MULTIFACETED METABOLOMICS APPROACHES FOR CHARACTERIZATION OF
LIGNOCELLULOSIC BIOMASS DEGRADATION PRODUCTS FORMED DURING
AMMONIA FIBER EXPANSION PRETREATMENT

By

Ramin Vismeh

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Chemistry

2012

ABSTRACT

MULTIFACETED METABOLOMICS APPROACHES FOR CHARACTERIZATION OF
LIGNOCELLULOSIC BIOMASS DEGRADATION PRODUCTS FORMED DURING
AMMONIA FIBER EXPANSION PRETREATMENT

By

Ramin Vismeh

Lignocellulosic biomass represents a rather unused resource for production of biofuels, and it offers an alternative to food sources including corn starch. However, structural and compositional impediments limit the digestibility of sugar polymers in biomass cell walls. Thermochemical pretreatments improve accessibility of cellulose and hemicellulose to hydrolytic enzymes. However, most pretreatment methods generate compounds that either inhibit enzymatic hydrolysis or exhibit toxicity to fermentive microorganisms. Characterization and quantification of these products are essential for understanding chemistry of the pretreatment and optimizing the process efficiency to achieve higher ethanol yields. Identification of oligosaccharides released during pretreatment is also critical for choosing hydrolases necessary for cost-effective hydrolysis of cellulose and hemicellulose to fermentable monomeric sugars.

Two chapters in this dissertation describe new mass spectrometry-based strategies for characterization and quantification of products that are formed during ammonia fiber expansion (AFEX) pretreatment of corn stover. Comparison of Liquid Chromatography Mass Spectrometry (LC/MS) profiles of AFEX-treated corn stover (AFEXTCS) and untreated corn stover (UTCS) extract shows that ammonolysis of lignin carbohydrate ester linkages generates a suite of nitrogenous compounds that are present only in the AFEXTCS extract and represent a loss of ammonia during processing. Several of these products including acetamide, feruloyl, coumaroyl and diferuloyl amides were characterized and quantified in the AFEXTCS extracts. The total

amount of characterized and uncharacterized phenolic amides measured 17.4 mg/g AFEXTCS. Maillard reaction products including pyrazines and imidazoles were also identified and measured in the AFEXTCS extract totaling almost 1 mg/g AFEXTCS. The total of quantified nitrogenous products that are formed during AFEX was 43.4 mg/g AFEXTCS which was equivalent to 45-50 % of ammonia that is lost during the pretreatment.

Methodology for identification, detection and quantification of various diferulate cross-linkers in forms of Di-Acids (Di-Ac), Acid-Amide (Ac-Am), and Di-Amides (Di-Am) in AFEX and NaOH treated corn stover using ultrahigh performance liquid chromatography/tandem mass spectrometry (LC/MS/MS) is presented. Characterization of isomeric diferulates was based on the distinguishing fragments formed upon collision induced dissociation (CID) of $[M+H]^+$ ions of each diferulate isomer.

LC separations combined with quasi-simultaneous acquisition of mass spectra at multiple collision energies provide fast spectrum acquisition using a time-of-flight (TOF) mass analyzer. This approach, called mux-CID, generates molecular and fragment ion mass information at different collision energies for molecular and adduct ions of oligosaccharides in a single analysis. Non-selective CID facilitated characterization of glucans and arabinoxylans in the AFEXTCS extracts. A LC/MS gradient based on multiplexed-CID detection was developed and applied to profile oligosaccharides in AFEXTCS extract. This method detected glucans with degree of polymerization (DP) from 2 to 22 after solid phase extraction (SPE) enrichment using porous graphitized carbon (PGC), which proved essential for recoveries of specific oligosaccharides. Arabinoxylans were also detected and partially characterized using this strategy after hydrolysis using xylanase. A relative quantification based on peak areas showed removal of almost 85% of the acetate esters of arabinoxylans after AFEX.

To all who have tried to make me a better person

ACKNOWLEDGEMENTS

I would like to extend my thanks to close friends that over the years at Michigan State have made my stay memorable. Roozbeh in particular, a great brother, friend and the best roommate I could ask for. Ali, Mahnaz, Benyamin (Benjool), Negah, Behrooz, Rafida, Tate, Farid, Meisam, Afra and Atefeh –thank you for being wonderful graduate school companions. I would also like to thank my friends which were supportive in many ways during my last year, especially when I was preparing this dissertation. Thanks to Afrand and Mersedeh (for keeping my caffeine levels high) and Nastaran, Bardia, and Hamideh--thank you guys for being wonderful friends.

I would like to thank my Best Friend, Dr. Behnaz Shafii. Behnaz, I know putting up with me is kind of hard, but you did it for five years and I am very grateful for that. You were by my side during times of happiness and sadness in East Lansing, and, as you always say, that's what true friends are for. Yes, it is hard to come by a true friend, and I consider myself very lucky to have found a true (and super smart) friend in you. I appreciate the opportunity to have learned so much from you. Thank you for all you did for me. I wish you happiness and success in whatever you do, wherever you are. You can always count on me as your life-long friend.

