pandey 2000).

Chemical proteomics is a multidisciplinary research field, including synthetic organic chemistry, cell biology, biochemistry, and mass spectrometry (Norin 2001). The main core of this field is to design and synthesize a library of chemical probes to study the functions of distinct enzyme families within the whole complex proteome. It is a bridge built up between genomics and drug discovery (Myers 2001). The pharmaceutical industry uses the chemical proteomics method to screen huge amount of compounds to identify and validate drug candidates and to improve efficiency in developing drugs (Adam 2002, Jessani 2002).

5. Mass spectrometry

Mass spectrometry is a rapid, sensitive and accurate quantitative high-throughput approach for the direct monitoring of enzymatic reactions (Hunt 1992, Chait 1993, Krebs 1995). It monitors the change of substrates by the change of molecular weight instead of by radiolabeling or use of chromophore, which may change the reaction kinetics and requires tedious synthesis, and thus provides one of most convenient analytical techniques in the research of biological macromolecules.

In recent years, mass spectrometry has become a powerful tool to determine the substrate specificity of enzymes. Electrospray ionization (ESI) (Matayoshi 1990) and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (Bothner 1998, Siuzdak 1998) were applied for the qualitative and quantitative characterization of viral capsid dynamics, as well as the hydrolysis of *p*-nitrophenyl substrates with glycosidases and lipases and achieved accurately quantitative kinetics and dynamics data compared with UV results (Bothner 2000). New methods using ESI-MS to calculate the kinetics parameters of carbohydrate activity enzymes have been developed recently; these new methods easily work for a range of substrates (Zea 2003, Zea 2004).

6. The traditional method to study glycosidases

Currently the only method to delineate the chemical function of putative glycosidases coded in genomes is by individually testing the new protein with every possible substrate (Lee 1989, Hsieh 1995), usually as their *p*-nitrophenol conjugates (Vocadlo 2004). For example, in the kinetic measurement of α-glucosidase, 120 μM *p*-nitrophenoyl-α-glucopyranoside was incubated with α-glucosidase at 25 °C in 25 mM sodium phosphate buffer (pH 7.5) (Bothner 2000). The α-glucosidase cleaved the glycosidic bond and released the *p*-nitrophenolate, which was monitored by UV at 400 nm (Figure 6).

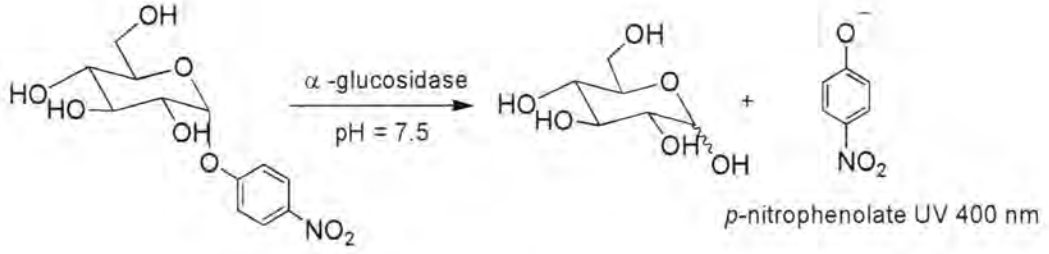


Figure 5. Using *p*-nitrophenyl-α-glucopyranoside to study the kinetics of α-glucosidase.

The disadvantage of this method is that all of the possible substrates need to be tested one by one and this increases the amount of required protein sample and time. Furthermore, the *p*-nitrophenol conjugate substrates are not stable at high temperature and decompose quickly and affect the accuracy when they are used to measure hyperthermophilic glycosidases.

II. Results and Discussion

To overcome the shortcomings of the current method, which relies on individual testing of possible *p*-nitrophenol conjugated substrates, we developed an alternative chemical proteomics approach using a synthetic mass-differentiated heat-stable substrate library with a mass spectrometry readout to discover the catalytic chemical function of the many putative glycosidases coded in genomes. Library components do not serve as reaction inhibitors and both primary and secondary enzyme substrates can be delineated.

The experiments include three parts: 1) design and synthesis of α and β-linked substrate libraries; 2) incubation of different glycosidases with α- or β-libraries at optimized conditions; and 3) monitoring of the enzymatic reactions by ESI-MS (Figure 7).

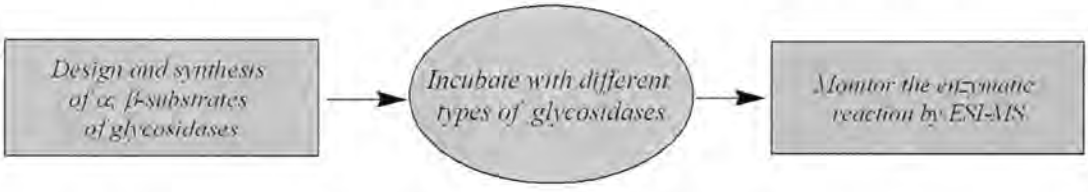


Figure 6. A chemical proteomics approach to study the function of putative glycosidases.

The basic strategy of a chemical proteomics approach (Figure 7) relies on the differentiation of all possible substrates for a given glycosidase by mass rather than by an ultraviolet, radioactive, or fluorescent tag, which would be difficult to distinguish in one reaction vial. In contrast to peptides that have established sequencing protocols (Wrighton 1996), monosaccharides are more difficult to identify as many structures differ only by stereochemistry and not by molecular weight or functional group (Veness 1996, Que 2003). To break this mass redundancy, carbohydrates with various mass-differentiated linkages would be needed. Unlike *p*-nitrophenol analogues, these glycosides should not be susceptible to high background hydrolysis rates, even at the elevated temperatures required to analyze putative proteins from hyperthermophilic sources. Finally, the library cannot contain any inhibitors of glycosidase activity for the success of this one-pot approach.

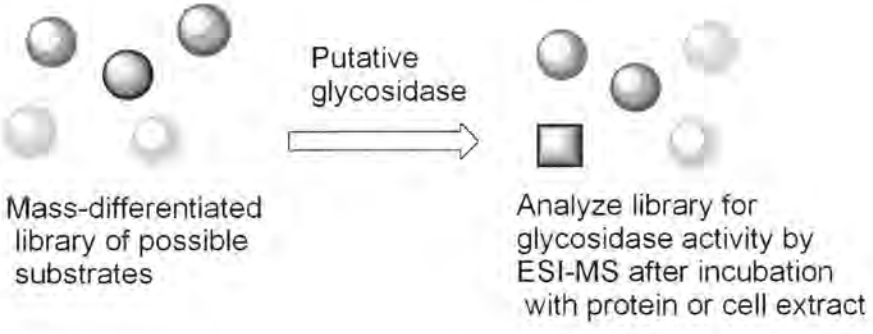


Figure 7. New strategy to determine the chemical function of the protein products of genes putatively assigned functions as glycosidases relies on incubation of the new protein with libraries of possible substrates that are inert to high background hydrolysis rates.

To test the feasibility of this strategy, small libraries of substrates were designed and synthesized. Although glycosidases are commonly assayed with bulky phenols as leaving groups, smaller anomeric leaving groups were desirable to avoid possible steric conflicts in the many potential active sites. Common sugars that would target the bulk of known glycosidases were chosen for an initial library and linkers were chosen for their stability as well as ease of synthesis.

The β-linked substrates were synthesized under Koenigs-Knorr conditions (Melvin 1999, Utille 2000, McGeary 2001, Davis 2000, Kikuchi 2001). First, the sugars were acylated by acetic anhydride and pyridine. Second, the acylated precursors were converted to the glycosyl bromide or chloride. Third, different alcohols were used to substitute the bromide or chloride groups at C-1 by using organometallic catalysts. Neighboring group participation assured high yields of the desired β-anomers (Figure 8).

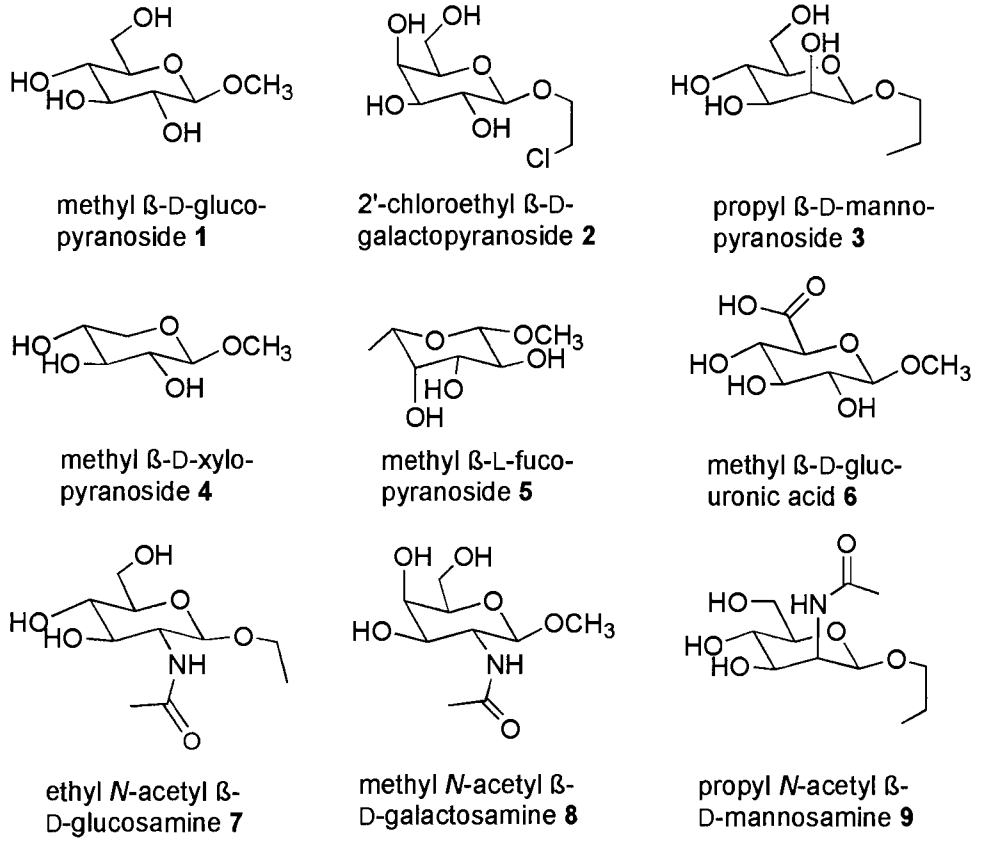
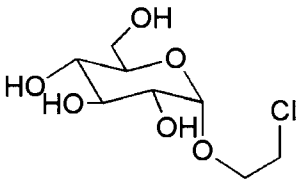


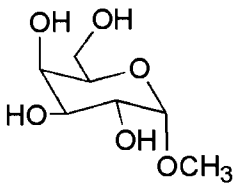
Figure 8. β-linked saccharides that are common known glycosidase substrates with various linkers to avoid any mass redundancy.

Fischer-Helferich conditions (Robyt 1998) were used and monosaccharides were heated with a variety of alcohols in presence of acidic resin, which served as a catalyst, to produce the mixture of α- and β-anomers. The α-pyranoside form is the predominant product can be

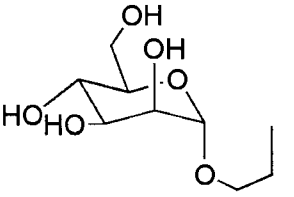
explained by the anomeric effect and the reaction is a thermodynamically controlled reaction While the furanose forms of the sugar react most rapidly and form the initial kinetic product, as the reaction progresses, the thermodynamically stable pyranosides are formed in which the α-pyranoside usually is the main product. This makes Fischer-Helferich method particularly useful in preparing the α-pyranoside anomer, which is difficult to synthesize by other methods. High yields of the α-pyranoside can be obtained by crystallization.



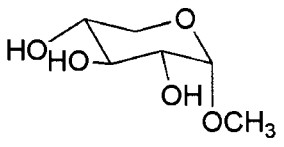
2'-chloroethyl α-D-glucopyranoside **10**



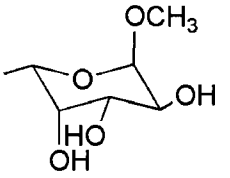
methyl α-D-galacto-pyranoside **11**



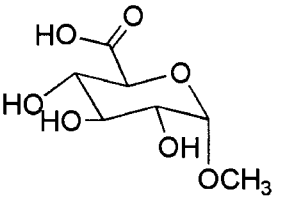
propyl α-D-manno-pyranoside **12**



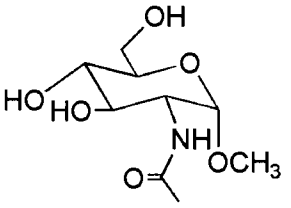
methyl α-D-xylo-pyranosidee **13**



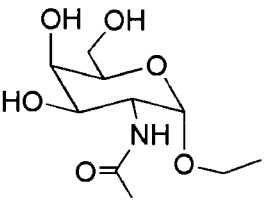
methyl α-L-fuco-pyranoside **14**



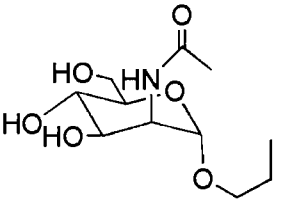
methyl α-D-gluc-uronic acid **15**



methyl *N*-acetyl-α-D-glucosamine **16**



ethyl *N*-acetyl-α-D-galactosamine **17**



propyl *N*-acetyl-α-D-mannosamine **18**

Figure 9. α-linked saccharides that are common known glycosidase substrates with various linkers to avoid any mass redundancy.

For several substrates, the α- and β-anomers can be separated by crystallization from different solvents. For example, the methyl α-D-xylopyranoside was crystallized from ethyl methyl ketone and methyl β-D-xylopyranoside was crystallized from ethyl acetate. For other substrates, such as D-glucose, D-galactose, and D-mannose substrates, their α- and β-anomers could not be separated directly by crystallization. Acetylation of all hydroxyl groups in the alkylated monosaccharides was accomplished by acetic anhydride and pyridine in order to dissolve them in organic solvents. After acetylation, the mixture of α- and β-anomers was separated using silica gel chromatography. Deacylation with sodium methoxide provided the desired α-linked saccharides (Figure 9).

With small libraries of both α- and β-linked carbohydrates in hand, we needed to discover conditions amenable for buffered enzymatic reactions to ionize compounds via electrospray for identification. Unfortunately, neutral carbohydrates are difficult to ionize (Sheeley 1998). The analysis of reaction aliquots diluted in aqueous trifluoroacetic acid was feasible in positive-ion mode; however, the *N*-acetylated sugars dominated the mass spectra. The amine-containing sugars (*m/z* = 258, 272, 286) were easily ionized and produced large peaks which dwarfed the peaks of neutral sugars (*m/z* = 187, 203, 217, 231, 245, 265) (Figure 10, a). Due to this signal suppression, we could not easily observe the change of the peaks by the MS spectrum.

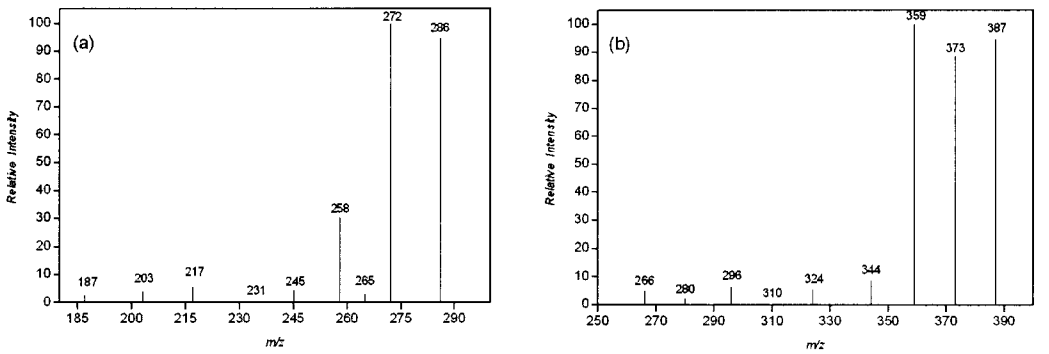
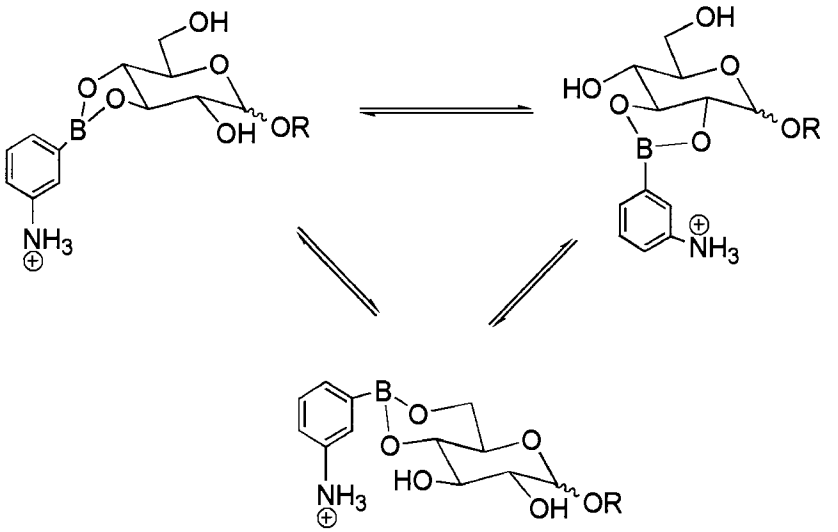


Figure 10. ESI-MS spectra in selected ion-monitoring positive-ion mode with β-linked library after quenching: (a) without 3-APBA; (b) with 3-APBA. Small neutral sugars are difficult to be ionized and their peaks were suppressed by the peaks of amine-containing sugars.

To solve the ionization problem, 3-aminophenylboronic acid (3-APBA) was added to the quench solution prior to analysis, 3-APBA which provides an easily ionizable amine for mass spectral analysis forms complexes with neutral sugars at low pH (Figure 11), This method improved the spectral readout of the sugar library dissolved in water as previously reported (Williams 2000) but worked poorly when the substrates were in typical enzyme reaction phosphate buffers. Even if we tried various quenching conditions, such as different concentrations of 3-APBA, different buffers, and different quenching solvents, the results were same (Figure 10, b).



R = H, alkyl group

Figure 11. The 3-Aminophenyl boronic acid (3-APBA) forms complexes with small neutral sugars at low pH (1% acetic acid/acetonitrile/water 1:50:50).

Interestingly, when β-*N*-acetyl glucosaminidase (from jack beans) was used in the enzymatic reaction, a very clear MS spectrum was obtained that showed all of the substrates at similar intensities. Comparison of the enzymatic reaction and quenching conditions of β-*N*-acetyl glucosaminidase with that of other glycosidases uncovered that the β-*N*-acetyl glucosaminidase was stored in an ammonium sulfate buffer.

To determine if ammonium sulfate buffer affects the ionization property of different carbohydrates, similar concentrations of ammonium sulfate buffer were added to the 3-APBA quenched enzymatic reactions. Both the amine-containing and neutral sugars were then found to be ionized at comparable magnitudes in positive-ion mode (Figure 12, a). In fact, altering the concentration of ammonium sulfate buffer changed the relative ratios of amine-containing to neutral sugars in the mass spectra. Peaks for sugars lacking a primary hydroxyl group such as xylose (**4**) and fucose (**5**) and negatively charged sugars such as glucuronic acid (**6**) ionize to a lesser degree as expected, but were still in the range of the other sugars for easy analysis. First, we used the positive-ion scan mode, which showed all ions of the compounds in the reaction mixtures, to detect the substrates in the library (Figure 12, b). After identifiing each of the substrates in the MS spectrum, we used selected ion-monitoring (SIM) positive-ion mode to monitor the enzymatic reactions (Figure 11, a).

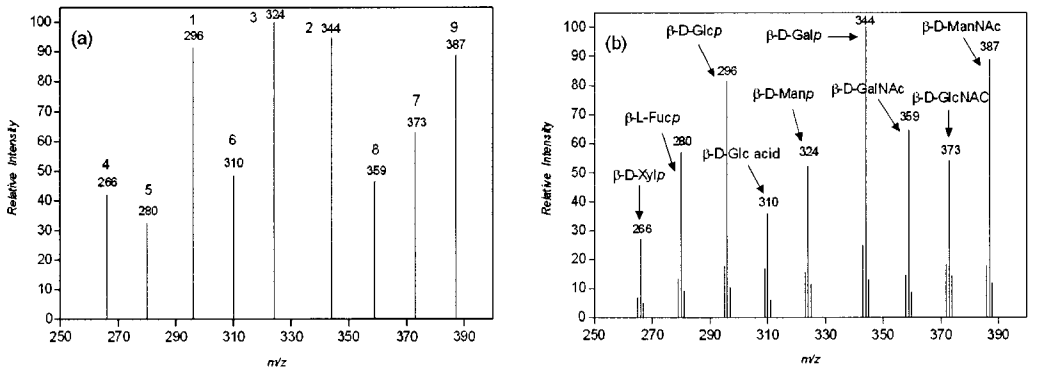


Figure 12. ESI-MS spectra with the β-linked library after quenching with 3-APBA and ammonium sulfate: (a) selected ion-monitoring positive-ion mode; (b) positive-ion scan mode.

We found there were three peaks for each of the substrates. The two smaller peaks besides the main peak are caused by isotopes of boron (Figure 12, b) as would be expected.

Carbohydrate libraries subsequently were tested for their ability to identify glycosidase enzymes by which substrate(s) they cleaved using the above method to consistently ionize the library components (Figure 13).

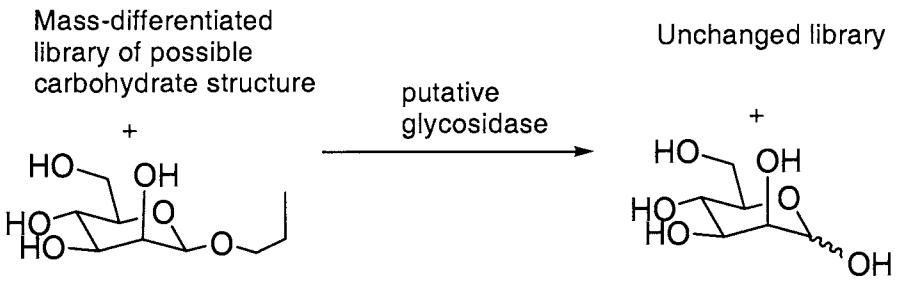


Figure 13. The glycosidase only selectively cleaves the specific substrate and leaves the other substrates in the library untouched.

The libraries were incubated with known glycosidases to probe four scenarios: (1) glycosidases with one known substrate (β-galactosidase from *Aspergillus oryzae*, α-glucosidase from rice, α-mannosidase from jack beans); (2) glycosidases with multiple known substrates or secondary activities (β-*N*-acetyl glucosaminidase from jack beans and the enzymes used in subsequent scenarios); (3) a glycosidase from crude versus pure sources (β-glucosidase from almond meal and chromatographically pure); and (4) a glycosidase from high-temperature sources (β-glucosidase from *Pyroccocus furious*). In each case, substrates for the enzymes significantly diminished after several hours of incubation with the libraries when compared to control reactions containing either no enzyme or heat-denatured enzyme, which were heated at 95 °C for 10 min to denature all of the glycosidases. At longer reaction times, secondary activities of the enzymes could be distinguished. Although not diagnostic, the cleaved sugar product could also be identified as both the mono- and diadduct of 3-APBA.

The structure of mono-adduct products are similar with the complexes formed between 3-APBA and alkylated monosaccharide (Figure 11), except the alkyl group was cleaved. The diadduct products are formed when one mole of monosaccharide forms complexes with two moles of 3-APBA (Figure 14)

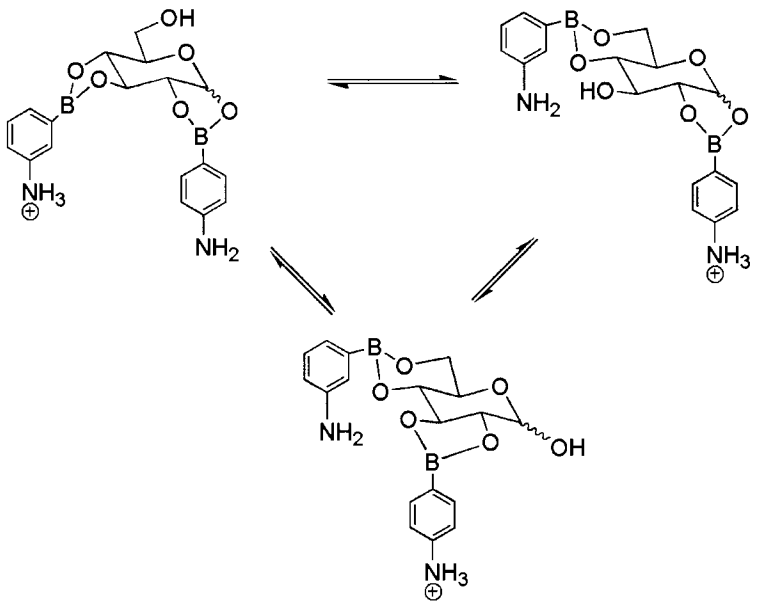


Figure 14. One mole of monosaccharide products forms complexes with two moles of 3-aminophenyl boronic acids (3-APBA) at low pH (1% acetic acid/acetonitrile/water 1:50:50).

More surprisingly, contaminating β-galactosidase activities and secondary activities of the enzymes could be distinguished easily from a chromatographically pure β-glucosidase, if the incubation was long enough, thereby providing a one-shot assessment of protein purity (Figure 15).

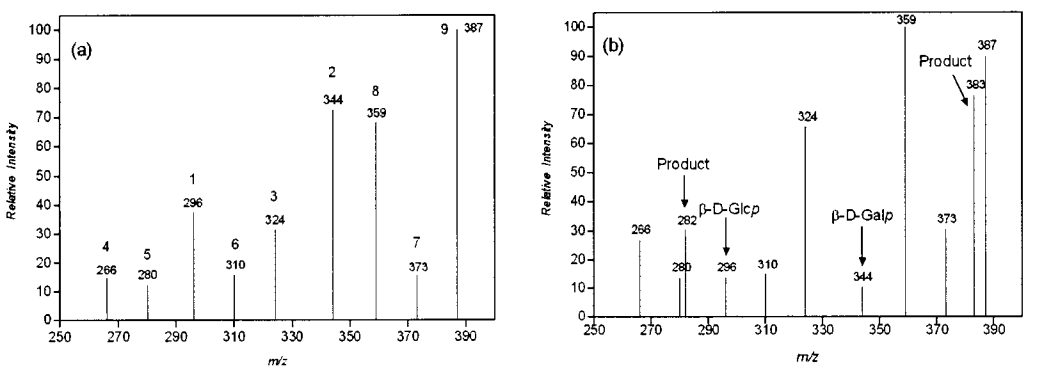


Figure 15. ESI-MS spectra in selected ion monitoring positive-ion mode of reaction after 10 h at 25 °C of β-glucosidase from almond with β-linked library after quenching with 3-APBA: (a) heat-killed control; (b) with enzyme. Glucose (**1**) is the specific substrate of this enzyme, but galactose (**2**) is cleaved, because of contaminated β-galactosidase impurity.

Mannosamine (**9**) also shows some change over time because of secondary activities of the enzymes.

We also measured a hyperthermophilic glycosidase, β-glycosidase I from *Pyrococcus furiosus*. In contrast to *p*-nitrophenol-linked sugar substrates (Figure 16), the libraries were stable for hours at 80 °C. The clear result using crude β-glucosidase from almond suggested that this library approach should also work with crude mixtures of glycosidases from cell extracts (See appendices for more results of enzymatic reactions).

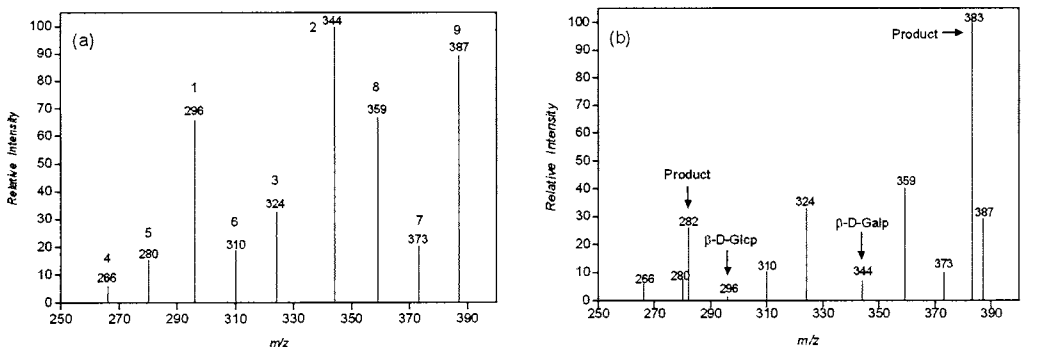


Figure 16. ESI-MS spectra in selected ion monitoring positive-ion mode of reaction after 3 h at 80 ° C of a glycosidase from *P. furiosus* with β-linked library after quenching with 3-APBA: (a) without enzyme; (b) with enzyme. Glucose (**1**) and galactose (**2**) are the primary substrates of this enzyme as seen by the significant diminishment of those peaks. *N*-Acetylmannosamine (**9**) also shows some change over time.

III. Conclusion

The new chemical proteomics approach can clearly identify specific substrates for glycosidases and thereby identify and name these newly sequenced putative glycosidases. In addition, these libraries are relatively stable even at high temperatures (80 °C - 90 °C) for several hours, which allow chemical function identification without requiring the optimal full-length substrate or buffer conditions of a glycosidase and to allow identification of

functions of proteins from hyperthermophilic sources. Once the cleaved glycosidic linkage is discovered for a particular enzyme, libraries containing that sugar linked to a wide variety of compounds can be used to further narrow down the substrate specificity of that enzyme. Protein mutants could also be screened for altered substrate specificities. Although the current libraries contain many of the sugars commonly found in mammals, the libraries can be expanded to include common plant and bacterial sugars. This chemical proteomics approach also can clearly identify glycosidase activity from cell extracts (almonds). All of these results encourage us to apply this strategy to the identification of bacteria. Because four bacterial strains previously have been differentiated only by their aminopeptidase activity profile (Basile 2002), a larger library of glycosidase substrates promises a mechanism to rapidly diagnose various species of cells by profiling their glycosidase activity. In addition, a mass-differentiated library approach suggests the application for the study of a wide range of other enzyme classes.

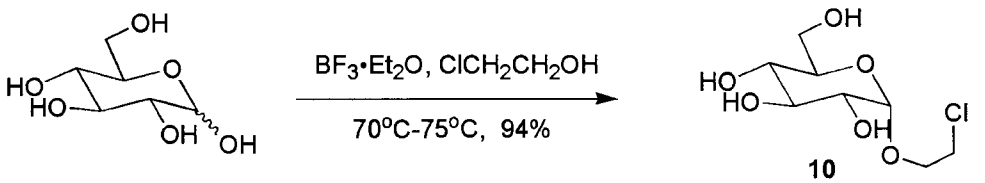
IV. Experimental Section

1. General methods:

Dichloromethane was distilled from calcium hydride after refluxing for more than 24 h under N_2. Diethyl ether was distilled from sodium metal and benzophenone under N_2. Nitromethane was distilled from calcium chloride under N_2. The alcohols were refluxed over magnesium metal and a catalytic amount of iodine before distillation. Amberlyst 15 ion-exchange resin, Dowex-50W X-8H^+-400 and Dowex WGR-2 ion exchange resin were washed repeatedly with methanol before use. All other commercial reagents and solvents were used as received without further purification except where noted. The reactions were monitored and the Rf values determined by analytical thin layer chromatography (TLC) with 0.25 mm EMScience silica gel plates (60-254). The developed TLC plates were visualized by immersion in *p*-anisaldehyde/EtOH/H_2SO_4 solution following by heating on a hot plate. Flash chromatography was performed with Selecto Scientific silica gel, 32-63 μM particle size.

All moisture-sensitive reactions were performed in flame or oven dried glassware under nitrogen atmosphere. Bath temperatures were used to record the reaction temperatures in all cases. All reactions were stirred magnetically at ambient temperature unless indicated otherwise. ^{1}H NMR and ^{13}C NMR spectra were obtained with a Bruker DRX300 at 300 MHz. All commercially available glycosidases—β-*N*-acetyl glucosaminidase (from jack beans), crude β-glucosidase and chromatographically pure β-glucosidase (both from almond meal), β-galactosidase (from *Aspergillus oryzae*), α-mannosidase (from jack beans), and α-glucosidase (from rice)—were purchased from Sigma (St. Louis, MO), except the β-glycosidase I from *P. furiosus*, which came from Calbiochem (La Jolla, CA).

Synthesis of *2'-chloroethyl α-D-glucopyranoside* (10)



Compound **10** was synthesized by a modification of a related procedure. BF_3•Et_2O (0.4 mmol, 58 μL) was added dropwise to a suspension of D-glucose (500 mg, 2.78 mmol) in $ClCH_2CH_2OH$ (5 mL, >5 equiv) under N_2. The reaction mixture was heated to 70-75 °C for 4 h. The reaction solution was cooled and the solvent was removed under reduced pressure. The residue was subjected to flash chromatography (5:1 chloroform/methanol) to obtain a mixture of anomers (572 mg, 2.36 mmol, 85%). The residue was dissolved in Ac_2O: pyridine (5 mL / 5 mL). After stirring for 15 h, the solvent was removed under reduced pressure to leave a white solid (940 mg, 2.27 mmol, 97%). The solid was recrystallized from methanol/H_2O to yield 2-chloroethyl-2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranoside (440 mg, 1.06 mmol, 47%). A solution of this compound (440 mg, 1.06 mmol) in anhydrous methanol (10 mL) was treated with sodium methoxide (4 mg) and the mixture was stirred for 3 h at ambient temperature, made neutral with Amberlite IR-120 (H^+) resin, and filtered. The filtrate was concentrated under reduced pressure and the resulting residue was subjected to

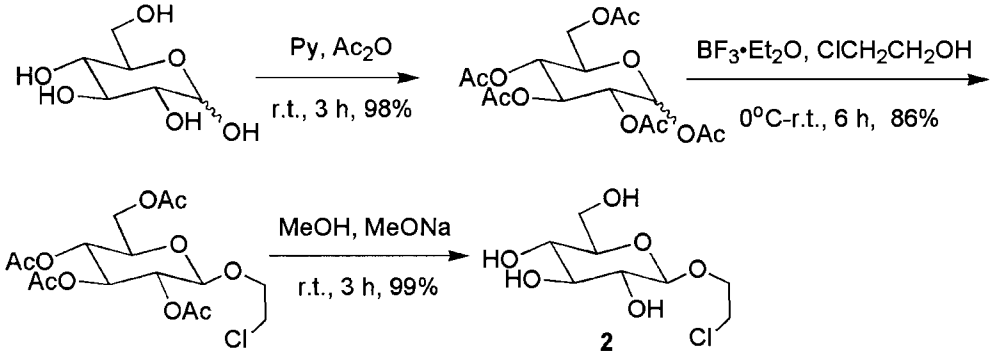
flash chromatography (5:1 chloroform/methanol) to yield **10** as a white solid (242 mg, 1.00 mmol, 94%).

^{1}H NMR (300 MHz, D_2O) δ 4.82 (d, $J_{1,2}$ = 3.6Hz, 1H), 3.82 (dt, J_t = 6.3 Hz, J_d = 10.8 Hz, 1H), 3.62-3.72 (m, 6H), 3.56 (t, J = 9.3 Hz, 1H), 3.41 (dd, $J_{1,2}$ = 3.9 Hz, $J_{2,3}$ = 9.9 Hz, 1H), 3.25 (t, J = 9.5 Hz, 1H).

^{13}C NMR (100 MHz, D_2O) δ 98.5, 73.2, 72.1, 71.4, 69.7, 68.4, 60.7, 43.5

MS (ESI-MS) *m/z* 265 $[M+Na]^+$.

Synthesis of *2'-chloroethyl* β-*D-galactopyranoside* (2)



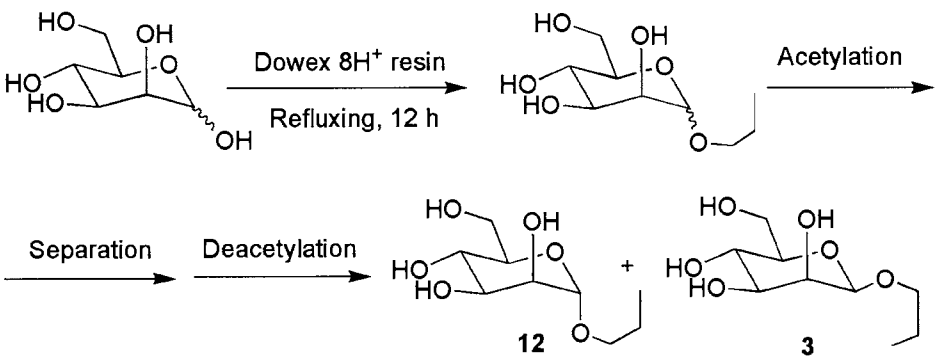
2'-chloroethyl-β-*D-galactopyranoside* (**2**) was synthesized as previously reported (Lee 1974).

^{1}H NMR (300 MHz, D_2O) δ 4.1 (dt, J_d = 11.4 Hz, J_t = 5.6 Hz, 1H), 3.82 (dt, J_d = 11.0 Hz, J_t = 6.2 Hz, 1H), 3.76 (dd, $J_{3,4}$ = 3.3 Hz, J4,5 = 0.9 Hz, 1H), 3.64-3.55 (m, 4H), 3.54 (m, 1H), 3.49 (dd, $J_{2,3}$ = 10.2 Hz, $J_{3,4}$ = 3.3 Hz, 1H), 3.38 (dd, $J_{1,2}$ = 7.7 Hz, $J_{2,3}$ = 10.2 Hz, 1H).

^{13}C NMR (100 MHz, D_2O) δ 103.3, 75.4, 72.8, 70.8, 70.3, 68.7, 61.1, 43.4

MS (ESI-MS) *m/z* 265 $[M+Na]^+$.

Synthesis of *propyl* β-*D-mannopyranoside* (3) and *propyl* α-*D-mannopyranoside* (12)



These two compounds were synthesized by a modification of a related procedure. A suspension of D-mannose (1.00 g, 5.56 mmol) in anhydrous 1-propanol (10 mL) containing 0.3 g Dowex-50 X-8H$^+$ resin was refluxed for 12 h. The resin was removed by filtration and washed with methanol. The filtrate and washings were combined and evaporated to yield 1-propyl-α/β-D-mannopyranoside as syrup (1.2 g, 5.4 mmol, 97%). The mixture was treated with excess acetic anhydride (5 mL) in pyridine (5 mL). After 6 h, the mixture was poured into 0.1M hydrochloric acid (20 mL) and extracted with ethyl acetate (3 x 15 mL). The organic layer was washed with saturated sodium bicarbonate solution (20 mL), water (20 mL), and brine (20 mL), dried over anhydrous magnesium sulfate, and concentrated under reduced pressure. The residue was purified by flash chromatography (3:2 hexane/EtOAc) to separate 1-propyl-2,3,4,6-*O*-tetra-acetyl-α-D-mannopyranoside and 1-propyl 2,3,4,6-*O*-tetra-acetyl-β-D-mannopyranoside. A solution of propyl 2,3,4,6-*O*-tetra-acetyl-β-D-mannopyranoside in anhydrous methanol (10 mL) was treated with sodium methoxide (3 mg) and the mixture was stirred for 3 h at ambient temperature, made neutral with Amberlite IR-120(H$^+$) resin, and filtered. The filtrate was concentrated under reduced pressure and the resulting residue was subjected to flash chromatography (7:1 chloroform/methanol) to yield compound **3** (790 mg, 3.56 mmol, 64% over two steps). The other anomer was subjected to the same deprotection conditions as compound **3** to yield compound **12** (123 mg, 0.55 mmol, 10% over two steps).

Propyl-β-D-mannopyranoside (**3**).

^{1}H NMR (300 MHz, D_2O) δ 4.71 (d, $J_{1,2}$ = 1.5 Hz , 1H), 3.78 (dd, $J_{1,2}$ = 1.8 Hz, $J_{2,3}$ = 3.6 Hz, 1H), 3.73 (dd, $J_{5,6'}$ = 1.2 Hz, $J_{6,6'}$ = 11.5 Hz), 3.65-3.56 (m, 2H, H-3), 3.54 (dd, $J_{5,6}$ = 5.7Hz, $J_{6,6'}$ = 11.5 Hz, 1H), 3.51-3.47 (m, 2H), 3.36 (dt, J_d = 9.6 Hz, J_t = 5.6 Hz, 1H), 1.46 (m, J = 7.5 Hz, 5.7 Hz, 2H), 0.76 (t, J = 7.5 Hz, 3H).

^{13}C NMR (100 MHz, D_2O) δ 99.7, 72.8, 70.8, 70.2, 69.6, 66.9, 61.1, 22.1, 10.1.

MS (ESI-MS) *m/z* 245 $[M+Na]^+$.

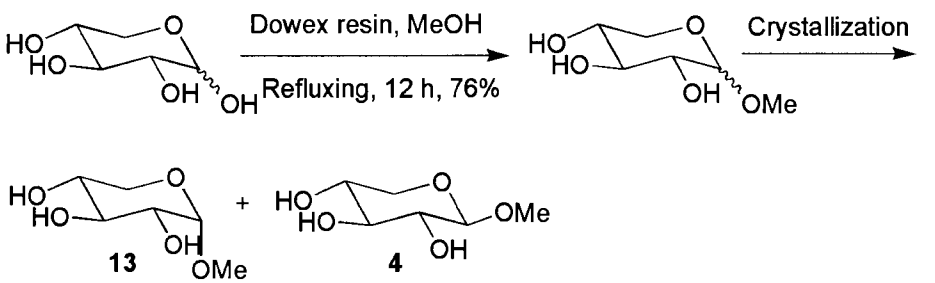
Propyl-α-D-mannopyranoside (**12**).

^{1}H NMR (300 MHz, D_2O) δ 4.55 (dd, $J_{1,2}$ = 0.9 Hz, 1H), 3.82 (dd, $J_{1,2}$ = 0.9Hz, $J_{2,3}$ = 3.2 Hz, 1H), 3.76 (dd, $J_{5,6'}$ = 2.1 Hz , $J_{6,6'}$ = 12.2 Hz, 1H), 3.67 (m, 1H), 3.56 (dd, $J_{5,6}$ = 6.4 Hz, $J_{6,6'}$ = 12.2 Hz, 1H), 3.50-3.41 (m, 2H), 3.40 (t, $J_{4,5}$ = 9.6 Hz, 1H), 3.20 (ddd, $J_{4,5}$ = 9.6 Hz, $J_{5,6}$ = 6.3 Hz, $J_{5,6'}$ = 2.1 Hz, 1H), 1.44 (m, 2H), 0.74 (t, J = 7.5 Hz, 3H).

^{13}C NMR (100 MHz, D_2O) δ 99.9, 76.4, 73.2, 71.8, 70.7, 67.0, 61.2, 22.3, 9.8.

MS (ESI-MS) *m/z* 245 $[M+Na]^+$.

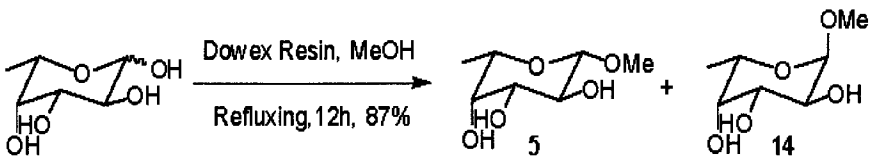
Synthesis of *methyl* β-*D-xylopyranoside* (4) and *methyl* α-*D-xylopyranoside* (13)



After D-xylose (1.00 g, 6.67 mmol) was refluxed with methanol (10 mL) and Dowex-50 X-8H$^+$ resin (0.5 g) for 10 h, fractional crystallization allowed separation of the resulting anomeric mixture. The β-anomer (367 mg, 2.24 mmol, 34%) was crystallized from ethyl acetate and the α-anomer (464 mg, 2.83 mmol, 42%) was crystallized from ethyl methyl ketone (Kvasnovsky 1998). NMR data matched those previously reported (Pouchert 1993).

MS (ESI-MS) *m/z* = 187 $[M+Na]^+$.

Synthesis of *methyl* β-*L-fucopyranoside* (5) and *methyl* α-*L-fucopyranoside* (14)



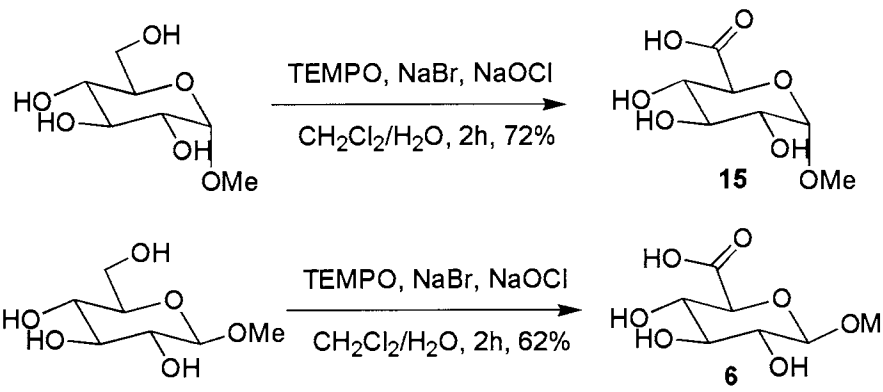
These two compounds were synthesized by a modification of a related procedure (Pouchert 1993, Zehavi 1972). After L-fucose (500 mg, 3.05 mmol) was refluxed with methanol (7 mL) and Dowex-50 X-8H$^+$ resin (0.25 g) for 12 h, compound **14** (277 mg, 1.28 mmol, 42%) was obtained by crystallization from ethanol. The mother liquid was concentrated under reduced pressure and the resulting residue was subjected to flash chromatography (7:1 chloroform/methanol) to yield methyl β-L-fucopyranoside **5** (70 mg, 0.39 mmol, 13%).

NMR data matched those previously reported (Knapp 1996, ThØgersen 1982).

MS (ESI-MS) *m/z* 201 $[M+Na]^+$.

Synthesis of *methyl* β-*D-glucuronic acid* (6) *methyl* α-*D-glucuronic acid* (15)

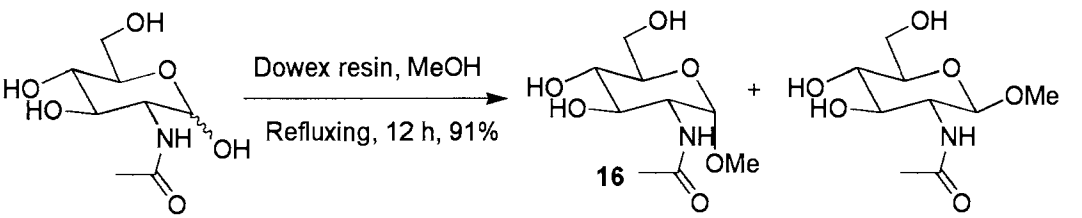
Methyl β-*D-glucuronic acid* (**6**) and *methyl* α-*D-glucuronic acid* (**15**) were synthesized as previously reported (Davis 1993, Melvin 1999).



NMR data matched those previously reported (Gorin 1975).

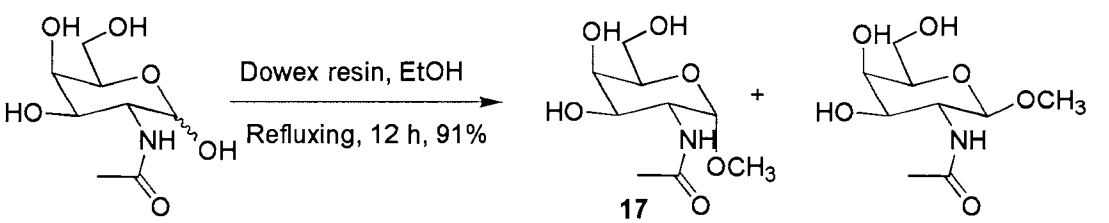
MS (ESI-MS) *m/z* 231 $[M+Na]^+$.

Synthesis of *methyl N-acetyl-α-D-glucosamine* (16)



Compound **16** was synthesized as previously reported (Galemmo 1983). The residue was subjected to flash chromatography (2:1 chloroform/methanol) to yield **16** (97 mg, 0.42 mmol, 91%). The ^1H NMR and ^{13}C NMR data of compound **16** matched those previously reported (Izumi 1987).

Synthesis of *ethyl N-acetyl-α-D-galactosamine* (17)



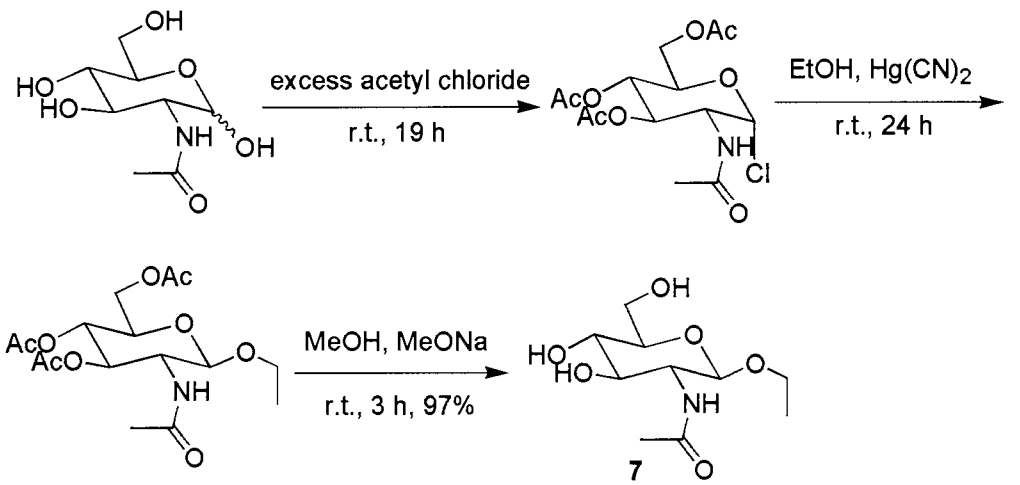
A solution of 2-acetamido-2-deoxy-D-galactose (100 mg, 0.46 mmol) in anhydrous ethanol (3 mL) was refluxed for 12 h with Dowex-50 X-8H$^+$ resin (230 mg). The resin was filtered and washed with ethanol (2 x 1 mL). The filtrate was evaporated and the residue was subjected to flash chromatography (3:1 EtOAc/EtOH) to yield **17** (71 mg, 0.29 mmol, 62%).

^{1}H NMR (300 MHz, D_2O) δ 4.75 (d, $J_{1,2}$ = 3.6 Hz, 1H), 4.00 (dd, $J_{2,3}$ = 10.6 Hz, $J_{3,4}$ = 3.6 Hz, 1H), 3.85-3.80 (m, 2H), 3.75 (dd, $J_{5,6}$ = 3.3 Hz, $J_{6,6'}$ = 11.8 Hz, 1H), 3.65-3.55 (m, 3H), 3.37 (m, 1H), 1.89 (s, 3H), 1.04 (t, J = 7.2 Hz, 3H).

^{13}C NMR (100 MHz, D_2O) δ 172.9, 100.2, 76.7, 74.4, 71.2, 62.8, 62.6, 54.4, 23.3, 15.2.

MS (ESI-MS) *m/z* 272 $[M+Na]^+$.

Synthesis of *ethyl N-acetyl-β-D-glucosamine* (7)



Compound **7** was synthesized by a modification of a related procedure. A mixture of 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (496 mg, 1.35 mmol) was dissolved in anhydrous ethanol (8 mL). The solution was treated with mercuric cyanide (1.2 equiv, 1.63 mmol, 413 mg) for 24 h at ambient temperature. The solvent was removed under reduced pressure. The resulting residue was taken up in chloroform (5 mL) and the

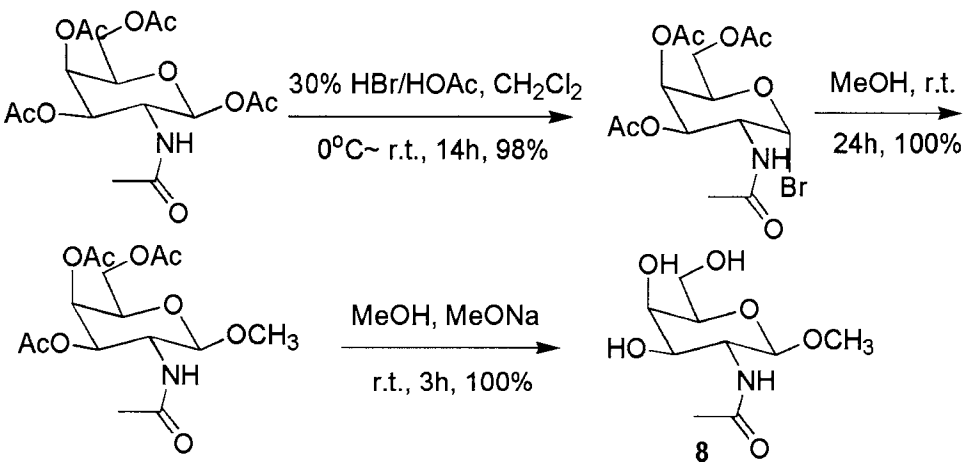
solid was removed by filtration. The filtrate was added chloroform (10 mL) and washed with water (5 mL). The organic layer was dried over anhydrous magnesium sulfate. The mixture was filtered and the solvent was removed under reduced pressure. The resulting residue was purified by flash chromatography (5:1 EtOAc/hexane) to yield ethyl 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranoside as a syrup (480 mg, 1.27 mmol, 94%). A solution of this compound (480 mg, 1.27 mmol) in anhydrous methanol (10 mL) was treated with sodium methoxide (3 mg) and stirred for 3 h at ambient temperature, neutralized with Amberlite IR-120(H+) resin, and filtered. The filtrate was concentrated under reduced pressure and the resulting residue was subjected to flash chromatography (3:1 chloroform/methanol) to yield **7** (270 mg, 1.08 mmol, 85%).

^{1}H NMR (300 MHz, D_2O) δ 4.38 (d, $J_{1,2}$ = 8.4 Hz, 1H), 3.80-3.70 (m, 2H), 3.59 (dd, *J*1,2 = 8.4Hz, *J*2,3 = 9.6, 1H), 3.53 (m, 2H), 3.38 (t, $J_{3,4}$ = 10.5Hz, 1H,), 3.30-3.24 (m, 2H), 1.88 (s, 3H), 1.00 (t, *J* = 7.1 Hz, 3H).

^{13}C NMR (100 MHz, D_2O) δ 174.7, 100.8, 76.0, 74.1, 70.0, 66.3, 60.9, 55.7, 22.3, 14.4.

MS (ESI-MS) *m/z* 272 $[M+Na]^+$

Synthesis of *methyl N-acetyl- β-D-galactosamine* (8)



Compound **8** was synthesized by a modification of a related procedure. 2-Acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-α-D-galactopyranoside (200 mg, 0.51 mmol) in CH_2Cl_2 (8 mL) was treated dropwise with 30% HBr/HOAc (0.51 mL, 5 equiv) at 0 °C for 1 h. The reaction mixture was stirred for 14 h as it warmed to ambient temperature. The reaction mixture was poured into ice water (10 mL). The organic layer was washed with saturated aqueous sodium bicarbonate (10 mL) and dried over anhydrous magnesium sulfate. Workup processes need to be finished within 15 min and all of the reagents and glassware were cooled to 4 °C prior to use to prevent decomposition. After filtration and solvent removed, the 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl bromide was stirred with anhydrous MeOH (5 mL) for 24 h. The reaction was neutralized with Dowex WGR-2 resin and filtrated; the filtrate was concentrated under reduced pressure to yield **8** as a white solid (118 mg, 0.50 mmol, 98%).

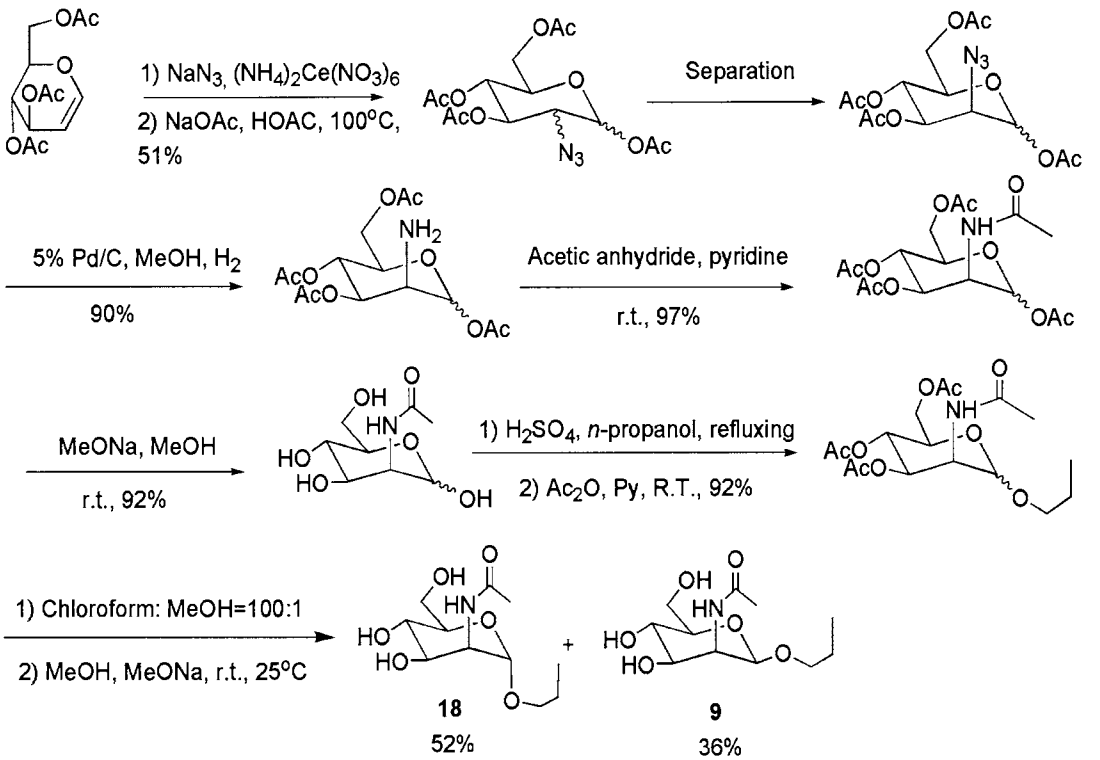
NMR data matched those previously reported (Grönberg 1994).

MS (ESI-MS) $m/z = 258$ $[M+Na]^+$.

Synthesis of *propyl N-acetyl-β-D-mannosamine* (9)

This compound was synthesized by a modification of a related procedure. 2-Acetamido-1,3,4,6-tetra-*O*-acetyl-2-deoxy-D-mannose (**19**)20 (600 mg, 1.53 mmol) was dissolved in 1-propanol (5 mL) and sulfuric acid (0.1 mL) was added dropwise to this solution. After being heated to 70 °C for 3 h, the mixture was treated with Dowex WGR-2(OH^-) resin until the pH reached ~10. The solvent was removed under reduced pressure and the resulting syrup was treated with excess acetic anhydride (3 mL) in pyridine (3 mL), since the acetyl groups were cleaved by the sulfuric acid. After 6 h, the mixture was poured into 0.1M hydrochloric acid (20 mL) and extracted three times with ethyl acetate (15 mL). The organic layer was washed with saturated aqueous sodium bicarbonate (25 mL), water (25 mL), and brine (25 mL), and then dried over anhydrous magnesium sulfate. The crude product was subjected to flash chromatography (gradient elution: 0 to 1% methanol/chloroform) to yield propyl 2-

acetamide-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-mannopyranoside (324 mg, 0.83 mmol, 54%) and propyl 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-mannopyranoside (216 mg, 0.55 mmol, 36%).



These two compounds were deacetylated individually by treatment with sodium methoxide (3 mg) in anhydrous methanol (10 mL) for 6 h at ambient temperature and neutralized with Amberlite IR-120(H^+) resin. The resin was removed by filtration and the solvent was removed under reduced pressure to yield **18** (197 mg, 0.75 mmol, 90%) and **9** (126 mg, 0.48 mmol, 87%).

Propyl N-acetyl-β-D-mannosamine (**9**).

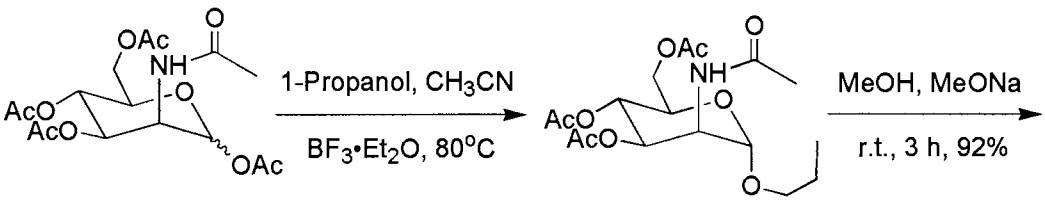
^{1}H NMR (300 MHz, D_2O) δ 4.95 (d, $J_{1,2}$ = 4.5 Hz, 1H), 4.22 (dd, $J_{1,2}$ = 4.5 Hz, $J_{2,3}$ = 2.7 Hz, 1H), 4.14 (t, J = 5.2 Hz, 1H), 3.97 (dd, $J_{2,3}$ = 2.7 Hz, $J_{3,4}$ = 8.7 Hz, 1H), 3.75 (ddd,

S10 1H), 3.66 (dd, $J_{5,6'}$ = 2.8 Hz, *J*6, 6' = 11.8 Hz, 1H), 3.56 (m, 1H), 3.47 (dd, $J_{5,6}$ = 4.8 Hz, $J_{6,6'}$ = 0.8 Hz, 1H), 3.36 (m, 1H), 1.90 (s, 3H), 1.42 (m, 1H), 0.72 (t, *J* = 7.5 Hz, 3H).

^{13}C NMR (100 MHz, D_2O) δ 174.5, 106.4, 80.1, 80.1, 71.6, 69.1, 63.3, 59.2, 22.3, 21.9, 9.7.

MS (ESI-MS) *m/z* 286 $[M+Na]^+$.

Synthesis of *propyl N-acetyl- α-D-mannosamine* (18)



18

This compound can also be synthesized by a modification of a related procedure.19 A solution of compound **19** (200 mg, 0.51 mmol) in CH_3NO_2 (1.5 mL), $BF_3{\cdot}OEt_2$ (8 μL), and 1-propanol (200 μL) was heated to 80 °C for 10 h. The mixture was cooled to ambient temperature and treated with CH_2Cl_2 (50 mL). The solution was washed with saturated aqueous sodium bicarbonate (20 mL) and water (20 mL). The organic layer was dried over anhydrous magnesium sulfate, filtered, and concentrated under reduced pressure. The residue was subjected to flash chromatography (4:1 EtOAc/hexane) to yield propyl 2-acetamide-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-mannopyranoside (148 mg, 0.38 mmol, 74%). This compound was treated with sodium methoxide (3 mg) in anhydrous methanol (4 mL) for 4 h at ambient temperature and neutralized with Amberlite IR-120(H^+) resin. The resin was removed by filtration and the solvent was removed by reduced pressure to yield compound **18** as a solid (92 mg, 0.35 mmol, 92%).

Propyl N-acetyl-α-D-mannosamine (18).

^{1}H NMR (300 MHz, D_2O) δ 4.63 (d, $J_{1,2}$ = 0.9 Hz, 1H), 4.17 (dd, $J_{1,2}$ < 1, $J_{2,3}$ = 4.8 Hz, 1H), 3.84 (dd, $J_{2,3}$ = 4.8, $J_{3,4}$ = 9.0, 1H), 3.72-3.61 (m, 2H), 3.54-3.42 (m, 3H, H-4), 3.34 (m, 1H), 1.88 (s, 3H), 1.45 (m, 2H), 0.76 (t, J = 7.5Hz, 3H).

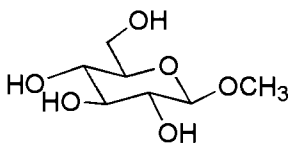
^{13}C NMR (100 MHz, D_2O) δ 174.8, 98.6, 72.4, 69.7, 69.3, 66.8, 60.5, 52.8, 22.1, 22.0, 10.0

MS (ESI-MS) *m/z* 286 $[M+Na]^+$.

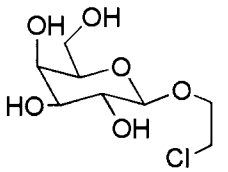
2. Mass Spectrometry Analysis

A Shimadzu LCMS 2010 quadrupole mass spectrometer (Shimadzu Scientific Instruments, Columbia, MD) equipped with an electrospray ionization (ESI) source was used in positive ion mode. The capillary temperature and the spray voltage were kept at 250 °C and 1.8 kV, respectively. The instrument was calibrated by direct infusion of polyethylene glycol (PEG) 200, 600, 1000 (1.5 μL/L, 2 μL/L, and 15 μL/L, respectively) and raffinose (50 mg/L) in water/methanol (1:1, v:v) containing ammonium acetate (0.19 mM), 0.1% formic acid, and 0.1% acetonitrile. For sample analysis, the solvent (50% water/acetonitrile) was constantly infused into the ion source at 250 μL/min by the attached Shimadzu HPLC pump and the samples were injected (5 μL). A preliminary MS chromatogram was obtained by scanning from 50-700 *m/z*. After analysis of the spectrum, the instrument was set for selected ion monitoring (SIM) in positive ion mode to increase the signal/noise ratio and all relevant *m/z* ions were monitored for analysis of the enzymatic reactions. Postrun software (LCMS Postrun version 2.02, Shimadzu Scientific Instrument, Columbia MD) was used to analyze the data from the ESI-MS chromatogram. Peaks were integrated to determine the relative intensity of each ion species.

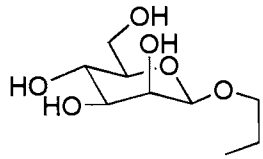
For the β-glycosidase, such as β-*N*-acetylglucosaminidase (from jack bean), β-glucosidase and chromatography pure β-glucosidase (both from almonds), β-galactosidase (from *Aspergillus oryzae*), each enzymatic reaction mixture contained:



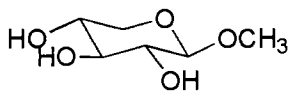
methyl ß-D-gluco-pyranoside **1**



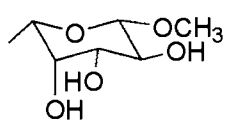
2'-chloroethyl ß-D-galactopyranoside **2**



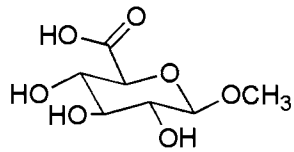
propyl ß-D-manno-pyranoside **3**



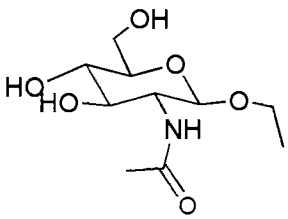
methyl ß-D-xylo-pyranoside **4**



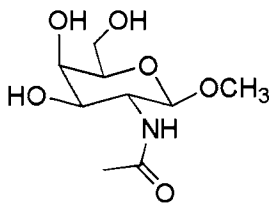
methyl ß-L-fuco-pyranoside **5**



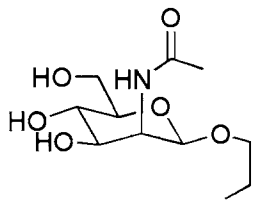
methyl ß-D-gluc-uronic acid **6**



ethyl *N*-acetyl ß-D-glucosamine **7**



methyl *N*-acetyl ß-D-galactosamine **8**



propyl *N*-acetyl ß-D-mannosamine **9**

a) β-Library containing nine substrates 7 μL (50 mM) for each substrate.

1) Methyl β-D-glucopyranoside, MW=194.18 *m/z*= 296 (M+3-APBA-2H_2O)

2) 2'-Chloroethyl β-D-galactopyranoside, MW=242.18 *m/z*=344 (M+102)

3) Propyl β-D-mannopyranoside, MW=222.25 *m/z*=324 (M+102)

4) Methyl β-D-xylopyranoside, MW=164.16 *m/z*=266 (M+102)

5) Methyl β-L-fucopyranoside, MW=178.18 *m/z*=280 (M+102)

6) Methyl β-D-glucuronic acid, MW=208.17 *m/z*=310 (M+102)

7) Ethyl *N*-acetyl β-D-glucosamine, MW=249.26 *m/z*=373 (M+102+Na-H)

8) Methyl *N*-acetyl β-D-galactosamine, MW=235.23 *m/z*=359 (M+124)

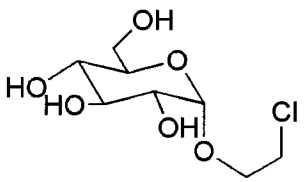
9) Propyl *N*-acetyl β-D-mannosamine, MW=263.29 *m/z*=387 (M+124)

b) Nanopure water 2 μL.

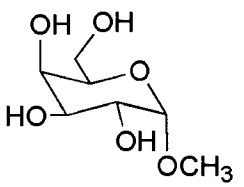
c) β-glycosidase 10 μL (0.5 U) in water.

d) Phosphate buffer 25 μL, 250 mM; the pH value is approximated based on the requirement of similar enzymes from that organism.

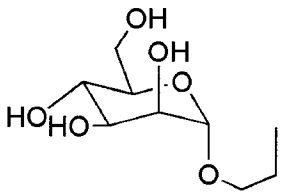
For the α-glycosidases, such as α-mannosidase (from jack jeans) and α-glucosidase (from rice), each enzymatic reaction included the α-library. This library contained:



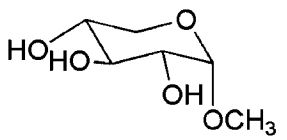
2'-chloroethyl α-D-glucopyranoside **10**



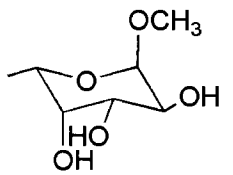
methyl α-D-galacto-pyranoside **11**



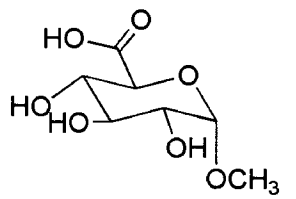
propyl α-D-manno-pyranoside **12**



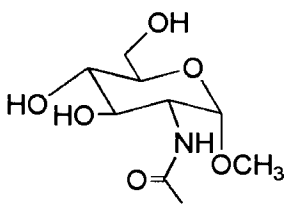
methyl α-D-xylo-pyranosidee **13**



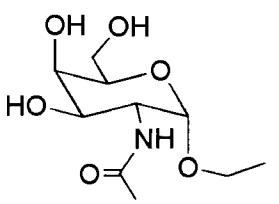
methyl α-L-fuco-pyranoside **14**



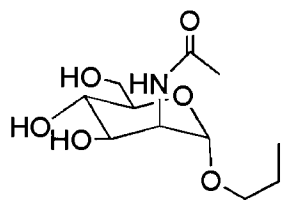
methyl α-D-gluc-uronic acid **15**



methyl *N*-acetyl-α-D-glucosamine **16**



ethyl *N*-acetyl-α-D-galactosamine **17**



propyl *N*-acetyl-α-D-mannosamine **18**

a) α-Library containing nine substrates 7 μL (50 mM) for each substrate.

1) 2'-Chloroethyl α-D-glucopyranoside, MW= 242.18 *m/z*=344 (M+3-APBA-$2H_2O$)

2) Methyl α-D-galactopyranoside, MW=194.18 *m/z* = 296 (M+102)

3) Propyl α-D-mannopyranoside, MW=222.25 *m/z* =324 (M+102)

4) Methyl α-D-xylopyranoside, MW=164.16 *m/z* =266 (M+102)

5) Methyl α-L-fucopyranoside, MW=178.18 *m/z* =280 (M+102)

6) Methyl α-D-glucuronic acid, MW=208.17 *m/z* =310 (M+102)

7) Methyl *N*-acetyl-α-D-glucosamine, MW=235.23 *m/z* =359 (M+102+Na-H)

8) Ethyl *N*-acetyl-α-D-galactosamine, MW=249.26 *m/z* =373 (M+124)

9) Propyl *N*-acetyl-α-D-mannosamine, MW=263.29 *m/z* =387 (M+124)

b) α-glycosidase, 10 μL (0.5 U) in water.

All other conditions mimicked those of the β-substrate library.

For the blank control sample, water (10 μL) was added instead of the glycosidase solution (10 μL). For the heat-killed control sample, the glycosidase was heated at 95°C for 10 min to denature the enzyme before addition to the reaction mixture. For the enzyme reaction sample, 0.5 units of enzyme were added. The reaction mixture was incubated at 25 °C for 5 min before it was initiated by the addition of the glycosidase. The reactions were incubated at a temperature normal for enzymes isolated from the particular organism and samples were quenched at 30 min, 1 h, 3 h, 6 h, or as noted. The mass spectrometry signal of the neutral sugars was small—only 2% as strong as that of the amine-containing sugars. The change of the intensity of neutral sugars could not be detected except by expansion of the mass spectrum peaks. Addition of 3-aminophenyl boronic acid (3-APBA, MW = 137 g/mol, Aldrich) to the quenched solution helped alleviate these ionization difficulties somewhat. The 3-APBA can form complexes with neutral sugars at low pH (1% acetic acid/acetonitrile) (Willams 2000). However, conditions as previously reported still resulted in weak ionization for the neutral sugars. We discovered that with addition of an ammonium sulfate $(NH_4)_2SO_4$ solution along with the 3-APBA to the quenched reaction solution the neutral sugars and amine sugars are ionized to the same extent

Furthermore, if the concentration of ammonium sulfate was changed from 20 mM to 100 mM, the degree of ionization of the neutral sugars relative to the aminosugars could be tuned. For example, when the concentration of ammonium sulfate is 100 mM, the peaks of neutral sugars are several folds higher than those of the amino-sugars. When the concentration of the ammonium sulfate is 20 mM, the peaks of the amino sugars are much higher than those of the neutral sugars. A concentration of ammonium sulfate from 50 mM to 70 mM allowed peaks of the different sugars to appear in the mass spectra at similar intensities. The peaks for the fucose, xylose and glucuronic acid library members usually still had weaker ionization intensities, however, because these sugars lack a hydroxyl group in position 5 or 6 and

cannot form strong complexes with 3-APBA. To quench the reaction mixtures, a 15 μL aliquot of the reaction mixture was added to 1% AcOH/acetonitrile (15 μL). After addition of 3-aminophenylboronic acid (15 μL, 50 mM stock solution in water), 2 μL aqueous ammonium sulfate (2.5 M) was added to the solution. The samples were centrifuged at 10,000 x g for 10 min to precipitate the protein before direct injection (5 μL) of the supernatant on the mass spectrometer as described above.

Due to the systemic variation of different batches of enzymatic reactions, we have set the threshold necessary to determine primary glycosidase activity to be a 50% reduction of the peak intensity. Secondary glycosidase activity can be detected by increasing the enzymatic reaction time or adding more glycosidase to the reactions.

Chapter 2

Glycosidase Activity Profiling by a Chemical Proteomics Approach for Bacterial Identification

I. Introduction

1. The importance of bacterial identification

Microorganisms play important roles in the field of agriculture, nutrition and food, biotechnology, and medicinal discovery (Okabe 2003, Personne 2004, Smith 2004). The detection of biological hazards in drinking water, food, and the environment; the medical diagnosis of pathogens; and the study of metabolic pathways require rapid, sensitive, and accurate methods to identify and characterize microorganisms. The requirement of building up an easy, fast and reliable method to identify the harmful bacteria or pathogens is very necessary, because the traditional method, which involves microscopy, serology, molecular tools and culture (Houpikian 2002), is time consuming and tedious.

2. The new microarray technologies of bacterial identification

To improve the efficiency and sensitive of microorganism identification and characterization, scientists have developed several new techniques, such as DNA and protein microarray technologies (Behr 1999, Call 2001, Harrington 2000), phenotype microarray methods (Hood 1996, Strauss 1997), and infrared spectroscopy (Al-Khaldi 2004). The most successful example of DNA microarray method uses universal primers to amplify highly conservative gene sequence, such as 16S rDNA encoding gene, to identify bacterial genotype (Relman 1999). This broad-range PCR technique has disclosed two novel bacteria to identify bacillary angiomatosis and Whipple disease (Kerkhoff 1999). Although these new technologies provide alternative methods to identify the bacteria, some limitations, such as high background interference (noise), differences in labeling of nucleic acid with different fluorescent dyes and controversial and irreproducible results still exist with these methods.

3. The MS-biomarker method of bacterial identification

Mass spectrometry (MS) offers another efficient approach to characterize and identify the bacteria pathogens by detecting chemotaxonomic biomarkers (Black 1994, Cole 1991), such as phospholipids, lipopolysaccharides, oligosaccharides, and proteins (Fang 1998, Krishnamurthy 1999, Liu 1998) in different strains of cells. Different MS equipment, such as fast atom bombardment (FAB), matrix-assisted laser adsorption ionization (MALDI) (Krishnamurthy 2000), and electrospray (ESI) (Goodacre 1999, Vaidyanathan 2001) were used to pursue optimized methodology which could identify different species of bacteria efficiently, but generated limited success, because of the shortcomings of FAB and MALDI-MS. For example, there are still spectral reproducibility problems in MALDI-MS from sample preparation errors (Domin 1999, Wang 1998). Liquid-Chromatography-Electrospray-Mass Spectrometry (LC-ESI-MS) was finally chosen to generate chromatographic profiles of specific protein biomarkers in a bacteria sample; this method is more applicable, stable and reproducible (Vaidyanathan 2001). The results of this system clearly demonstrate that LC/ESI-MS bacteria identification and analysis can be performed rapidly with high resolution and provide a promising on-line sample preparation and totally automatic analytical process in the future (Krishnamurthy 1999). However, the quantity and number of protein biomarker components change greatly when different cell extraction conditions are used and the protein profiles obtained from cell suspensions also changed significantly depending on freeze-thaw and lyophilization operations. The standard cell extraction and disruption methods are necessary for this system to prevent inhomogeneous results in different research groups and equipment and increase the reproducibility of the experimental data (Vaidyanathan 2001).

4. The enzyme activity profiles method of bacterial identification

Rapid identification of bacteria by investigating microbial enzyme activity profiles appeared more than twenty year ago (Bascomb 1975) by using enzyme activities such as phosphatases, esterases, aminopeptidases and glycosidases to identify bacteria and fungi (Basile 1993, Kilian 1976, Lee 1975, Maddocks 1975). The potential specific enzymes in the

bacteria can serve as another promising target besides the protein biomarkers. The main idea is to investigate the activity of the enzyme instead of enzymes molecule themselves to identify bacteria by incubating the bacteria extract or whole bacteria, which contain the corresponding specific enzymes, with chromogenic or fluorogenic substrates and monitor the enzymatic reactions by UV or fluorescence spectroscopy. This method assumes all of the cleavage or change of the specific substrates are caused by corresponding enzymes instead of any other factors.

In fact, this method provides a rapid, sensitive, and automated system for bacteria detection and identification. The results of enzymatic reactions are reliable and reproducible. Even if the amount of enzymes is not very high in cells, enzymes still can amplify their functions by cleaving a large number of specific substrates within a short time and generate the same homogenous results. All of these characteristics make the activity of enzymes sensitive and easy to be detected and this system is convenient to be developed for a kit or automated system. Nevertheless, there is little work using enzyme activity profiles to detect bacteria in recent years, because this method usually needs more than ten chromogenic or fluorogenic substrates of the enzymes to incubate with the cell suspension individually. Although each of the enzymatic reactions only needs a small amount of bacteria, all of the tests with the potential substrates increases the amount of bacteria the same fold as the number of substrates which need be tested. This disadvantage increases the cost and time of this method.

In order to overcome the current shortcoming of enzyme activity profiling, we decided to try a chemical proteomics approach to identify bacteria. Glycosidases are the enzymes which hydrolyze carbohydrate chains and they are an extremely large class of hydrolases that widely exist from prokaryotes to eukaryotes. (More background is in chapter 1). For example, *E. coli* and *Bacillus cereus* contain gene sequences for β-glucosidase, β-galactosidase and β-mannosidase (Carbohydrate-active enzyme database, http://afmb.cnrs-mrs.fr/CAZY). Glycosidases enzyme profiles have been used in the identification and differentiation of the different species of bacteria. Several experiments which used around ten *p*-nitrophenyl substrates (chromogenic Substrates) or 4-methylumbelliferone substrates

(fluorogenic substrates) to study the glycosidases activity profiles and rapidly identify different species or subspecies of bacteria have been published (Kilian 1976, Kilian 1978, Godsey 1981). In these studies, the bacterial were suspended directly with the respective *p*-nitrophenyl substrates or 4-MeU substrates at optimized temperature and pH. The enzyme activity was monitored by detection of the color produced by *p*-nitrophenol or fluorescence of 4-methylumbelliferone. This method is fast, convenient and reproducible but labor-intensive.

II. Results and Discussion

To improve the enzyme activity profile method, we present an alternative chemical proteomics approach using a synthetic mass-differentiated substrate library with mass spectrometry to readout the glycosidase activity present in a cellular extract. *Escherichia coli* K12, *Bacillus cereus,* and *Psuedomonas aeruginosa* were chosen as the target bacteria, because they could be easily identified and differentiated by profiling with aminopeptidase activity (Basile 2002). *E. coli* is a type of bacteria that normally lives in the digestive tract of humans and animals (Todar 2002). Some strains of *E. coli* can cause diarrhea and gastrointestinal problems. *Bacillus cereus* causes two types of food-borne intoxications which are characterized by nausea, vomiting, abdominal cramps and diarrhea. *Pseudomonas aeruginosa* causes many various infections, such as respiratory system infections, soft tissue infections, bone and joint infections, and gastrointestinal infections. *Pseudomonas aeruginosa* infection is a serious problem in patients hospitalized with fatality rate up to 50 percent in cancer, cystic fibrosis, and burn patients (Todar 2002).

In chapter 1, a small α- and β-substrate library for glycosidases containing eighteen monosaccharides with different molecular weight was designed and synthesized. Millimolar concentrations of each substrate were used to incubate with 0.2~0.5 Unit of known glycosidases in the same reaction vial at optimized temperature and pH in buffer. ESI-MS was used to monitor the enzymatic reactions at specified time intervals (Yu, 2004). In each case, substrates for the enzymes significantly diminished after several hours of incubation with the libraries when compared to control reactions. More surprisingly, contaminating

glycosidase activities could be identified from a crude protein mixture and different glycosidase activities involved in a crude cell extract (almond) could be detected clearly (See the details in chapter 1). All of the promising results encourage us to try this strategy for the identification of different types of bacteria by studying their glycosidase activities.

The enzyme activity profiling experiments were set up and the reaction mixture was incubated at optimized conditions for 24 h in total, but there were no further changes after 6 h or 12 h. The mass spectrums of enzyme activity profiles for *E. coli* K12, *Bacillus cereus* are shown in Figure 1 and 2 (See page 32-34 for mass assignment).

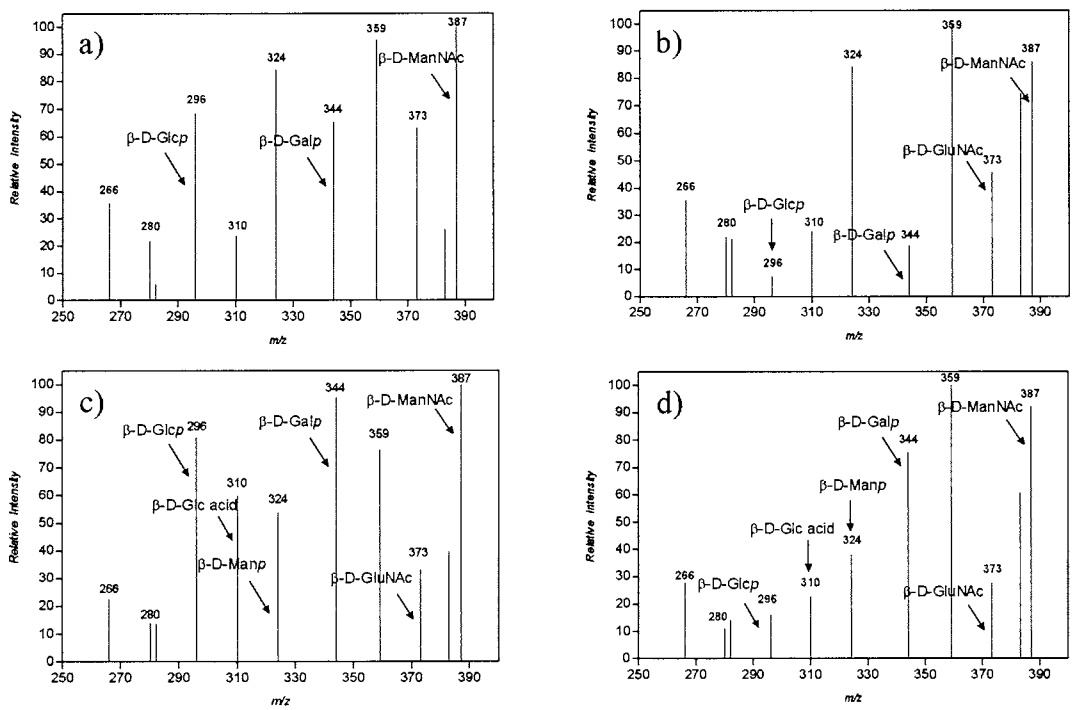


Figure 1. ESI-MS spectrum in selected ion-monitoring positive-ion mode of β-linked substrates after 12 h at 37 °C incubation with cell extract of *E. coli* K12 or *Bacillus cereus*: (a) heat-killed control of *E. coli* K12; (b) with cell extract of *E. coli* K12. (c) heat-killing control of *Bacillus cereus*; (b) with cell extract of *Bacillus cereus*. The cell extract of *E. coli* K12 selectively cleaved β-D-glucose, β-D-galactose, β-D-glucosamine, and β-D-mannosamine substrates in different intensities; in addition to these four substrates, the cell extract of *Bacillus cereus* also selectively cleaved β-D-glucuronic acid and β-D-mannose substrates.

Since different alkyl groups were introduced in the substrates to eliminate mass redundancies, there is no signal overlap for all of the substrates. After each signal of the substrates was identified by ESI-MS scanning positive mode, we used the selected ion monitoring positive-ion mode to detect the reduction of substrate concentrations as the indication of the enzymatic activity for the cell extract. Same molecular weight of monosaccharides and corresponding alcohols are the products after the cleavage of each different substrate. The alcohol products did not be detected because of their low molecular weights and boiling points. The monosaccharide product showed two peaks (*m/z* = 282, 383) in the mass spectrum and were identified as both the mono- and diadduct of 3-APBA.

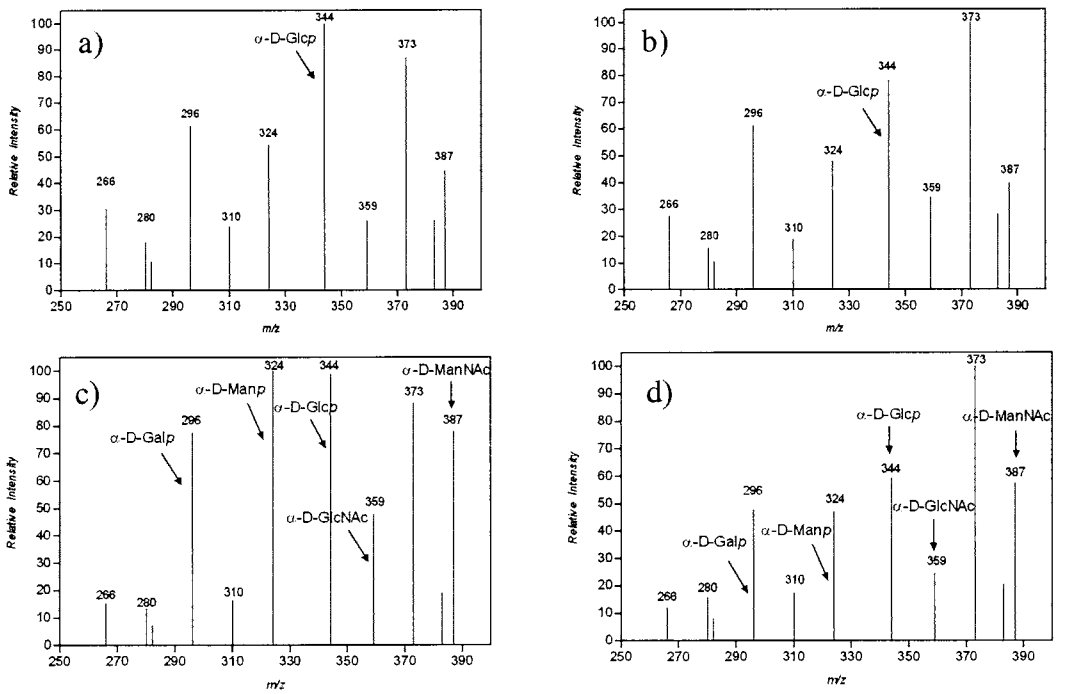


Figure 2. ESI-MS spectrum in selected ion-monitoring positive-ion mode of α-linked substrates after 6 h at 37 °C incubation with cell extract of *E. coli* K12 or *Bacillus cereus*: (a) heat-killed control of *E. coli* K12; (b) with cell extract of *E. coli* K12. (c) heat-killing control of *Bacillus cereus*; (b) with cell extract of *Bacillus cereus*. The cell extract of *E. coli* K12 only cleaved α-D-glucose substrates; the cell extract of *Bacillus cereus* selectively cleaved α-D-glucose, α-D-galactose, α-D-mannose, α-D-glucosamine, and α-D-mannosamine substrates in different intensities.

There are also *m/z* = 282, 383 peaks in the mass spectrum of heat-killed control group, because the cell extract contains a large number of diverse molecules. The signals of some other chemicals overlapped with the peaks of monosaccharide product/3-APBA complex. We still can observe the intensity increasing of the product peaks in the enzymatic reaction groups.

The enzyme activity profiling monitored by ESI-MS method can not only identify which substrates are cleaved but also show the change of the signal intensities of each of the cleaved substrates. Figure 3 shows the percentage decreasing of the signal intensity of cleaved substrates. The substrates which were not cleaved by the enzymes served as the internal references.

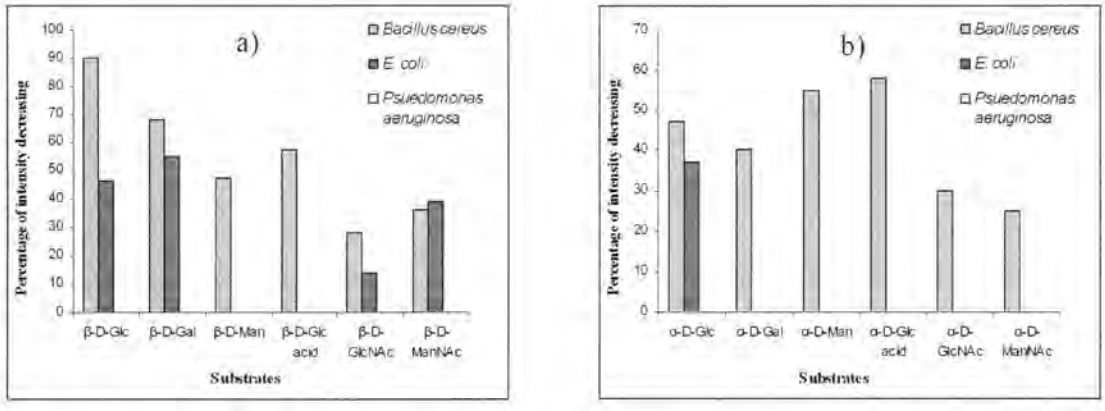


Figure 3. Enzyme activity profiles of *Bacillus cereus* and *E. coli* K12. a) The percentage decrease of the peak intensities of cleaved β-linkage substrates. B) The percentage decrease of the peak intensities of cleaved α-linkage substrates. There is no substrate cleavage in both the β- and α-library for *Psuedomonas aeruginosa* cell extract. The unchanged substrates served as the standard.

Clear identification among these three bacteria already has been obtained only by their glycosidases activity. The change of the substrate signal intensity could potentially be used to differentiate the subspecies of the different bacteria which hydrolyze the same substrates. The difference of intensity changes of the same substrate makes it possible to use less

substrate to identify and differentiate bacteria. This method improved the accuracy and sensitivity of the qualitative positive/negative enzyme activity profile test.

We also used traditional method (*p*-nitrophenyl substrates) to compare with the ESI-MS data and reliability of our method and obtained homogenous results. Table 1 and 2 compare the results between traditional *p*-nitrophenyl substrates enzyme activity profile test and the new ESI-MS monitoring enzyme activity profile method.

Table 1. The results of *p*-nitrophenyl substrates enzyme activity profile test and ESI-MS monitoring enzyme activity profile method of β-library.

β-substrates		Glc	Gal	Man	Xyl	L-Fuc	Glc acid	Glc NAc	Gal NAc	Man NAc
E. coli	ESI	+	+	-	-	-	-	+	-	+
	UV	+	+	-	-	-	N/A	+	-	N/A
Bacillus	ESI	+	+	+	-	-	+	+	-	+
	UV	+	+	-	-	-	N/A	+	-	N/A
Psued	ESI	-	-	-	-	-	-	-	-	-
monas	UV	-	-	-	-	-	N/A	+	-	N/A

Table 2. The results of *p*-nitrophenyl substrates enzyme activity profile test and ESI-MS monitoring enzyme activity profile method of α-library.

α-substrates		Glc	Gal	Man	Xyl	L-Fuc	Glc acid	Glc NAc	Gal NAc	Man NAc
E. coli	ESI	+	-	-	-	-	-	-	-	-
	UV	+	+	-	-	-	N/A	-	-	N/A
Bacillus	ESI	+	+	+	-	-	+	-	-	+
	UV	+	+	-	-	-	N/A	-	-	N/A
Psued	ESI	-	-	-	-	-	-	-	-	-
monas	UV	-	-	-	-	-	N/A	-	-	N/A

Most of the results match each other in this comparison, but a few of substrates are different. The main reason we regard is that the interaction between the high hydrophobic *p*-nitrophenyl group and some glycosidases, such as β-glucosaminidase in *Psuedomonas*

aeruginosa, is different with that of the small alkyl groups. The linkage between *p*-nitrophenyl group and the monosaccharides is more labile compared with the corresponding linkage of the small alkyl groups with the monosaccharides. The condition and substrates in our method are closer to the "real" situation of substrate-enzyme interaction, because DMF were not used and a strong glycosidic bond was introduced. Furthermore, the multiple substrates in the same reaction vial and slight pH variation between our library method and *p*-nitrophenyl test also affect the activities of different glycosidases.

III. Conclusion

ESI-MS combined with our enzyme activity profiling chemical proteomics method provides a way to monitor enzyme activity in bacteria directly instead of by chromogenic or fluorogenic methods, thereby improving the efficiency and sensitivity of the traditional enzyme activity test for the identification of bacteria. This technique does not require enzyme purification and immobilization by using the cell extract directly to detect the intracellular glycosidase activities and successfully differentiate bacteria. Moreover, this method suggests a more efficient and accurate assay by applying it to detect multiple enzyme activities together. It is easier to avoid the mass redundancy of the substrates for various enzymes, but the substrates of different enzymes should not inhibit each other. Expansion of the substrate pool can potentially allow differentiation of even greater numbers of bacteria.

IV. Experimental Section

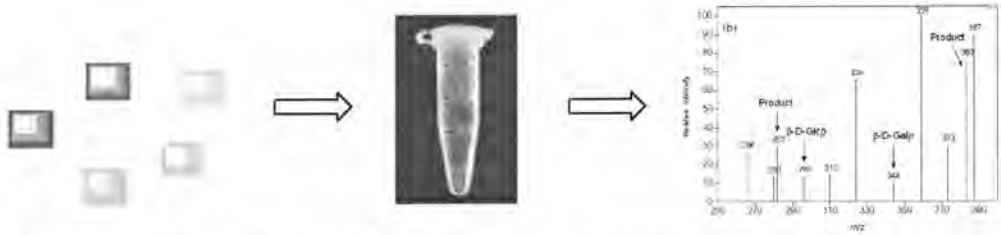
Chemicals. All of the substrates in α- and β-library were designed and synthesized by our lab (the synthesis procedures were described in chapter 1). The *p*-nitrophenyl substrates and other compounds were purchased from Sigma-Aldrich (Milwaukee, WI).

Bacterial Sample Preparation. All of the bacteria, *E. coli* K12, *Bacillus cereus* (ATCC14579), and *Psuedomonas aeruginosa* (ATCC47085) were grown on LB agar plate at 37 ^{o}C for 18 h and transferred from the plate into 4 mL LB Broth in two test tubes by a

sterilized platinum loop and incubated at 37 °C in a shaker for 12 h. The cultured solution was transferred into 200 mL LB broth in 1 L Erlenmeyer flasks and cultured for another 18 h at 37 °C in a shaker. The bacteria were collected by centrifuging at 1, 0000 X g for 10 min. The precipitate was stored at -80 °C at least 30 min before analysis. The cell pellet (from 50 mL cell culture OD_{600} = 1.2) was thawed on ice-bath and 500 μL Tris buffer (50 mM, pH = 7.5) and 2 μL PMSF (1 mg/mL) in ethanol were added. The cells were disrupted by sonication. After the unbroken cells and debris were removed by centrifugation at 10,000 X g at 4 °C for 30 min, the supernatant was stored at 0 °C as the cell extract for enzymatic reactions. 250 μL cell extract solution was transferred to another 1 mL microcentrifuge tube and heated at 95 °C for 5 min. The suspension was centrifuged under 10,000 X g for 10 min and the supernatant was stored as the heat-killed sample for the control reactions.

ESI-MS Detection. A Shimadzu LCMS 2010 quadrupole mass spectrometer (Shimadzu Scientific Instruments, Columbia, MD) equipped with an electrospray ionization (ESI) source was used in positive ion mode. Other details are described in chapter 1.

Enzymatic activity monitored by ESI-MS. The enzymatic reaction: 7 μL of each α or β substrate (50 mM in nanopure water), 10 μL phosphate buffer (500 mM, pH = 7.5), 30 μL of cell extract (10 μL DMF and 20 μL of cell extract were used for the groups which compared with *p*-nitrophenyl substrates tests to obtain the exactly same reaction condition.) were loaded in a 1 mL microcentrifuge tube. The reaction mixture was incubated at 37 °C and monitored by ESI-MS at 1 h, 3 h, 6 h, and 12 h. The same number and quantity of reagents were loaded in the control reactions, but heat-killed sample was added instead of cell extract. To quench the reaction mixtures, a 15 μL aliquot of the reaction mixture was added to 1% AcOH/acetonitrile (15 μL). After addition of 3-aminophenylboronic acid (15 μL, 50 mM stock solution in water), 2 μL aqueous ammonium sulfate (2.5 M) was added to the solution (the reason and details are explained in chapter 1). The samples were centrifuged at 10,000 x g for 10 min to precipitate the protein and other impurity before direct injection (5 μL) of the supernatant in the mass spectrometer as described above and obtained the enzyme activity profiling. The α- and β-linkage substrates and their ion (*m/z*) generated in the ESI-MS spectrum were listed in chapter 1.



1) Synthesis of carbohydrate substrates with different mass

2) Culture different bacteria

3) Simultaneously incubate all of the substrates with the different bacteria extract

4) Monitoring the cell extract reaction with ESI-MS and obtain the enzyme activity profile for the bacteria

Figure 4. The procedure of the enzyme activity profiling measurement monitored by ESI-MS

Tests with *p*-nitrophenyl derivatives. The enzymatic reaction conditions: 7 μL of each *p*-nitrophenyl substrates (50mM in 15% DMF/H_2O solution), 73 μL of phosphate buffer (50 mM pH = 7.5), 20 μL cell extract were loaded in a 750 μL microcentrifuge tube and the reaction mixture was incubated at 37 ^{o}C for 6 h and monitored them by UV scanning plate under 400 nm. The control group was prepared and analyzed by a same manner, but heat-killed sample (95 ^{o}C, 5 min) was added instead of cell extract.

References

Adam, G. C.; Sorensen, E. J.; Cravatt, B. F.; Proteomic profiling of mechanistically distinct enzyme classes using a common chemotype. *Nat. Biotechnol.* **2002**, *20*, 805-809.

Al-Khaldi, SF.; Mossoba, M. M.; Ismail, A.A.; Fry, F. S. Accelerating bacterial identification by infrared spectroscopy by employing microarray deposition of microorganisms. *Foodborne Pathogens and Disease* **2004**, *1*, 172-177.

Ashwell, G. and Harford, J. Carbohydrate-specific receptors of the liver. *Annu. Rev. Biochem.* **1982**, *51*, 531.

Baranski, T. J.; Koelsch, G.; Hartsuck, J. A.; Kornfeld S. Mapping and molecular modeling of a recognition domain for lysosomal enzyme targeting. *J. Biol. Chem.* **1991,** *266*, 23365.

Bascomb, S.; Grantham, C. A. Application of automated assay of asparaginase and other ammonia-releasing enzymes to the identification of bacteria. *Soc. Appl. Bacteriol. Tech. Ser.* **1975**, *8* (Some Methods Microbiol. Assay), 29-54.

Basile, F.; Ferrer, I.; Furlong, E. T.; Voorhees, K. J. Simultaneous multiple substrate tag detection with ESI-Ion trap MS for in vivo bacterial enzyme activity profiling. *Anal. Chem.* **2002**, *74*, 4290-4293

Basile, F.; Hughes, K. D.; Wisniowski, P. E.; Gorenstein, D. G.; Lytle, F. E.; McCay-Buis, T. S.; Huber, D. M.; Hemming, B. C. Fast and Sensitive Laser-Based Enzymatic Detection of the Lactose Operon in Microorganisms, *Anal. Biochem.* **1993**, *211*, 55-60.

Baum, L. G. Developing a taste for sweets. *Immunity*, **2002,** *16*, 5-8.

Behr, M. A.; Wilson, M. A.; Gill, W. P. *et al.* Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **1999**, *284*, 1520.

Bevilacqua, M.; Butcher E.; Furie, B.; Gllstin M. *et al.* Selectins: a family of adhesion receptors. *Cell* **1991**, *67*, 233.

Black, G. E; Fox, A; Fox, K; Snyder, A. P.; Smith, P. B. Electrospray tandem mass spectrometry for analysis of native muramic acid in whole bacterial cell hydrolysates. *Anal. Chem.* **1994**, *66*, 4171.

Bothner, B.; Dong, X. F. A.; Bibbs, L. D.; Johnson, J. E.; Siuzdak, G. Evidence of viral capsid dynamics using limited proteolysis and mass spectrometry. *J. Biol. Chem.* **1998**, *273*, 673-676.

Bothner, B.; Chavez, R.; Wei, J.; Siuzdak, G.; *et al.* Monitoring enzyme catalysis with mass spectrometry. *J. Biol. Chem.* **2000**, *275*, 13455-13459.

Call, D.R.; Brockman, F. J.; Chandler, D. P. Detecting and genotyping *Escherichia coli* O157: H7 using multiplexed PCR and nucleic acid microarrays. *Int. J. Food Microbiol.* **2001**, *67*, 71.

Chait, B. T.; Wang, R.; Beavis, R. C.; Kent, S. B. Protein ladder sequencing. *Science* **1993**, *262*, 89-92.

Chapman, E.; Best, M. D.; Hanson, S. R.; Wong, Chi-Huey. Sulfotransferases: Structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed.* **2004**, *43*, 3526-3548.

Cole, M. J.; Enke, C. G. Direct determination of phospholipid structures in microorganisms by fast atom bombardment triple quadrupole mass spectrometry. *Anal. Chem.* **1991**, *63*, 1032.

Coutinho, P. M.; Stam, M.; Blanc, E.; Henrissat, B. Why are there so many carbohydrate-active enzyme-related genes in plants? *Trends Plant Sci.* **2003**, *8*, 563-565.

Cumming, D. A. Glycosylation of recombinant protein therapeutics: control and functional implication. *Glycobiology* **1991**, *1*, 115-130.

Dalbey, R. E.; Chen, M. Y. Sec-translocase mediated membrane protein biogenesis. *Biochimica et Biophysica Acta* **2004**, *1694*, 37-53.

Daniels, M. A.; Devine, L.; Miller, J. D. *et al.* CD8 binding to MHC class I molecules is influenced by T cell maturation and glycosylation. *Immunity* **2001**, *15*, 1051-1061.

Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. Glycomethanethiosulfonates: powerful reagents for protein glycosylation. *Tetrahedron: Asymmetry* **2000**, *11*, 245-262.

Davies, G.; Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **1995**, *3*, 853-859.

Davis, N. J.; Flitsch, S. L. Selective oxidation of monosaccharide derivatives to uronic acids. *Tetrahedron Lett.* **1993**, *34,* 1181-1184.

Delorme, E.; Lorenzini, T.; Giffin, J.; Martin, F.; Elliott, S.; *et al.* Role of Glycosylation on the secretion and biological activity of erythropoietin. *Biochemistry* **1992**, *31*, 9871-9876.

Disney, M. D; Zheng, J.; Swager, T. M.; Seeberger, P. H. Detection of bacteria with carbohydrate-functionalized fluorescent polymers. *J. Am. Chem. Soc.* **2004**, *126*, 13343-6.

Domin, M. A.; Welham, K. J.; Ashton, D. S. The effect of solvent and matrix combinations on the analysis of bacteria by MALDI-TOF MS. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 222–226.

Drickamer K. Two distinct classes of carbohydrate-recognition domains in animal lectins. *J. Biol. Chem.* **1988**, *263*, 9557.

Dunn, A. D.; Coris, C. M.; Myers, H. E.; Dunn, J. T. Tyrosine 130 is an important outer ring donor for thyroxine formation in thyroglobulin. *J. Biol. Chem.* **1998**, *273*, 25223-25229.

Dunn, A. D.; Cruthfield, H. G.; Dunn, J. T. Thyroglobulin processing by thyroidal proteases: Major sites of cleavage by cathepsins B, D, and L. *J. Biol. Chem.* **1991**, *266*, 20198-20204.

Egrie, J. C.; Strickland, T. W.; Lane, J.; Aoki K.; Cohen, A. M.; Browne, J. K. and Hines, D. K. *et al.* Characterization and biological effect of recombinant human erythropoietin. *Immunobiology* **1986**, *172,* 213-224.

Egrie, J. C.; Dwyer, E.; Lykos, M.; Hitz, A.; Brown, J. K. Novel erythropoiesis stimulating protein (NESP) has a longer serum half-life and greater in vivo biological activity than recombinant human erythropoietin (rHuEPO). *Blood* **1997**, *90*, 56a.

Estrada, A. et al. Immunomodulatory activities of oat β-glucan *in vitro* and *in vivo*. *Microbiol. Immunol.* **41**, 991-998.

Fang, J.; Barcelona, M. J. Structural determination and quantitative analysis of bacterial phospholipids using liquid chromatography electrospray ionization mass spectrometry. *J. Microb. Methods.* **1998**, *33*, 23–35.

Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I. Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC SIGNR. *Science* **2001**, *294*, 2163-2166.

Feizi, T. Demonstration by monoclonal antibodies that carbohydrate structures of glycoprotein and glycolipids are onco-developmental antigens. *Nature* **1985**, *314*, 53.

Fukuda, M. Cell surface glycoconjugates as onco-differentiation makers in hematopoietic cells. *Biochim. Biophys. Acta.* **1985**, *780*,119.

Galemmo, R. A.; Horton, D. Preparative routes to methyl 2-acetamido-2,6-dideoxy-α-D-glucopyranoside. *Carbohydr. Res.* **1983**, *119*, 231-240.

Gebler, J.; Gilkes, N. R.; Claeyssens, M.; Whiter, S.G.; Stereo selective hydrolysis catalyzed by related β-1,4-glucanases and β-1,4-xylanases. *J. Biol. Chem.* **1992**, *267*, 12259-12262.

Godsey, J. H.; Matteo, M. R.; Shen, D.; Tolman, G.; Gohlke, J. R. Rapid identification of *Enterobacteriaceae* with microbial enzyme activity profiles. *J. Clin. Microbiol.* **1981**, *13*, 483-90.

Goodacre, R.; Heald, J. K.; Kell, D. B. Characterization of intact microorganisms using electrospray ionization mass spectrometry. *FEMS Microbiol. Lett.* **1999**, *176,* 17-24.

Gorin, P. A. J.; Mazurek, M. Further studies on the assignment of signals in 13C magnetic resonance spectra of aldoses and derived methyl glycosides. *Can. J. Chem.* **1975**, *53*, 1212-1217.

Grönberg, G.; Nilsson, U.; Bock, K.; Magnusson, G. Nuclear magnetic resonance and conformational investigations of the pentasaccharide of the Forssman antigen and overlapping di-, tri-, and tetra-saccharide sequences. *Carbohydr. Res.* **1994**, *257*, 35-54.

Harrington, C. A.; Rosenow, C.; Retief, J. Monitoring gene expression using DNA microarrays. *Curr. Opin. Microbiol.* **2000**, *3*, 285.

Hart, G. W.; Haltiwanger, R. S.; Holt, G. D.; Kelly W. G. Glycosylation in the nucleus and cytoplasm. *Annu. Rev. Biochem.* **1989**, *58*, 841.

Helenius, A.; Aebi, M. Intracellular functions of N-linked glycans. *Science* **2001**, *291*, 2364-2369.

Henrissat, B.; Bairoch, A. Update the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **1996**, *316*, 695-696.

Hirichs, W.; Saenger, W. Crystal and molecular structure of the hexasaccharide complex (p-nitrophenyl α-maltohexaoside). *J. Am. Chem. Soc.* **1990**, *112*, 2789-2796.

Hood, D.W.; Deadman, M. E.; Jennings, M. P. *et al.*, DNA repeats identify novel virulence genes in *Haemophilus influenzae*. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11121.

Houpikian P.; Raoult, D. Traditional and molecular techniques for the study of emerging bacterial diseases: one laboratory's perspective. *Emerg. Infect. Disea.* **2002**, *8*, 122-131.

Hsieh, F. Y. L.; Tong, X.; Wachs, T.; Ganem, B.; Henion, J. Kinetic monitoring of enzymic reactions in real time by quantitative high-performance liquid chromatography-mass spectrometry. *Anal. Biochem.* **1995**, *229*, 20-25.

Hunt, D. F.; Michel, H.; Dickinson, T. A.; Sette, A. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-Ad. *Science* **1992**, *256*, 1817-1820.

Imberty, A.; Chanzy, H.; Perez, S.; Buleon A.; Tran, V. The double-helical nature of the crystalline part of A-starch. *J. Mol. Biol.* **1988**, *201*, 365-378.

Izumi, K. N.M.R. spectra (^{1}H- and ^{13}C-) of methyl-2-acetamido-2-deoxy-D-hexopyranosides in the presence of lanthanide ions. *Carbohydr. Res.* **1987**, *170,* 19-26.

Jessani, N.; Liu, Y.; Humphrey, M.; Cravatt, B. F. Enzyme activity profiles of the secreted and membrane proteome that depict cancer cell invasiveness. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10335-10340.

Kerkhoff, F. T.; Bergmans, A. M.; Vander, Z. A.; Rothova, A. Demonstration of Bartonella grahamii DNA in ocular fluids of a patient with neuroretinitis. *J. Clin. Microbiol.* **1999**, *37*, 4034-4038.

Kikuchi, H.; Saito, Y.; Komiya, J.; Takaya, Y.; Honma, S.; Nakahata, N.; Ito, A.; Oshima, Y. Furanodictine A and B: Amino sugar analogues produced by cellular slime mold d*ictyostelium discoideum* showing neuronal differentiation activity. *J. Org. Chem.* **2001**, *66*, 6982-6987.

Kilian, M. Rapid identification of *Actinomycetaceae* and related bacteria. *J. Clin. Microbiol.* **1978**, *8*, 127-133.

Kilian, M.; Bulow, P. Rapid diagnosis of *Enterobacteriaceae*. I. Detection of bacterial glycosidases. *Acta. Pathol. Microbiol. Scandi. Sec B: Microbiology* **1976**, *84*, 245-51.

Knapp, S.; Gore, V. K. Synthesis of the shimofuridin nucleoside disaccharide. *J. Org. Chem.* **1996**, *61*, 6744-6747.

Kornfeld, S. Trafficking of lysosomal enzymes. *FASEB J.* **1987**, *1*, 462.

Kornfeld, R.; Korfeld S. Assembly of asparagines-linked oligosaccharide. *Annu. Rev. Biochem.* **1985**, *54*, 631.

Krantz, S. B. Erythropoietin. *Blood* **1991**, *77*, 419-434.

Krebs, J. F.; Siuzdak, G. H.; Dyson, J.; Stewart, J. D.; Benkovic, S. J. *Biochem.* **1995**, *34*, 720.

Krishnamurthy, T.; Rajamani, U.; Ross, P.L.; Jabhour, R.; Nair, H.; Eng, J.; Lee, T. D.; *et al.* Mass spectral investigations on microorganisms. *J. Toxicol. Toxin Rev.* **2000**, 19, 95–117.

Krishnamurthy, T.; Davis, M.; Stahl, D.; Lee, D. Liquid chromatography/microspray mass spectrometry for bacterial investigations. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 39-49.

Kvasnovsky, J.; Kozak, J.; Lokaj, V.; Svec, J.; Magdolen, P. Slovak Patents, SK 279045 B6 19980603, 1998.

Lander, E. S.; Linton, L. M.; Birren, B.; Fitzhugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860-921.

Larsen, R. D.; Ernst, L. K.; Nair, R. P.; Lowe, J. B. Molecular cloning, sequence, and expression of a human GDP-L-fucose: β-D-galactoside 2-α-L-fucosyltransferase cDNA that can form the H blood group antigen. *Pro. Natl. Acad. Sci. USA* **1990**, *87*, 6674.

Lawson, S. L. *et al.* Mechanistic consequences of replacing the active-site nucleophile Glu-358 in *Agrobacterium* sp. β-glucosidase with a cysteine residue. *Biochem. J.* **1998**, *330*, 203-209.

Lee, E. D.; Muck, W.; Henion, J. D.; Covey, T. R. Real-time reaction monitoring by continuous-introduction ion-spray tandem mass spectrometry. *J. Am. Chem. Soc.* **1989**, *111*, 4600-4604.

Lee, K. L; Rega, M. E.; Watson, R. R.; Campbell, C. C. Identification of yeast phase of pathogenic fungi by the specificity of their aminopeptidase(s). *Sabouraudia* **1975**, *13*, 132-141.

Lee, R. T.; Lee, Y. C. Synthesis of 3-(2-aminoethylthio) propyl glycosides. *Carbohydr. Res.* **1974**, *37*, 193-201.

Liu, C. L.; Hofstadler, S. A.; Bresson, J. A.; Udseth, H., R.; Tsukuda, T.; Smith, R. D.; Snyder, A. P. On-line dual microdialysis with ESI-MS for direct analysis of complex biological samples and microorganism lysates. *Anal. Chem.* **1998**, *70*, 1797–1801.

Love, K. R.; Seeberger, P. H. Automated solid-phase synthesis of protected tumor-associated antigen and blood group determinant oligosaccharides. *Angew. Chem. Int. Ed.* **2004**, *43*, 602-605.

Ly, H.; Withers, S. G.; Mutagenesis of glycosidases. *Annu. Rev. Biochem.* **1999**, 487-522.

Mackenzie, L. F. *et al.* Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* **1998**, *102,* 5583-5584.

Maddocks, J. L.; Greenan, M. J. Rapid method for identifying bacterial enzymes. *J. Clinic. Pathol.* **1975**, *28*, 686-7.

Maeda, Y. Y. and Chihara, G. Letinan, a new immunoacceleratorof cell-mediated responses. *Nature* **1971**, *229*, 634.

Matayoshi, E. D.; Wang, G. T.; Kraft, G. A.; Erickson, J. Novel fluorogenic substrates for assaying retroviral proteases by resonance energy transfer. *Science* **1990**, *247*, 954-958.

McCarter, J.; Wither, S. G. Mechanism of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885-892.

McGeary, R. P.; Wright, K.; Toth, I. Conversion of glucosamine to galactosamine and allosamine derivatives: control of inversions of stereochemistry at C-3 and C-4. *J. Org. Chem.* **2001**, *66*, 5102-5105.

Melvin, F.; McNeill, A.; Henderson, P. J. F.; Herbert, R. B. The improved synthesis of β-D-glucuronides using TEMPO and t-butyl hypochlorite. *Tetrahedron Lett.* **1999**, *40*, 1201-1202.

Moody, A. M.; Chui, D.; Reche, P. A.; Priatel, J. J.; Marth, J. D.; Reinherz, E. L. Developmentally regulated glycosylation of the CD 8αβ coreceptor stalk modulates ligand binding. *Cell* **2001**, *107*, 501-512.

Myers, S.; Baker, A. Drug discovery-an operating model for a new era. *Naure Biotechnol.* **2001**, *19*, 727-730.

Narhi, L. O.; Arakawa T.; Aoki, K. H.; Strickland T. W.; et al. The effect of carbohydrate on the structure and stability of erythropoietin. *J. Biol. Chem.* **1991**, *266*, 23022-23026.

Norin, M.; Sundstrom, M.; Structural proteomics: developments in structure-to-function predictions. *Trends Biotech.* **2001**, *20*, 79-84.

Okabe, S.; Ito, T.; Satoh, H. Sulfate-reducing bacterial community structure and their contribution to carbon mineralization in a wastewater biofilm growing under microaerophilic conditions. *Appl. Microbio. Biotech.* **2003**, *63*, 322-334.

Pandey, A.; Mann, M.; Proteomics to study genes and genomes. *Nature* **2000**, *405*, 837-846.

Perugino, G.; Trincone A.; Rossi, M.; Moracci, M.; Oligosaccharide synthesis by glycosynthases. *Trend in Biotech.* **2004**, 22, 31-37

Pouchert, C. J.; Behnke, J. *The Aldrich Library of 13C and 1H FT NMR Spectra.* Aldrich Chemical Company, Inc. **1993**, *1*, 301-304.

Que, A. H.; Novotny, M. V. Structural characterization of neutral oligosaccharide mixtures through a combination of capillary electrochromatography and ion trap tandem mass spectrometry *Anal. Bioanal. Chem.* **2003**, *375*, 599-608.

Relman, D. A. The search for unrecognized pathogens. *Science* **1999**, *284*, 1308-1310.

Reiss, T. Drug discovery of the future: the implications of the human genome project. *Trends Biotechnol.* **2001**, *19*, 496-499.

Robyt, J. F. *Essentials of Carbohydrate Chemistry*. Springer-Verlag New York, Inc. **1998**, Chapter *6*, Page 157.

Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. Glycosylation and the immune system. *Science* **2001**, *291*, 2370-2376.

Sawada, R.; Tsuboi, S.; Fukuda, M. Differential E-selectin-dependent adhesion efficiency in sublines of a human clone caner with distinct metastatic potentials. *J. Biol. Chem.* **1994**, *269*, 1425.

Seeberger, P. H.; Soucy, R. L.; Kwon, Y.U.; Snyder, D. A.; Kanemitsu, T. A convergent, versatile route to two synthetic conjugate anti-toxin malaria vaccines. *Chem. Comm.* **2004**, *15*, 1706-1707.

Seto, N. O. L.; Palcic, M. M.; Compston, C. A.; Li, H.; Bundle, D. R.; Narang, S. A. Sequential interchange of four amino acids from blood group B to blood group A

glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes *J. Biol. Chem.* **1997**, *272*, 14133-14138.

Sheeley, D. M.; Reinhold V. N. Structural characterization of carbohydrate sequence, linkage, and branching in a quadrupole ion trap mass spectrometer: Neutral oligosaccharides and N-linked glycans. *Anal. Chem.* **1998**, *70*, 3053-3059

Siuzdak, G. Probing viruses with mass spectrometry. *J. Mass Spectrom.* **1998**, *33*, 203-211.

Sinnott, M. L.; Catalytic mechanisms of enzymatic glycosyl transfer. *Chem. Rev.* **1990**, *90*, 1171-1202.

Smith, J. L.; Fratamico, P. M.; Novak, J. S. Quorum sensing: A primer for food microbiologists. *J. Food. Protect.* **2004**, *67*, 1053-1070.

Strauss, E. J. and Falkow, S. Microbial pathogenesis: genomics and beyond. *Science* **1997**, *276*, 707.

ThØgersen, H.; Lemieux, R. U. Further justification for the exo-anomeric effect. Conformational analysis based on nuclear magnetic resonance spectroscopy of oligosaccharides. *Can. J. Chem.* **1982**, *60,* 44-50.

Todar, K. *Todar's Online Textbook of Bacteriology*, Department of Bacteriology, University of Wisconsin-Madison, **2002**, http://www.textbookofbacteriology.net/ (November 30th, 2004).

Tomer, K. B. Separations combined with mass spectrometry. *Chem. Rev.* **2001**, *101*, 297-328.

Utille, J. P.; Priem, P. Synthesis of allyl 2-*O*-(α-L-arabinofuranosyl)-6-*O*-(α-D-mannopyranosyl)-β-D-mannopyranoside, a unique plant *N*-glycan motif containing arabinose. *Carbohydr. Res.* **2000**, *329*, 431-439.

Vaidyanathan, S.; Kell, D. B.; Goodacre, R. Flow-injection electrospray ionization mass spectrometry of crude cell extracts for high-throughput bacterial identification. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 118-128.

Vaidyanathan, S.; Rowland, J. J.; Kell, D. B.; Goodacre, R. Discrimination of aerobic endospore-forming bacteria via electrospray-ionization mass spectrometry of whole cell suspensions. *Anal. Chem.* **2001**, *73*, 4134–4144.

Varki, A. Biological Roles of oligosaccharides-all of the theories are correct *glycobio.* **1993**, *3*, 97-130.

Veness, R. G.; Evans, C. S. Identification of monosaccharides and related compounds by gas chromatography-Fourier transform infrared spectroscopy of their trimethylsilyl ethers. *J. Chromatogr. A* **1996**, *721*, 165-172.

Venter, J. C.; Adams, M. D.; Myers, E. W.; Holt R. A.; *et al.* The sequence of the human genome. *Science* **2001**, *291*, 1304-1351.

Vestal, M. L. Method of ion generation *Chem. Rev.* **2001**, 361-375.

Vocadle, D. J. Bertozzi, C. R. A strategy for functional proteomic analysis of glycosidase activity from cell lysates *Angew. Chem. Int. Ed.* **2004**, 43, 5338-5342.

Wang, Q.; Graham, R.W.; Trimbur D.; Warren, R. A.; Withers S. G. Change enzymatic reaction mechanisms by mutagenesis: conversion a retaining glycosidase to an inverting enzyme. *J. Am. Chem. Soc.* **1994**, *116*, 11594-11595.

Wang, Z. P.; Russon, L.; Li, L.; Roser, D. C.; Long, S. R. Investigation of spectral reproducibility in direct analysis of bacteria proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy. *Rapid. Commun. Mass. Spectrom.* **1998**, *12*, 456–464.

Wells, L.; Vosseller, K.; Hart, G. W. Glycosylation of nucleocytoplasmic protein: signal transduction and O-GlcNAc. *Science* **2001**, *291*, 2376-2378.

White, A.; Rose D. R. Mechanism of catalysis by retaining β-glycosyl hydrolases. *Curr. Opin. Struct. Biol.* **1997**, *7*, 645-651.

White, A.; Tull, D.; Johns, K.; Withers, S. G.; Ross, D. R. Crystallographic observation of a covalent intermediate in a β-glycosidase. *Nature. Struct. Biol.* **1996**, *3*, 149-154.

Willams, D.; Lee, T. D.; Dinh, N.; Young, M. K. Oligosaccharide profiling: the facile detection of mono-, di- and oligosaccharides by electrospray orthogonal time-of-flight mass spectrometry using 3-aminophenylboronic acid derivatization. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1530-1537.

Willims, D., Young, M. K. Analysis of neutral isomeric low molecular weight carbohydrates using ferrocenyl boronate derivatization and tandem electrospray mass spectrometry. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 2083-2091.

Withers, S. G. *et al.* Mechanistic consequences of mutation of the active site nucleophile Glu 358 in *Agrobacterium* β-glucosidase. *Biochem.* **1992**, *31*, 9979-9985.

Wolfenden, R.; Lu, X.; Young, G.; Spontaneous hydrolysis of glycosides, *J. Am. Chem. Soc.* **1998**, *120*, 6814-6815.

Wrighton, N. C.; Farrell, F. X.; Chang, R.; Dower W. J. Small peptides as potent mimetics of the protein hormone erythropoietin. *Science* **1996**, *273,* 458.

Wu, J.; Takayama, S.; Wong, C. H.; Siuzdak, G. Quantitative electrospray mass spectrometry for the rapid assay of enzyme inhibitors. *Chem. Biol.* **1997**, *4,* 653-657.

Xu, R.; Hanson, S. R.; Zhang, Z.; Yang, Y.; Schultz, P. G.; Wong, Chi-Huey. Site-specific incorporation of the mucin-type *N*-acetylgalactosamine-α-*O*-threonine into protein in Escherichia coli. *J. Am. Chem. Soc.* **2004**, *126*, 15654-15655.

Yamamoto, F. I.; Clausen, H.; White, T.; Marken, J.; Hakomori, S. I. Molecular genetic basis of the histo-blood group ABO system. *Nature* **1990**, *345*, 229.

Yamamoto, F. I.; Hakomori, S. I. Sugar-nucleotide donor specificity of Histo-blood group A and B transferases is based on amino acid substitutions. *J. Biol. Chem.* **1990**, *265*, 19257.

Yu, Y.; Ko, K. S.; Zea, C. J.; Pohl, N. L.; Discovery of the Chemical Function of Glycosidases: Design, Synthesis, and Evaluation of Mass-Differentiated Carbohydrate Libraries, *Org. Lett.* **2004**, *6,* 2031-2033.

Zea, C. J.; MacDonell, Stephen W.; Pohl, N. L. Discovery of the archaeal chemical link between glycogen (Starch) synthase families using a new mass spectrometry assay. *J. Am. Chem. Soc.* **2003**, *125*, 13666-13667.

Zea, C. J.; Pohl, N. L. Kinetic and substrate binding analysis of phosphorylase b via electrospray ionization mass spectrometry: a model for chemical proteomics of sugar phosphorylases. *Anal. Biochem.* **2004**, 327, 107-113.

Zea, C. J.; Pohl, N. L. General assay for sugar nucleotidyltransferases using electrospray ionization mass spectrometry. *Anal. Biochem.* **2004**, *328*, 196-202.

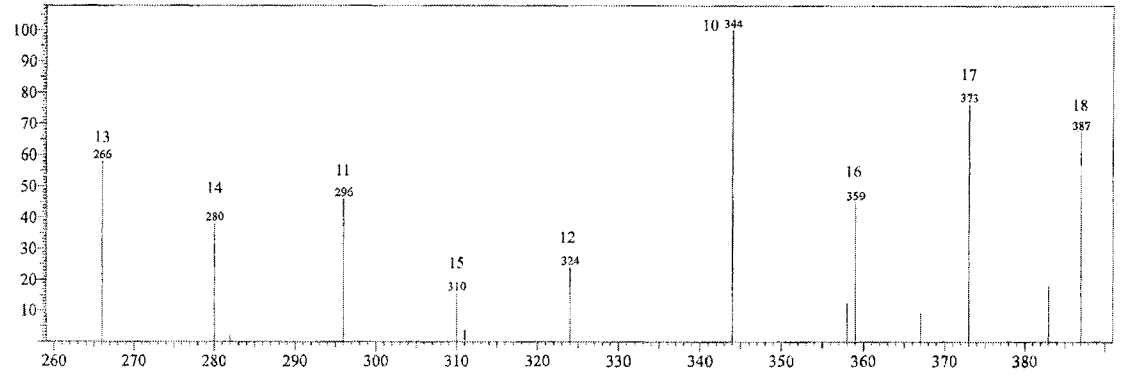
Zehavi, U.; Sharon, N. The synthesis of methyl 2,4-diacetamido-2,4,6-trideoxy hexopyranosides. *J. Org. Chem.* **1972**, *37*, 2141-2145.

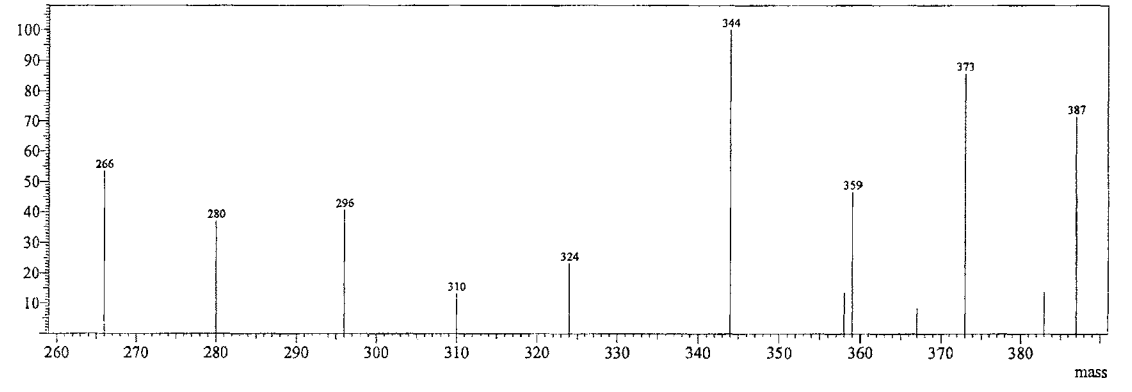
Appendices

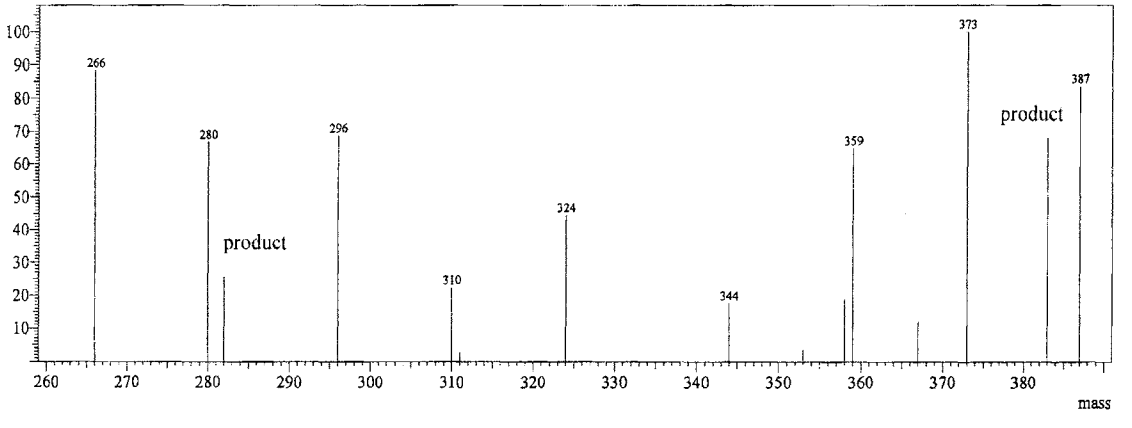
Nuclear Magnetic Resonance Spectra

& MS Spectra of ESI-MS Monitoring Enzymatic Reactions

α-glucosidase (from rice)

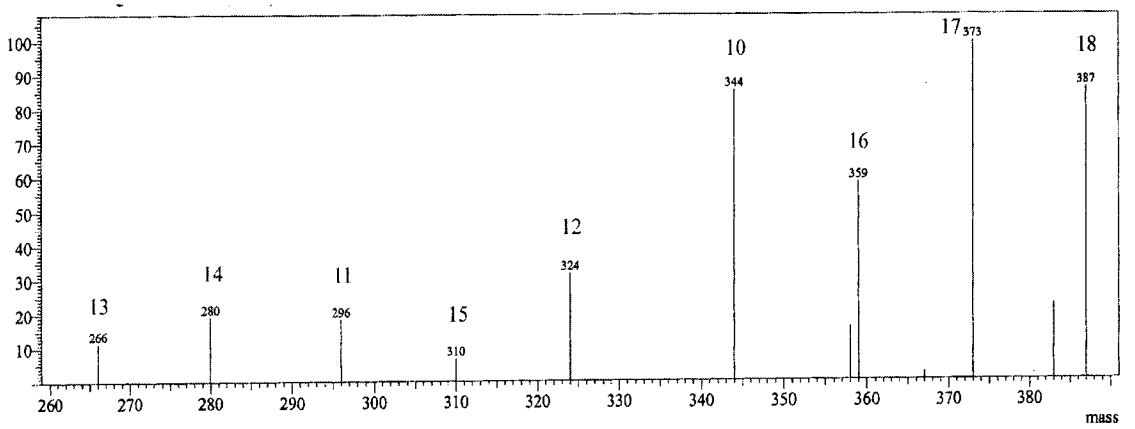


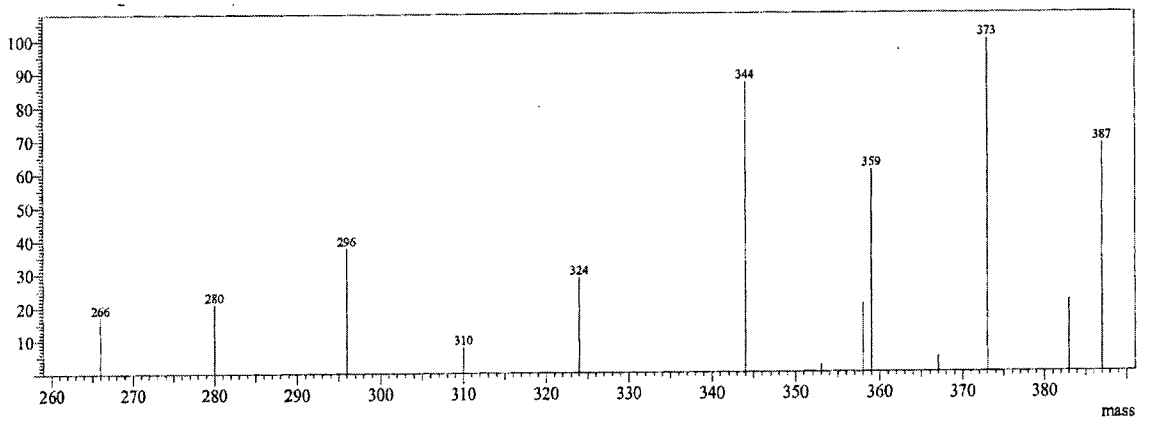


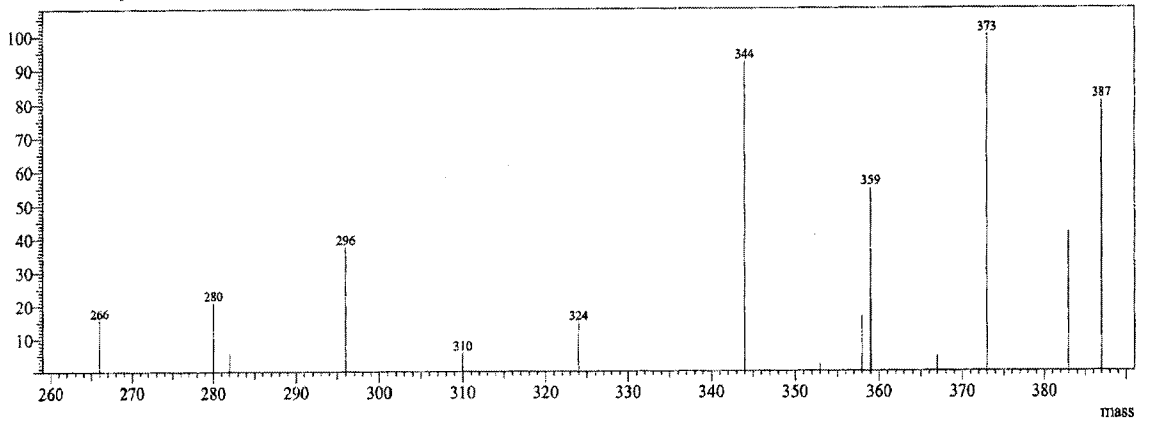


a) Time, 0 hour, b) Time, 10 hours, control, c) Time, 10 hours, reaction

α-mannosidase (from jack bean)

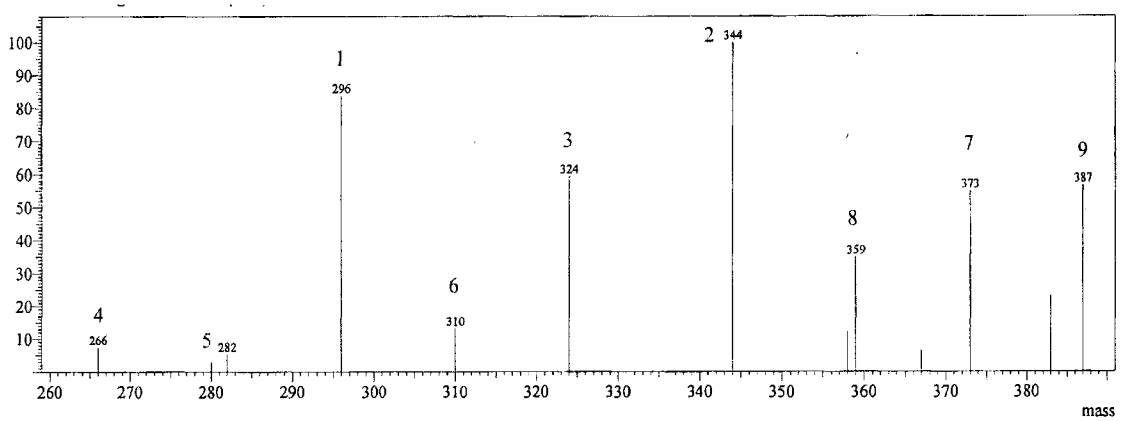


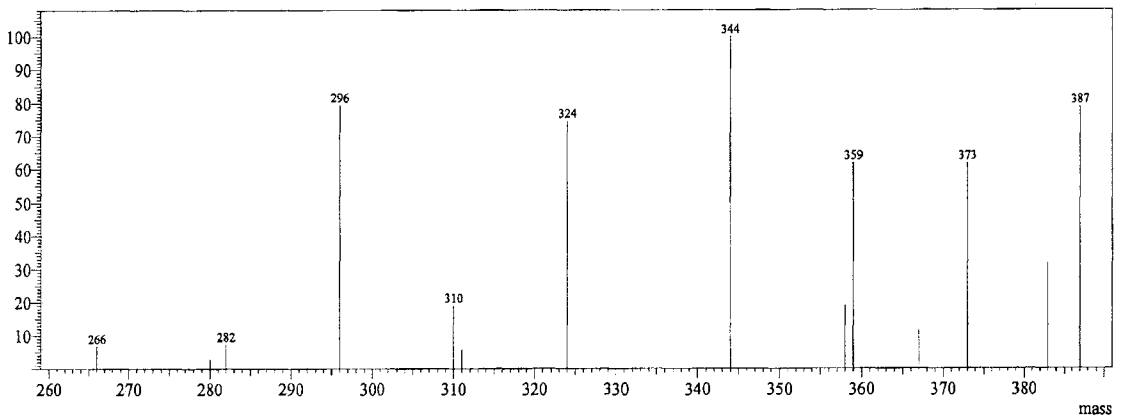


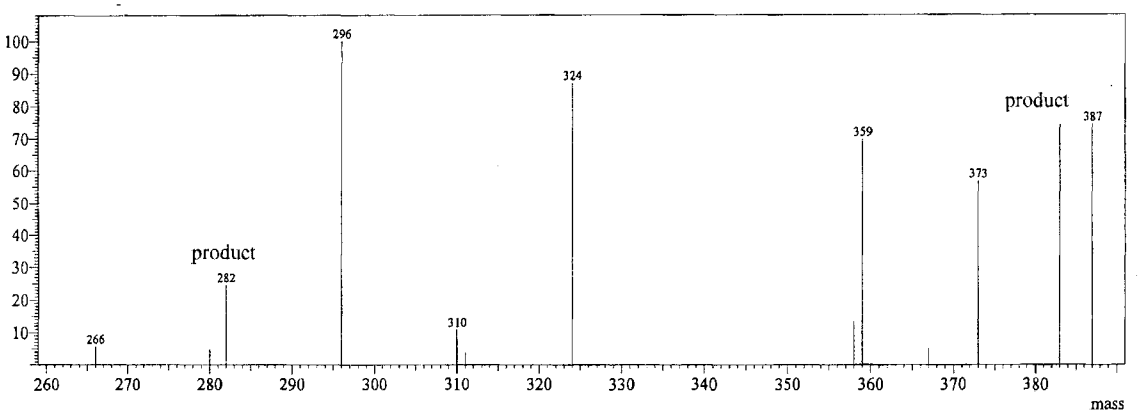


a) Time, 0 hour, b) Time, 10 hours, control, c) Time, 10 hours, reaction

β-galactosidase (from *Aspergillus oryzae*)



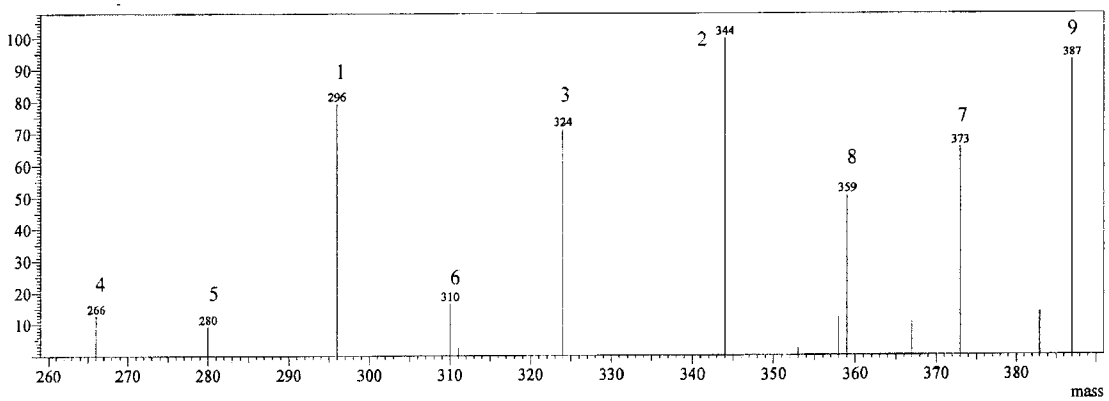




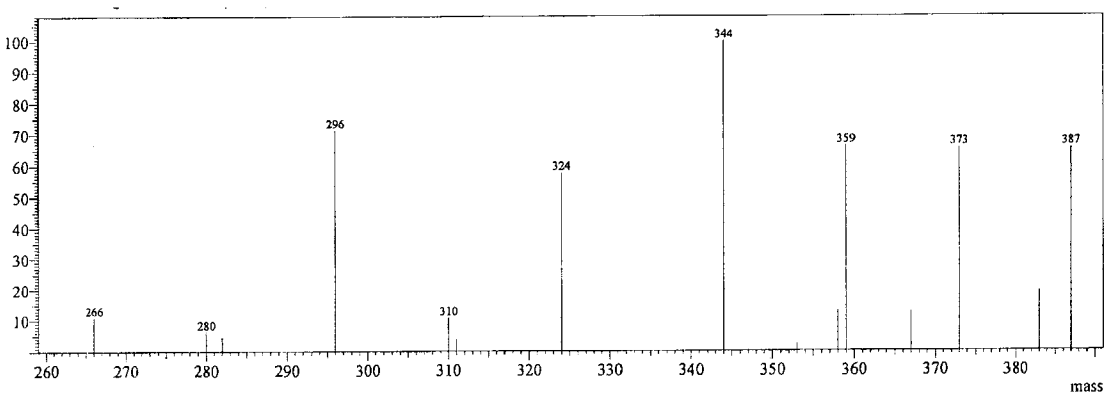
a) Time, 0 hour, b) Time, 1 hour, control, c) Time, 1 hour, reaction

β-glucosidase (crude, from Almond)

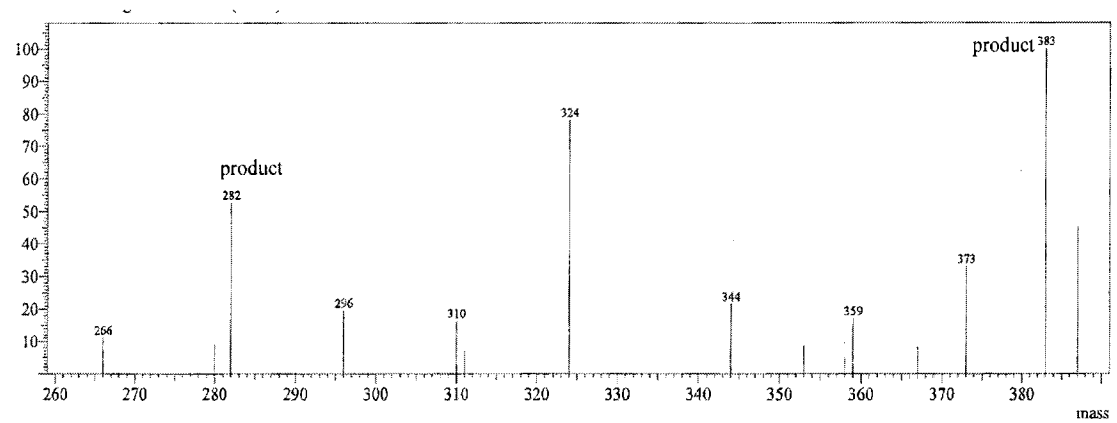
a)



b)

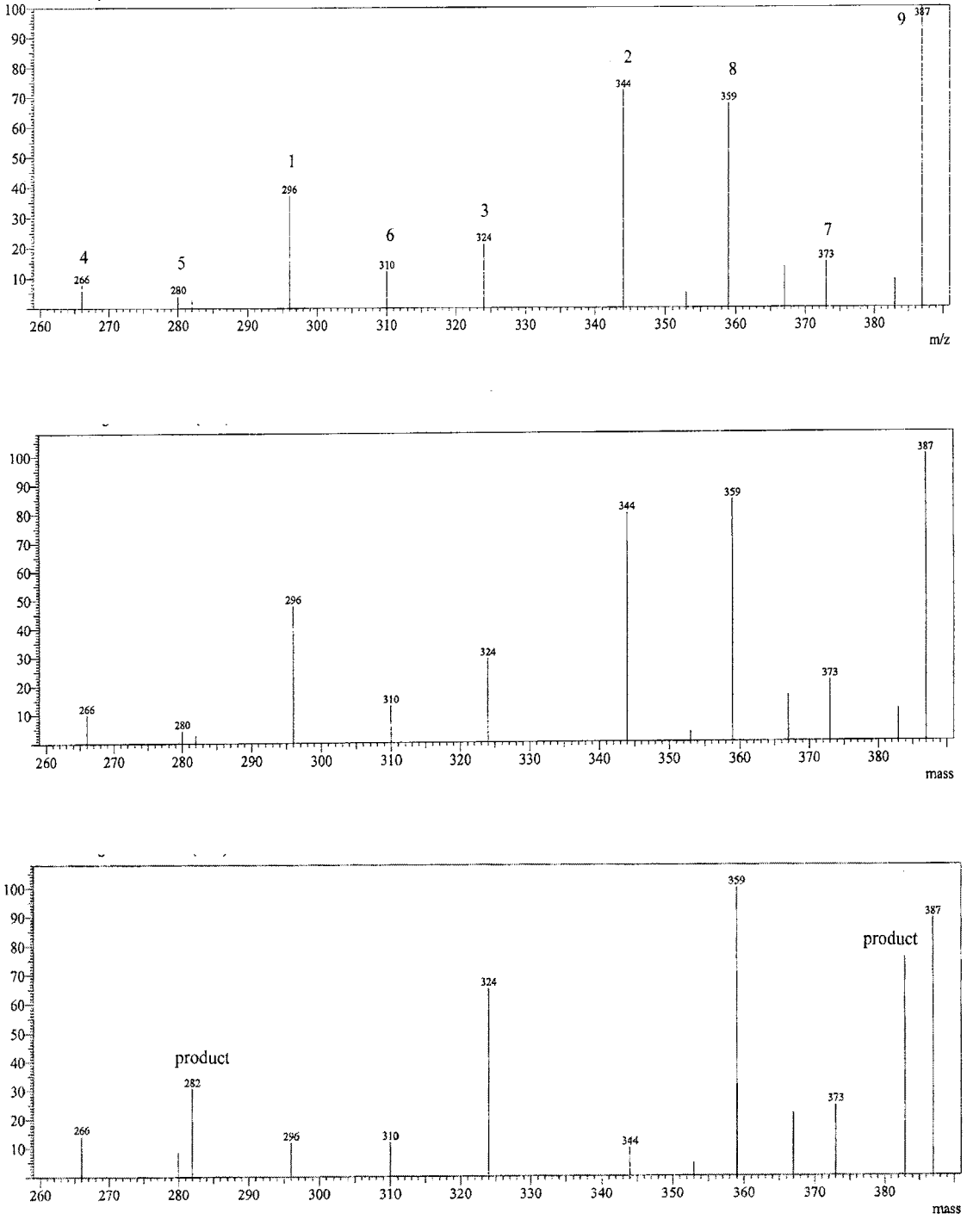


c)



a) Time, 0 hour, b) Time, 8 hours, control, c) Time, 8 hours, reaction

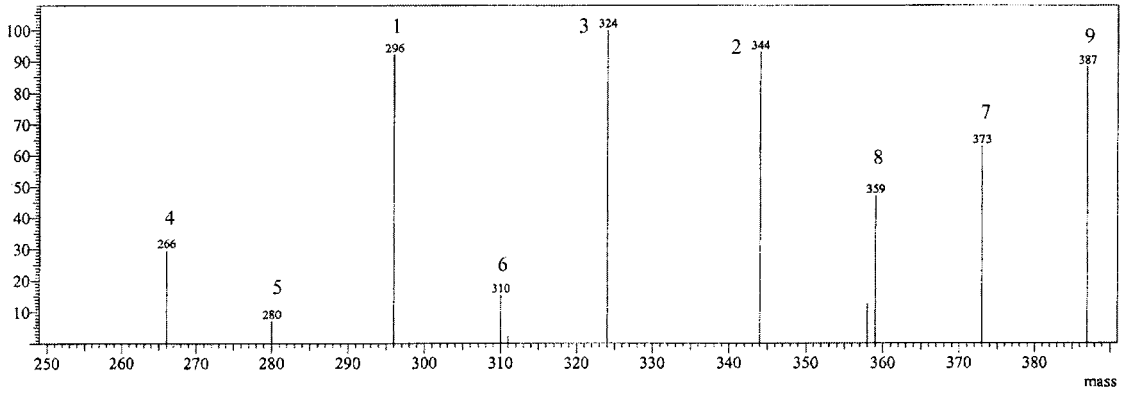
β-glucosidase (chromatographically pure, from Almond)



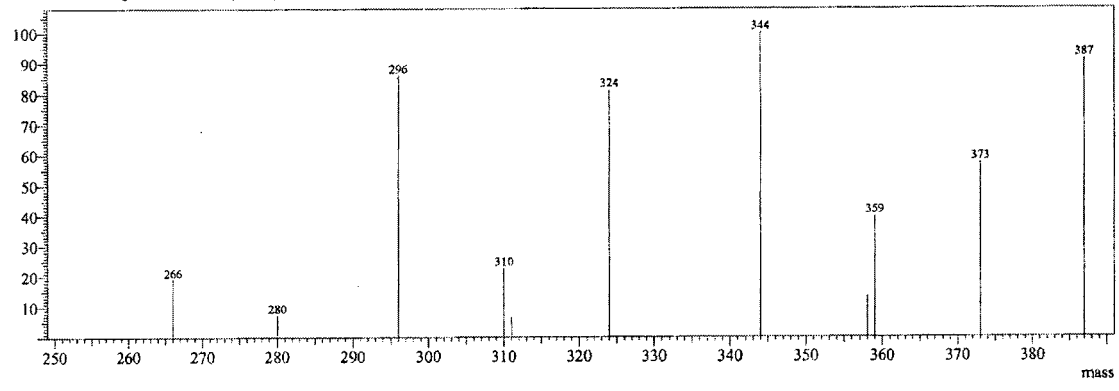
a) Time, 0 hour, b) Time, 12 hours, control, c) Time, 12 hours, reaction

β-*N*-acetyl glucosaminidase (from Jack Bean)

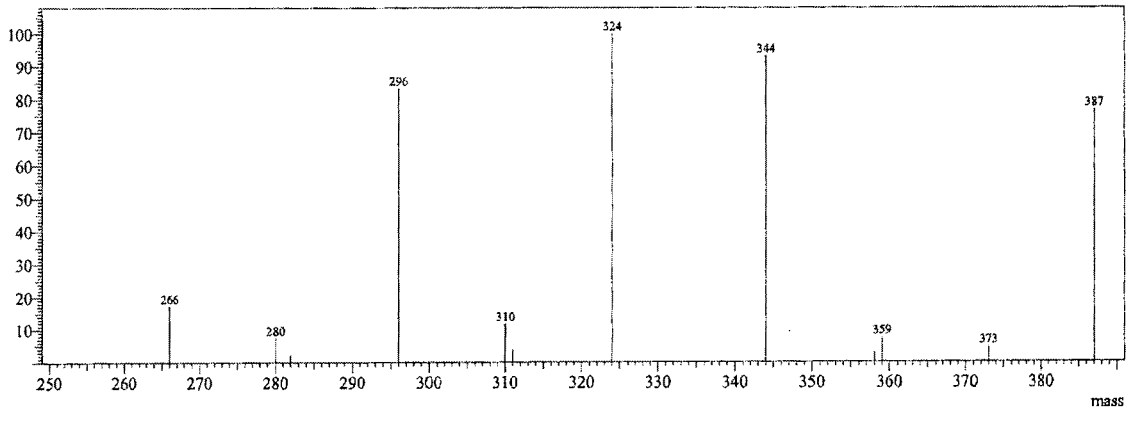
a)



)

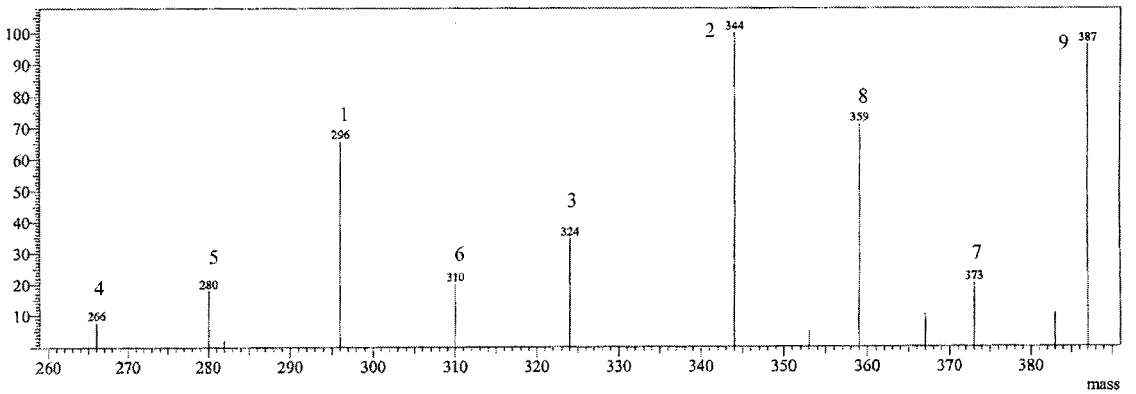


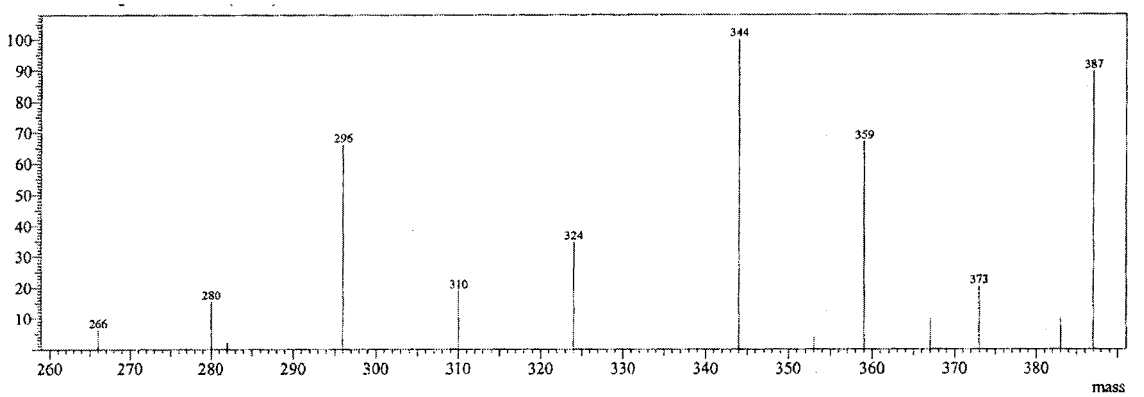
)



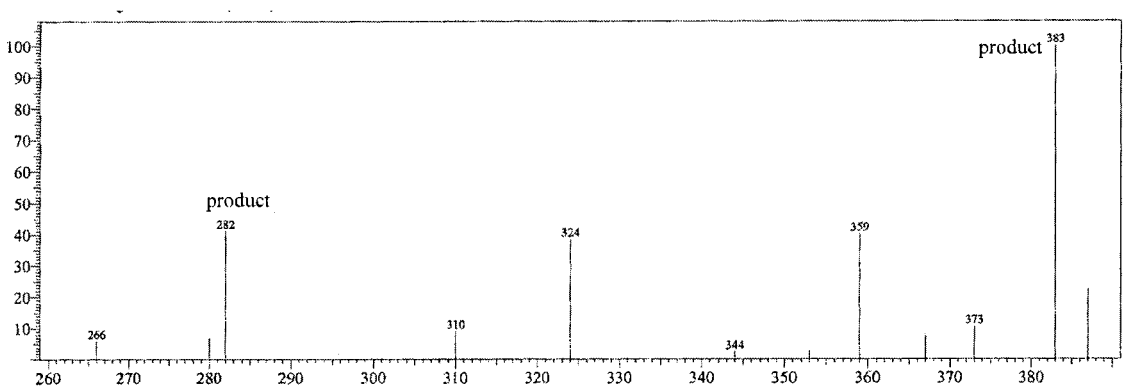
a) Time, 0 hour, b) Time, 1 hour, control, c) Time, 1 hour, reaction

β-glycosidase I (from *Pyrococcus.furiosus*)

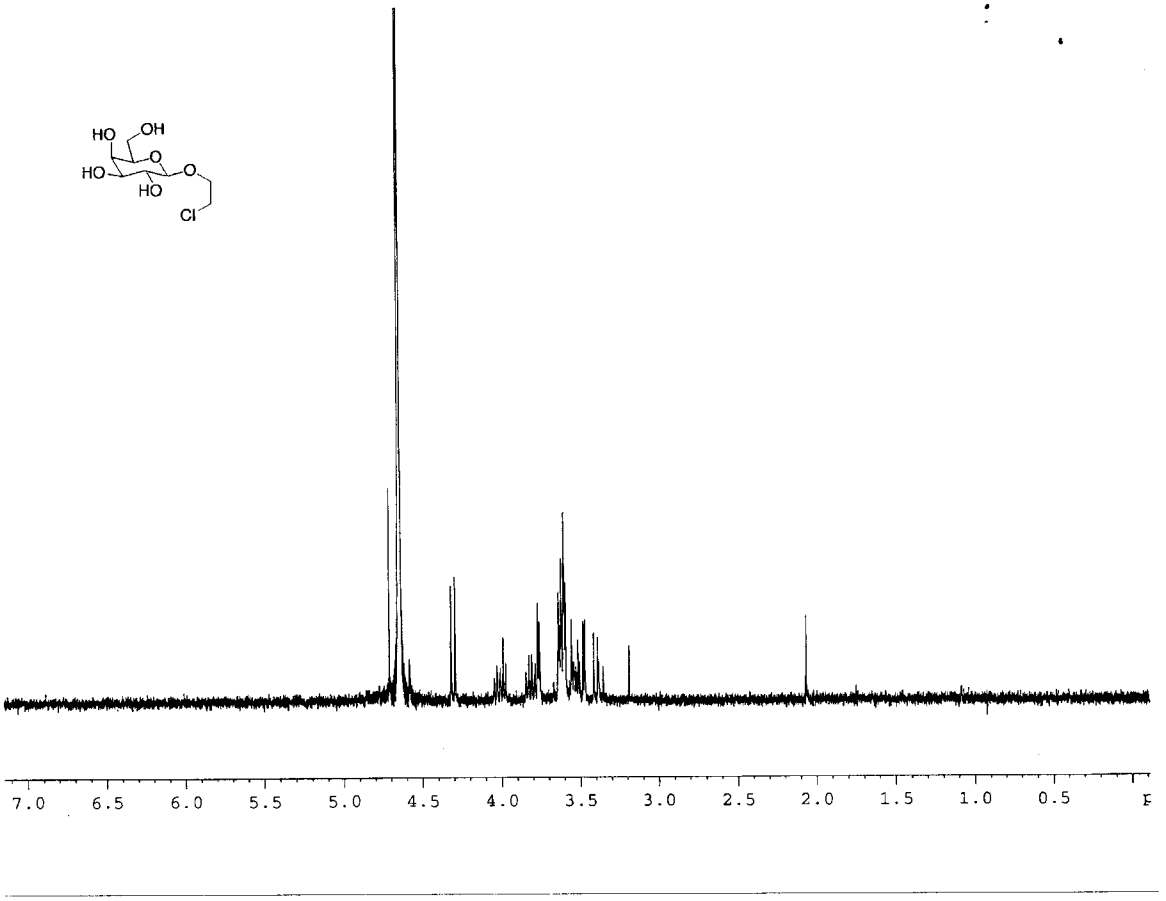




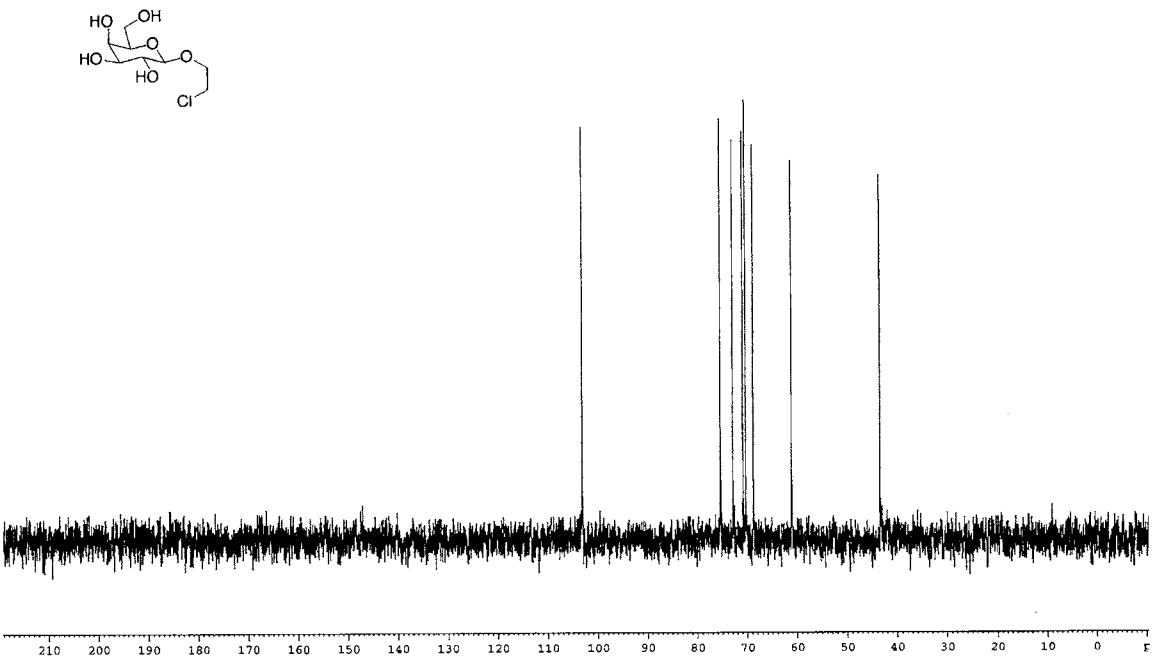
c)



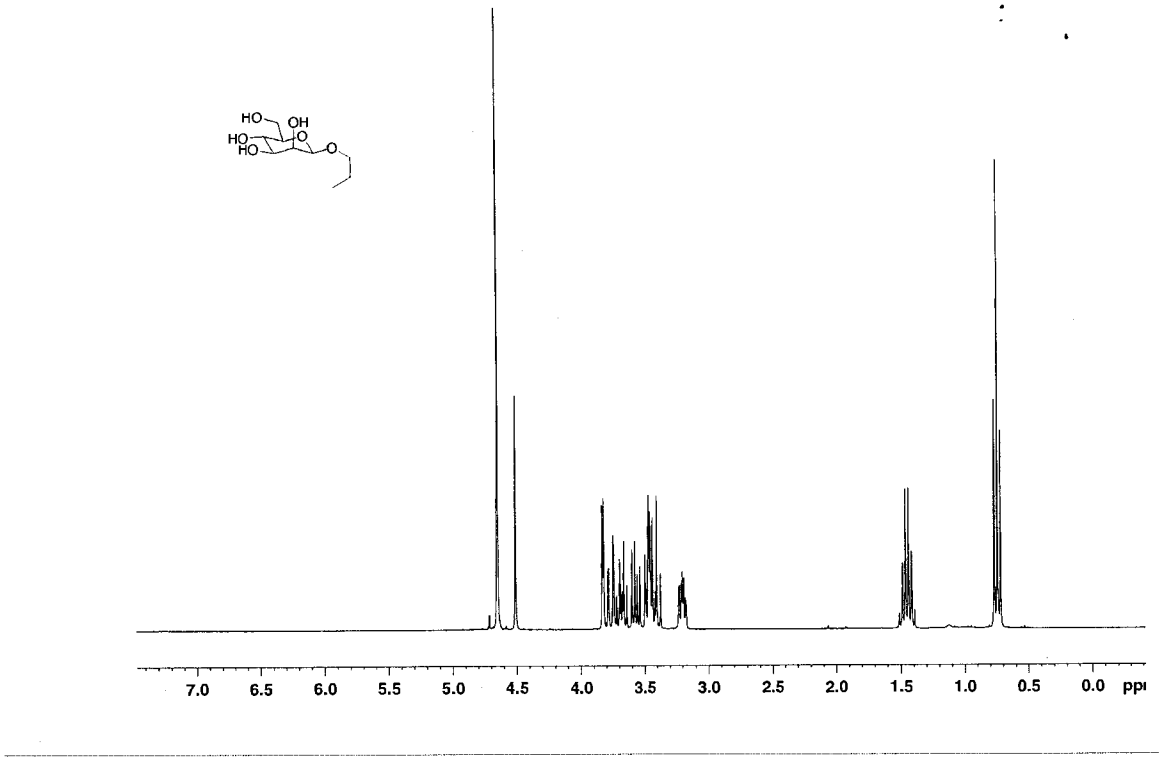
a) Time 0 hour, b) Time 3 hours, control, c) Time 3 hours, reaction



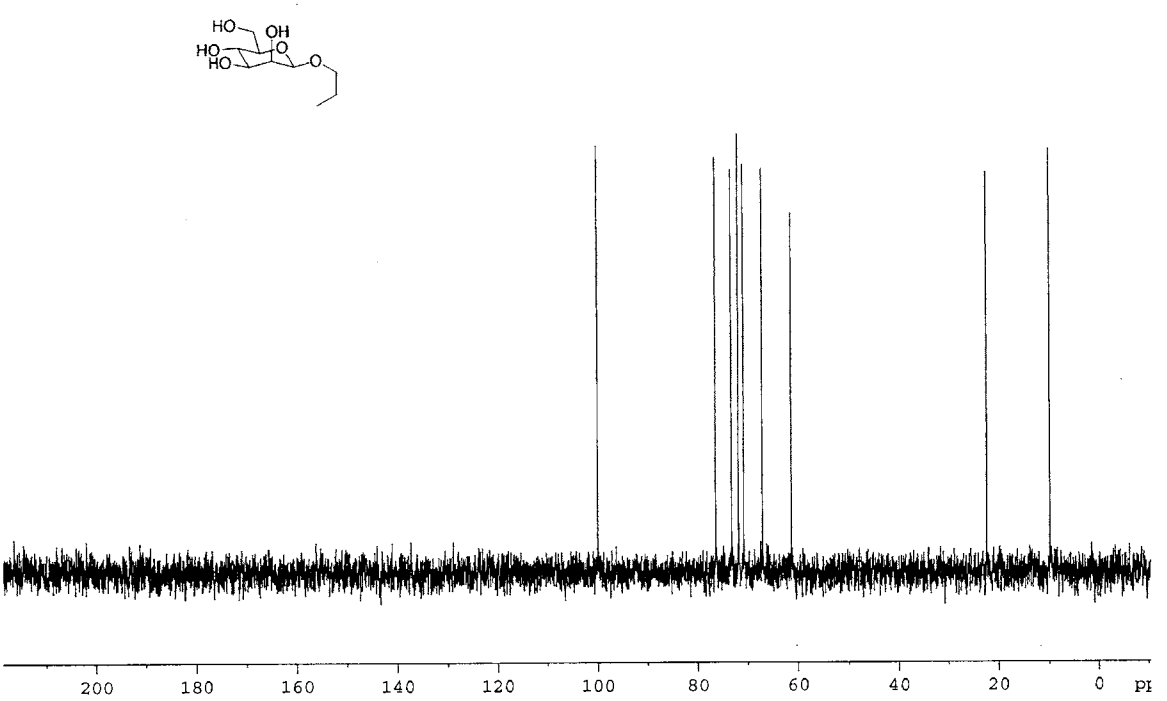
The ^{1}H NMR of 2'-chloroethyl β-D-galactopyranoside



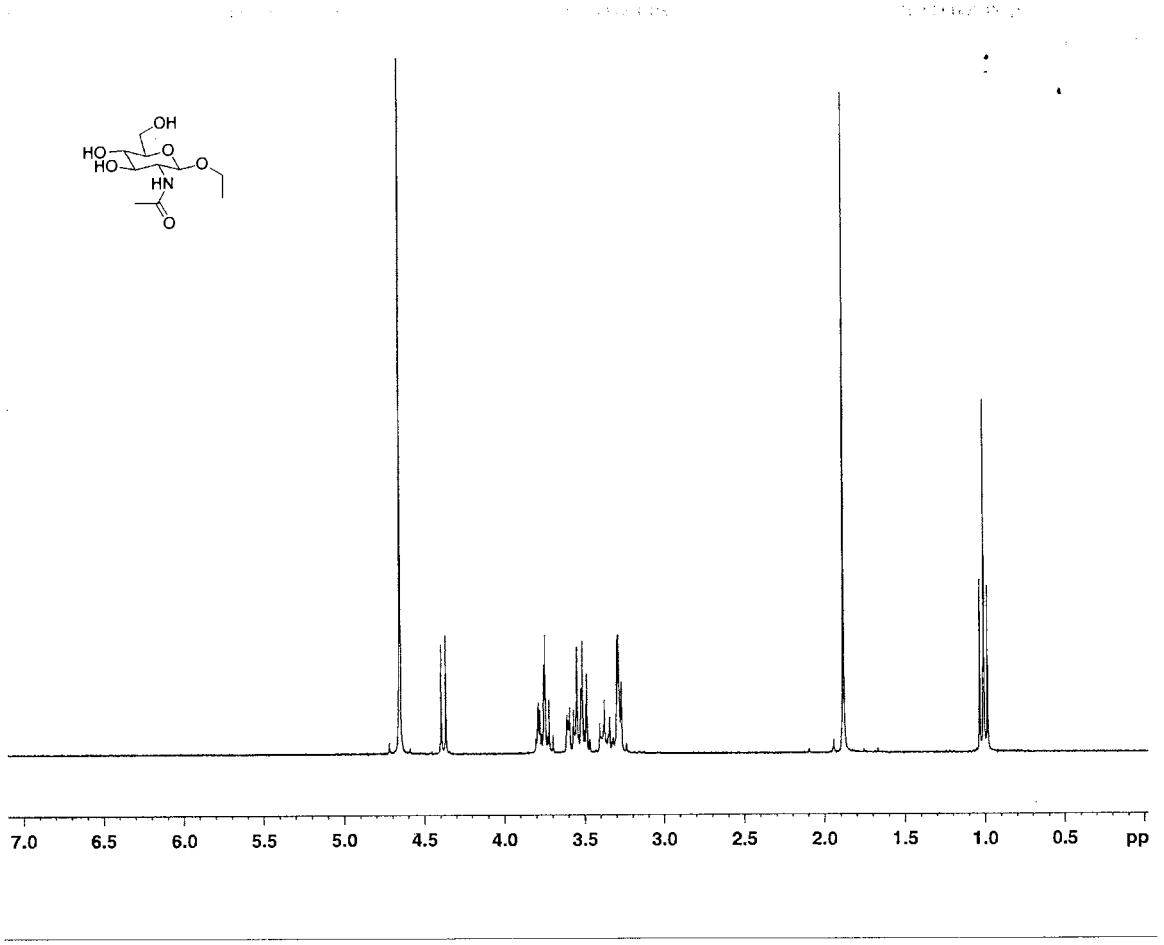
The ^{13}C NMR of 2'-Chloroethyl β-D-galactopyranoside



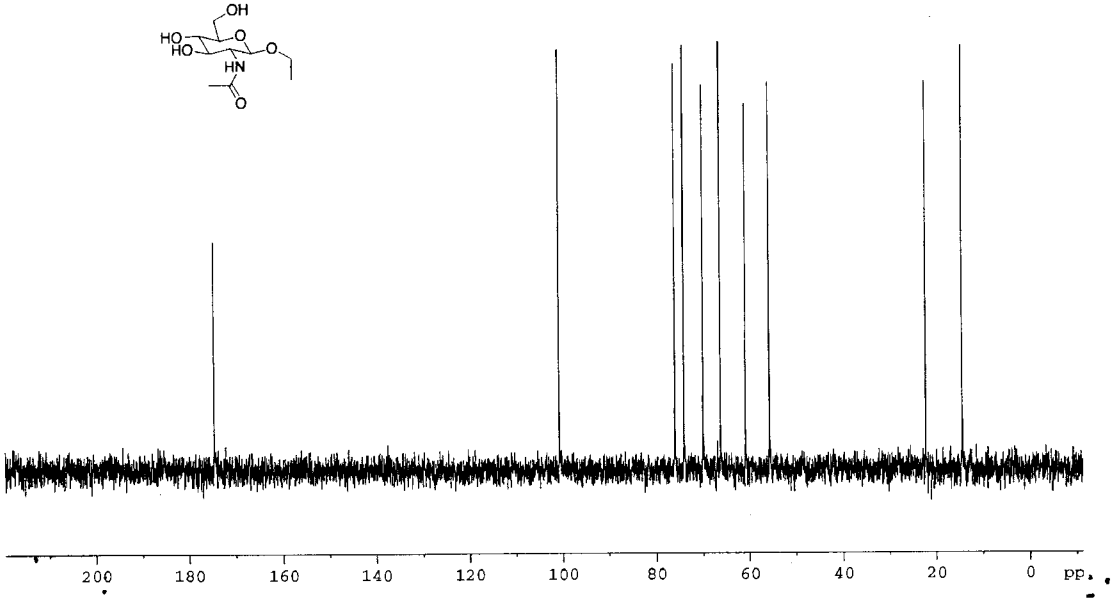
The ^{1}H NMR of propyl β-D-mannopyranoside



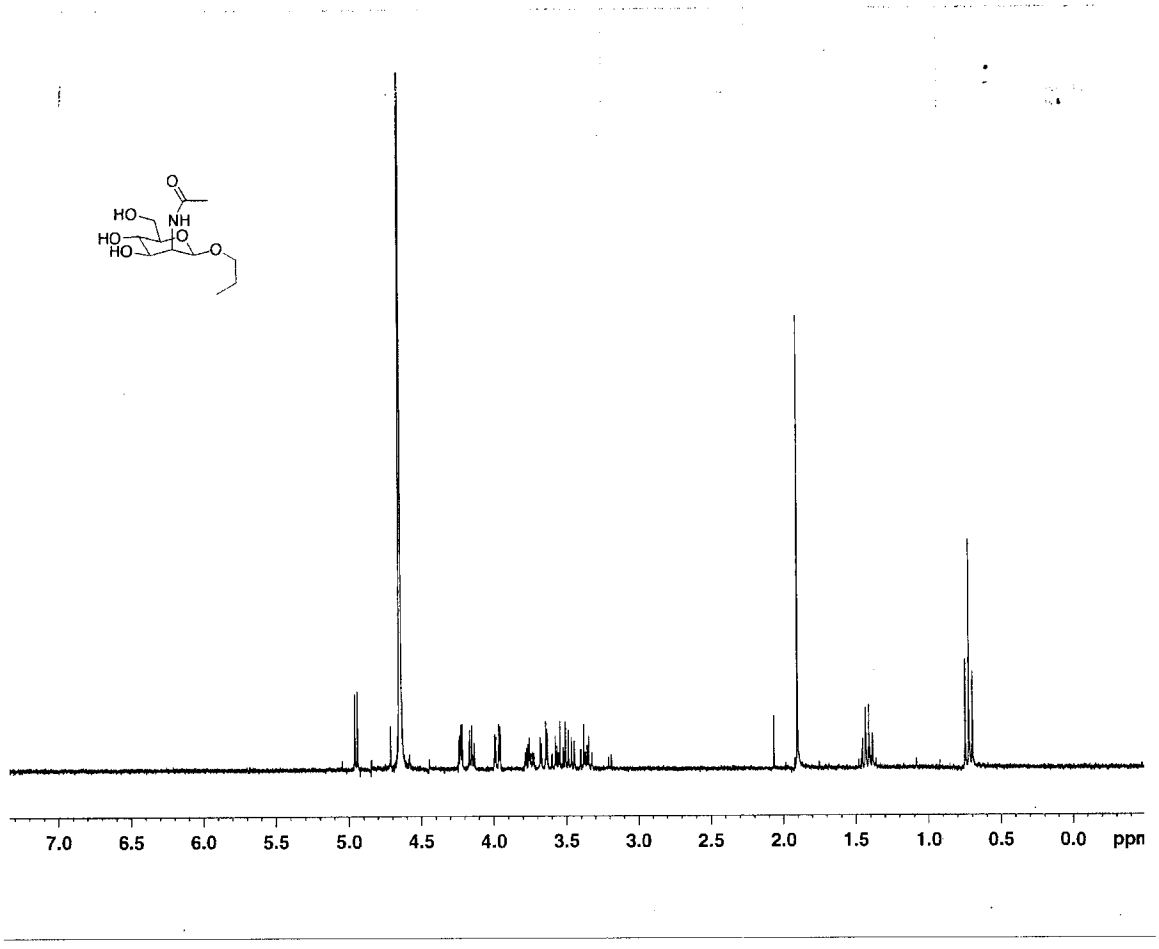
The ^{13}C NMR of propyl β-D-mannopyranoside



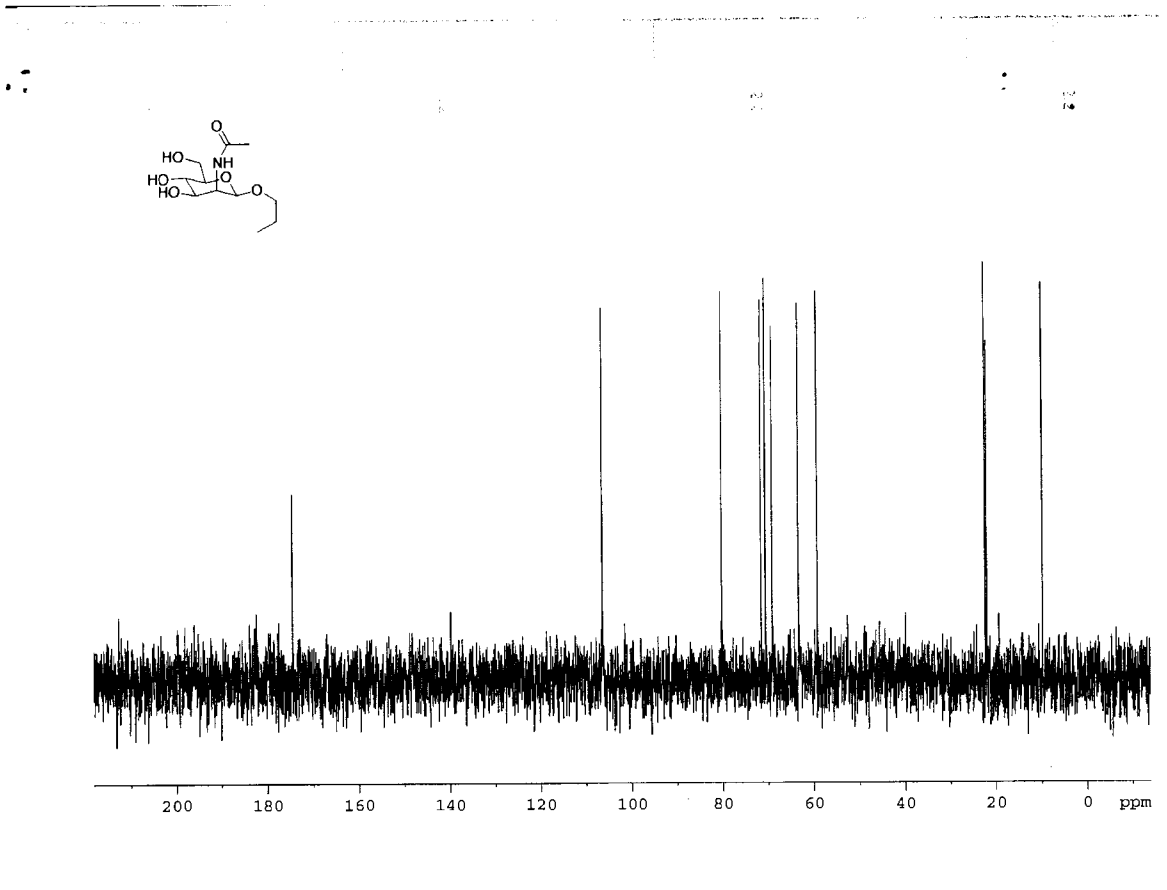
The ^{1}H NMR of ethyl *N*-acetyl-β-D-glucosamine



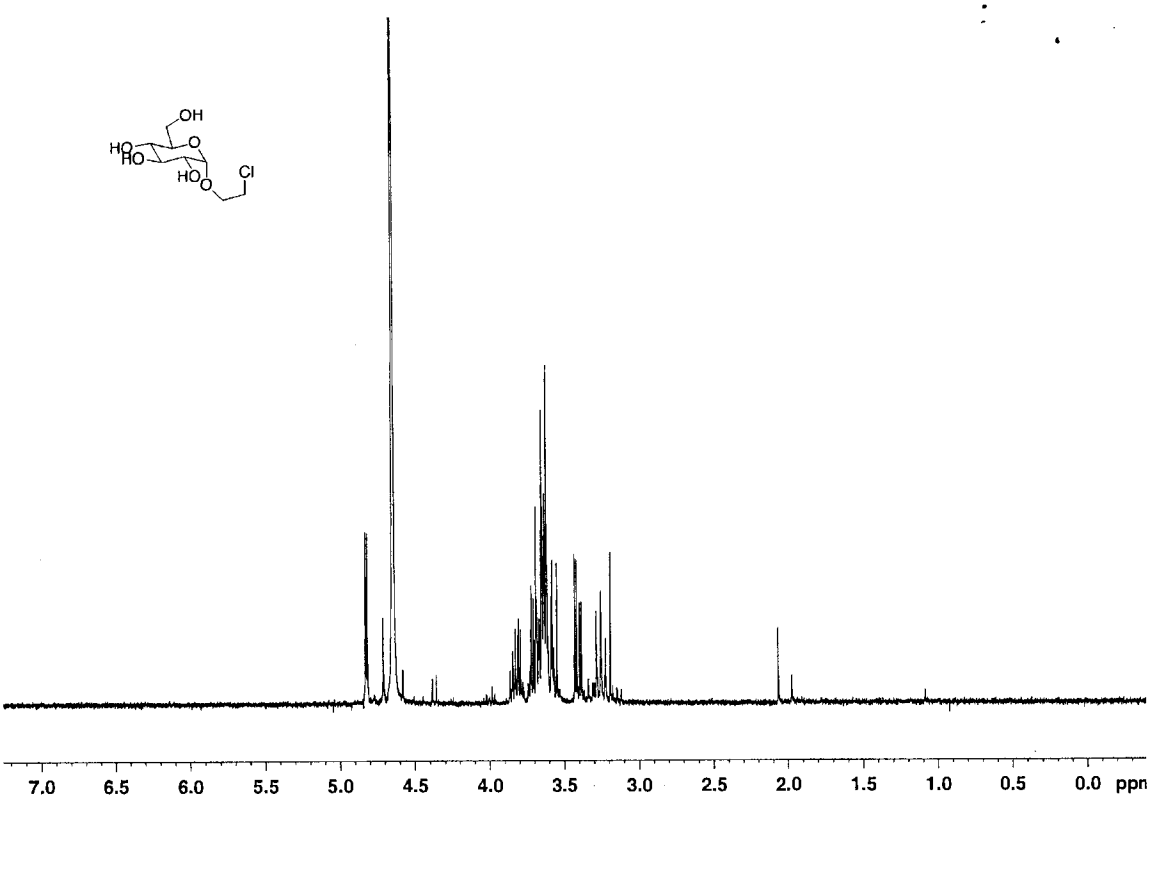
The ^{13}C NMR of ethyl *N*-acetyl-β-D-glucosamine



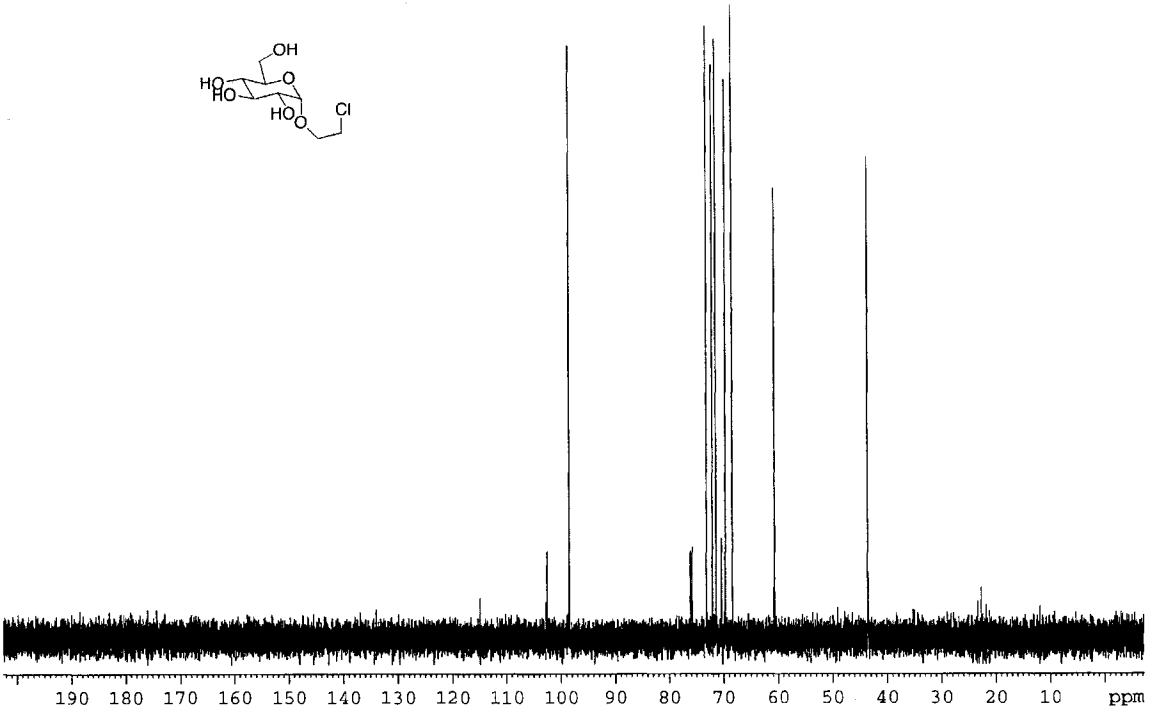
The ^{1}H NMR of propyl *N*-acetyl-β-D-mannosamine



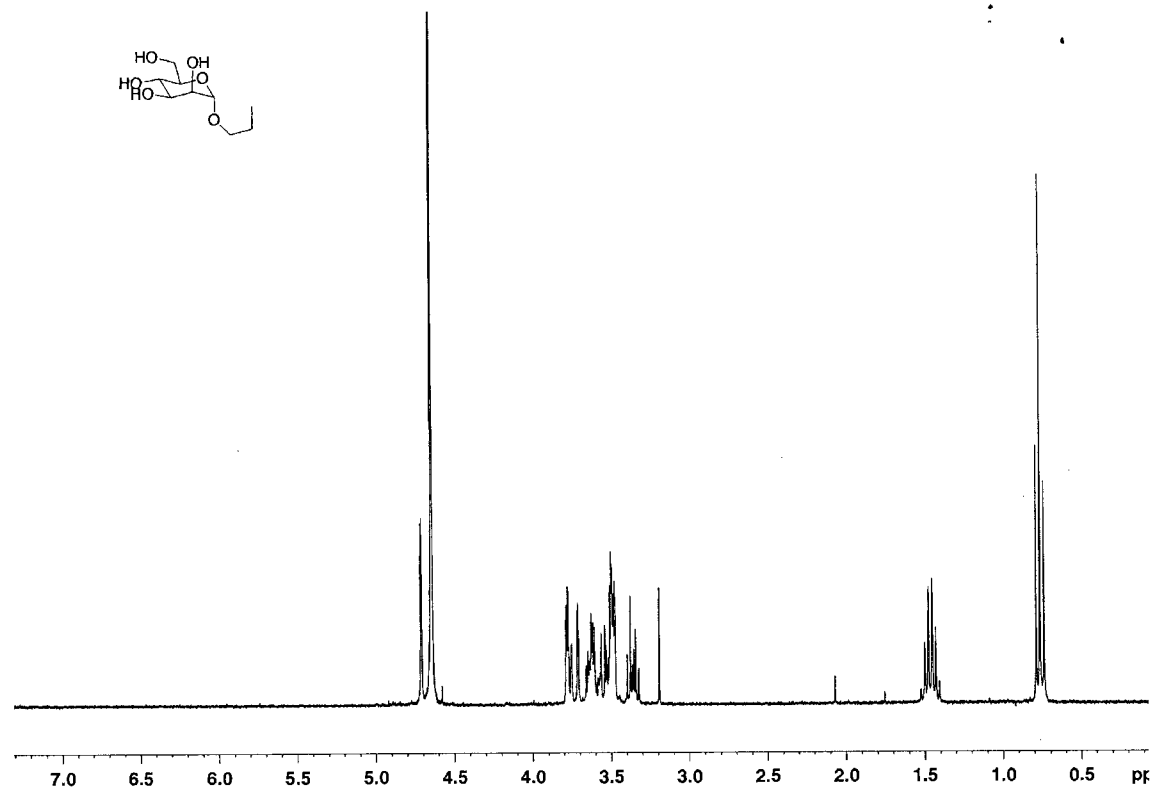
The ^{13}C NMR of propyl *N*-acetyl-β-D-mannosamine



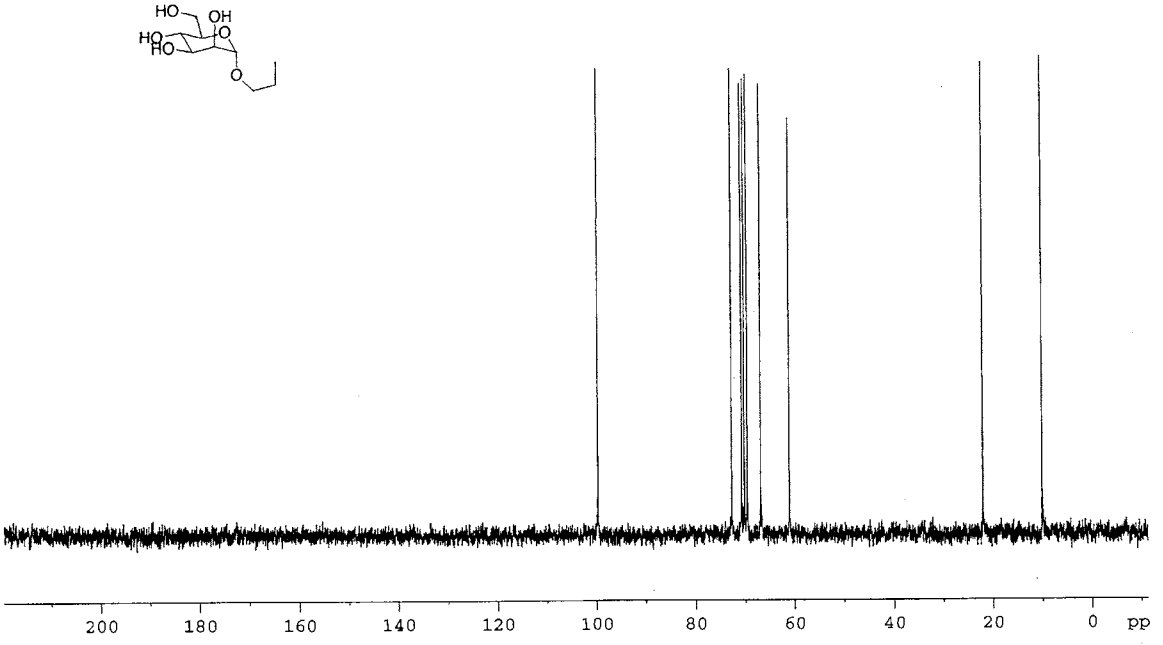
The ^{1}H NMR of 2'-chloroethyl α-D-glucopyranoside



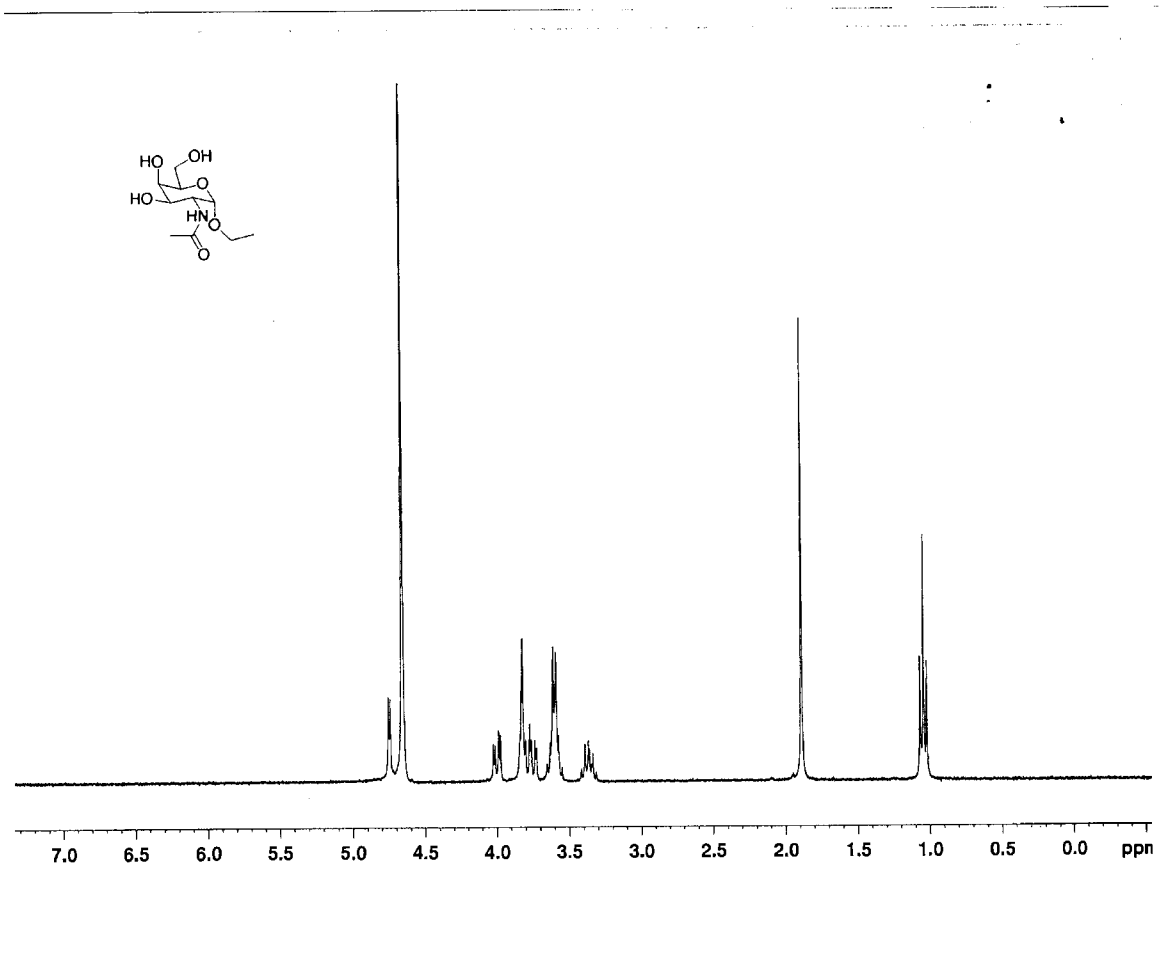
The ^{13}C NMR of 2'-chloroethyl α-D-glucopyranoside



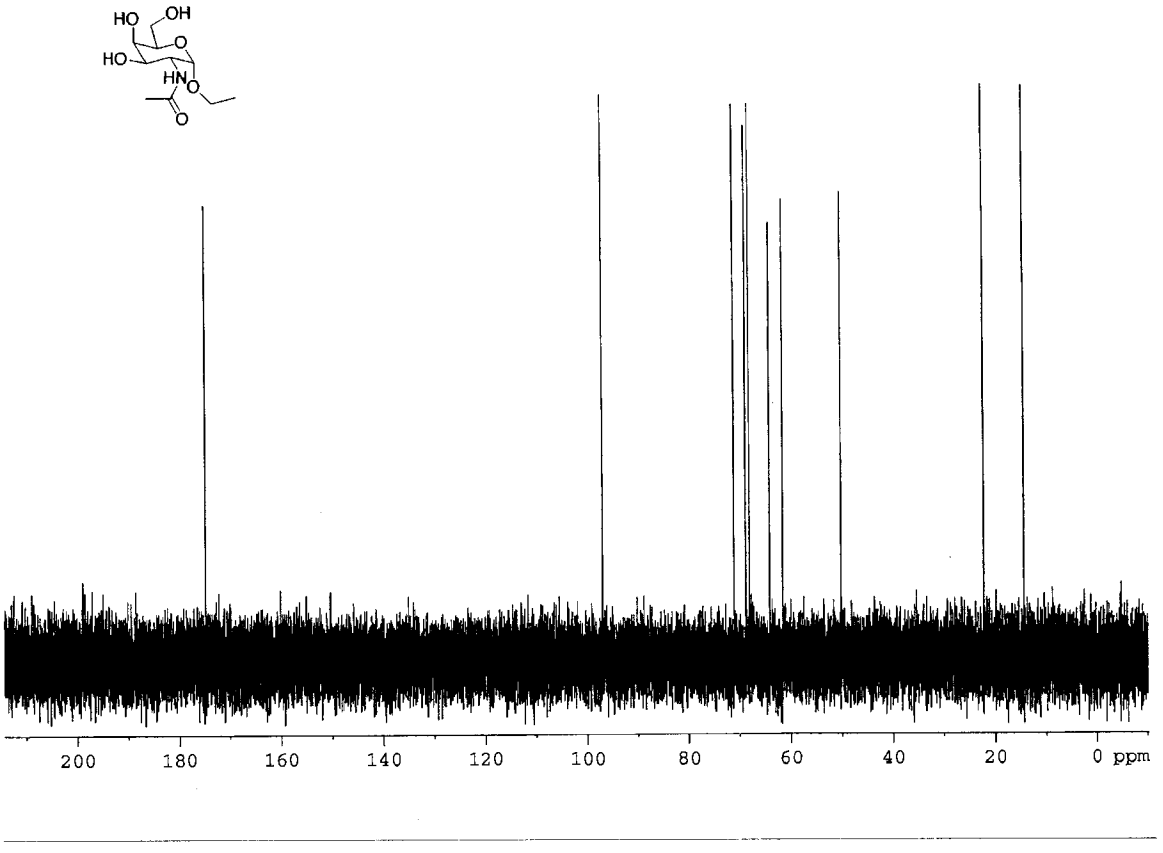
The ^{1}H NMR of propyl α-D-mannopyranoside



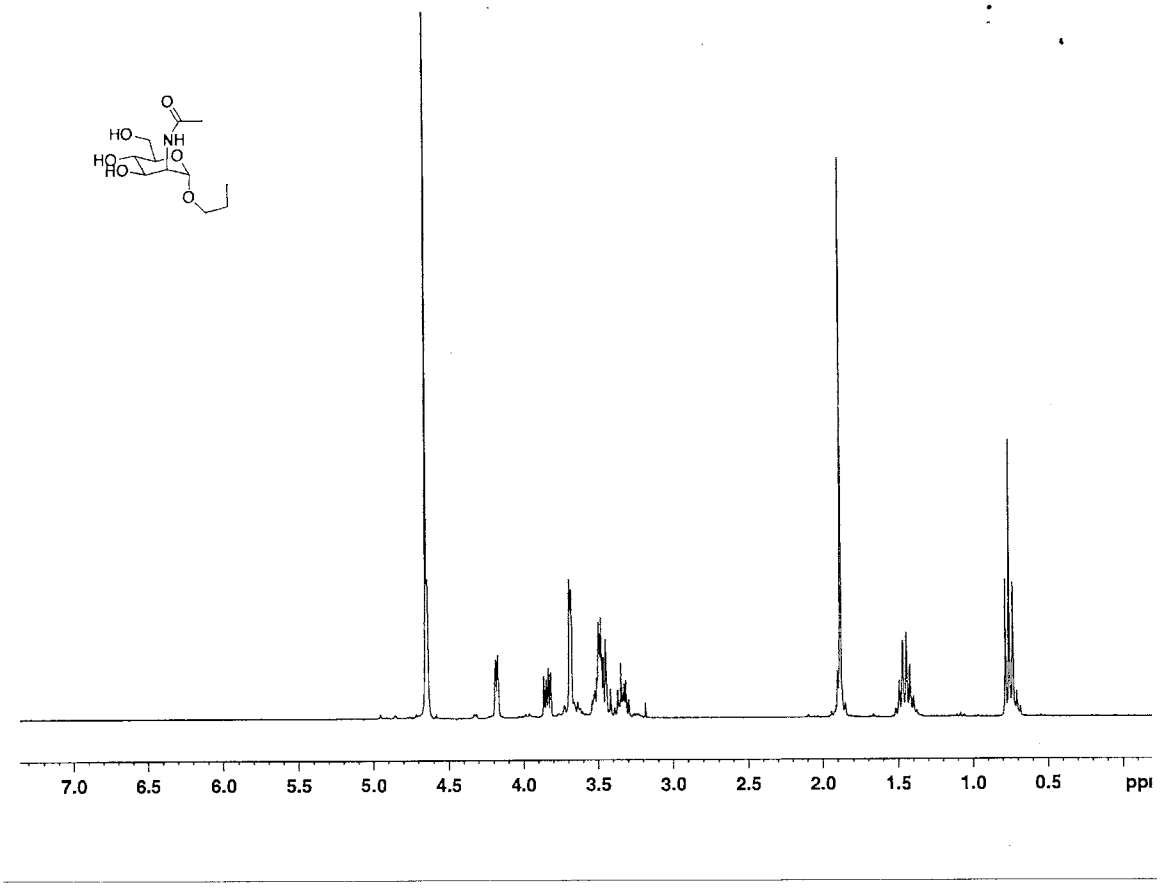
The ^{13}C NMR of propyl α-D-mannopyranoside



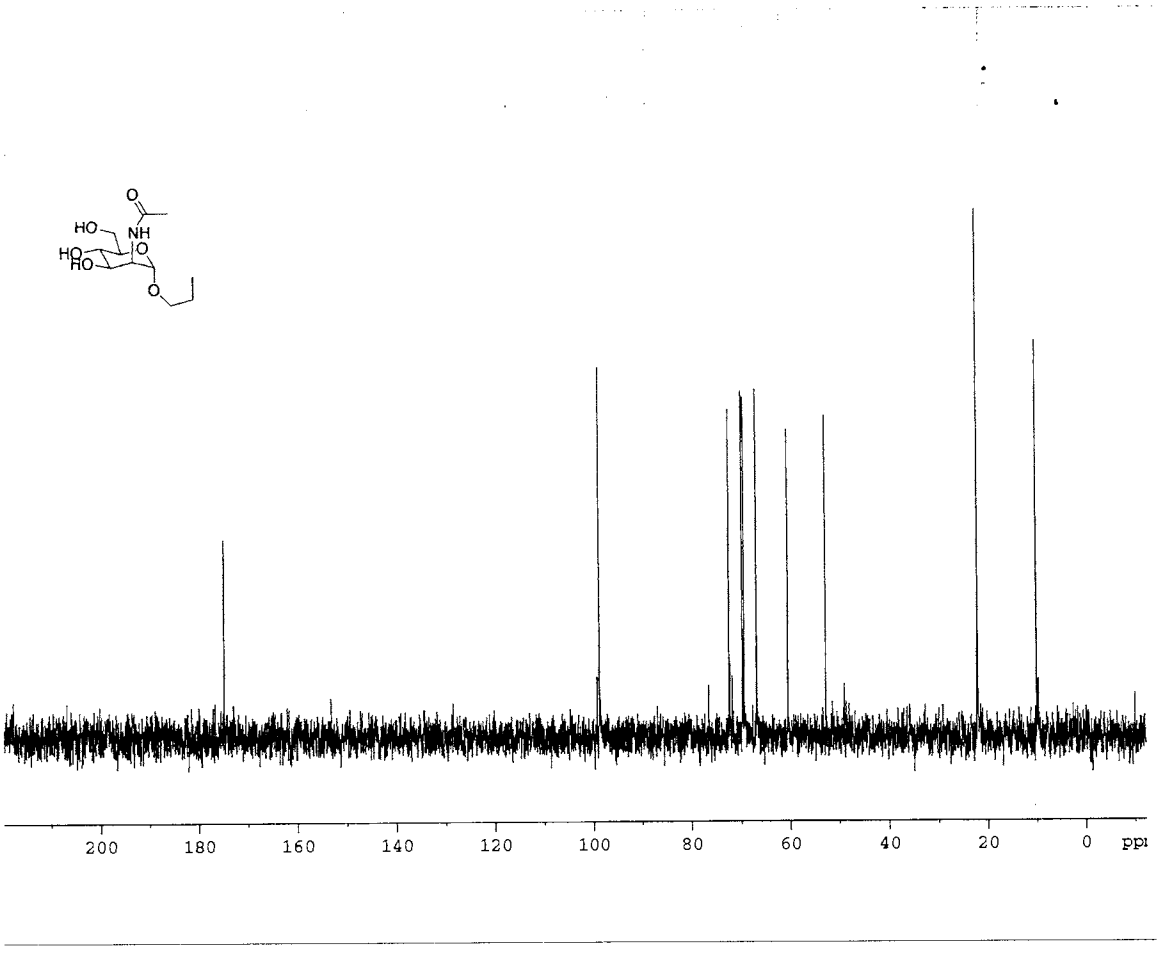
The ^{1}H NMR of ethyl *N*-acetyl-α-D-galactosamine



The ^{13}C NMR of ethyl *N*-acetyl-α-D-galactosamine



The ^{1}H NMR of propyl *N*-acetyl-α-D-mannosamine



The ^{13}C NMR of propyl *N*-acetyl-α-D-mannosamine

Acknowledgments

I greatly appreciated my adviser Dr. Nicola Pohl, who instructed me not only biological and chemical knowledge and skill but also how to work carefully like a professional chemist. Her instruction will benefit me for my whole life. I am very grateful to my senior Ko Kwang-seuk and Corbin Zea who gave me numerous suggestions and assistance. I need to thank the Chemical Instrumentation Facility and the Department of Chemistry at Iowa State University. We also acknowledge an NSF-CAREER award, the Plant Sciences Institute, the USDA, a Shimadzu University Research Grant, and the Herman Frasch Foundation (ACS-PRF) for support of this research.