Lastly, this thesis would be incomplete without showing my heartfelt thanks for my family more than anyone else: my wonderful parents ("Houshang" and "Pouran") my crazy, handsome, smart and passionate brother "Milad", and the new member of our family, "Parisa". Thank you for your love, support and encouragement along the way. I could have not made this journey without you. For those who I have lost, I thank you as well, and I keep your everlasting memories with me! Thank you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

5-HMF	**5-Hydroxy Methyl Furfural**
AFEX	**Ammonia Fiber Expansion**
AFEXTCS	**Ammonia Fiber Expansion Treated Corn Stover**
ATCS	**Acid Treated Corn Stover**
CID	**Collision Induced Dissociation**
EI	**Electron Ionization**
EPI	**Enhanced Product Ion**
ESI	**Electrospray Ionization**
GC/MS	**Gas Chromatography Mass Spectrometry**
HILIC	**Hydrophilic Interaction Chromatography**
LC	**Liquid Chromatography**
LC/MS	**Liquid Chromatography Mass Spectrometry**
LCC	**Lignin Carbohydrate Complex**
MALDI	**Matrix Assisted Laser Desorption Ionization**
MRM	**Multiple Reaction Monitoring**
MS	**Mass Spectrometry**
mux-CID	**Multiplexed Collision Induced Dissociation**
NMR	**Nuclear Magnetic Resonance**
OPLS-DA	**Orthogonal Projections to Latent Structures Discriminant Analysis**
PCA	**Principal Components Analysis**

PGC	**Porous Graphitized Carbon**
PLS-DA	**Partial Least Squares Discriminant Analysis**
q-TOF	**Quadrupole-Time-of-Flight**
RI	**Refractive index**
RMD	**Relative Mass Defect**
RP	**Reversed Phase**
SIM	**Selected Ion Monitoring**
SPE	**Solid Phase Extraction**
TMS	**Trimethylsilylation**
TOF	**Time of Flight**
UHPLC	**Ultra High Performance Liquid Chromatography**
UTCS	**Untreated Corn Stover**
UV	**Ultraviolet**

Chapter One: Driving Forces behind Improved Analytical Strategies for Development of Renewable Liquid Fuels

1.1 Introduction

A continuous increase in petroleum demand and efforts to reduce greenhouse gas emissions have ignited interest in development of alternative renewable energy resources. Biofuels such as bioethanol represent sustainable and renewable sources of energy that depend on plant feed-stocks. Bioethanol can contribute significantly to the transportation sector fuel supply with minimal contribution to the net carbon dioxide in the atmosphere.

The current "food to fuel deal" discussion [1-3] has been the subject of controversial debate because production of bioethanol from food resources such as cornstarch increases scarcity of food resources, and in turn generates economic pressure for higher food prices. Hence using lignocellulosic (woody) biomass from rapid-growing plant materials including corn stover, switchgrass and poplar for biofuel production has gained more attention among researchers over the last decade.

Converting lignocellulosic biomass to biofuel remains challenging despite advances in biotechnology, due to the recalcitrance of plant cell walls and barriers to production of fermentable sugars from cellulose and hemicellulose. The conversion process includes some kind of pre-treatment that disrupts structures of lignin and hemicellulose, followed by enzymatic hydrolysis of sugar polymers and fermentation of monosaccharides to form liquids including ethanol. Deep understanding of the chemical composition of the process mixtures in each process step is crucial for optimizing process efficiency. Products of interest to be identified in the entire process include a vast array of small or large phenolic compounds derived from degradation of lignin, carbohydrate-derived byproducts including furans and pyrazines, carbohydrates ranging from monosaccharides to large polysaccharides, metabolites in microorganisms such as yeast, and an assortment of lipids and proteins. Mass spectrometry has been the leading analysis technique for characterization of these products because it is capable of analyzing an incredibly

broad range of substances.

Although there have been great advances in characterization of proteins (proteomics), lipids (lipidomics), metabolites (metabolomics), and oligosaccharides (glycomics), less focus have been directed toward identification of compounds derived from degradation of lignin, which is one of the most abundant natural polymers. More specifically, after treatment of lignocellulosic biomass, literature reports of analysis of degradation products have been largely confined to a list of about 60-70 phenolic aldehydes and acids, carboxylic acids, and furans some of which are known to be inhibitors to hydrolytic enzymes or toxic to fermentation microorganisms [4]. Given the complexity of plant matter, it is unlikely that these constitute the entire range of potential process inhibitors. Furthermore, some of these may have potential value as products with industrial usage, but have been overlooked. Electrospray ionization mass spectrometry offers enormous potential for a more global characterization of process streams since it efficiently ionizes phenolic compounds and provides molecular and fragment mass information that can drive compound identification.

Attention in the biomass research community turned recently toward the objective of achieving a new degree of understanding of the role of degradation products including phenolic compounds from lignin degradation or oligosaccharides released from cell wall sugar polymers. This chapter focuses on mass spectrometry-based platforms that can be used to identify degradation products formed during lignocellulosic biomass treatment, and covers a thorough review of current technologies and challenges in using mass spectrometry for such purposes as well as advancements and recent related literature on sequencing of lignin. This discussion also presents shortcomings and areas that need more attention as well as suggestions for high-throughput analysis of degradation products using LC/MS/MS.

1.2 Biofuels

The growth in world energy consumption demands supplementation, and eventual replacement, of fossil fuels with alternative sources of renewable energy. Biofuels, which are energy resources produced from renewable materials, are considered promising alternatives for fossil fuels. Many biofuels are biodegradable and exhibit less toxicity than fossil fuels, and have prospects to play important economic roles worldwide. Various biofuels including bioethanol, biodiesel, biobutanol and green gasoline are being targeted as liquid transportation fuels. Bioethanol and biodiesel contributed to about 2.7% of the world transportation energy sector in 2010 [5]. It is predicted that this contribution could reach 27% by 2050. Although these numbers are promising, the great challenge that plagues the biofuel industry derives from the "food vs. fuel" competition. For example, more than 95% of bioethanol, the most common biofuel in the United States, is produced directly from corn starch. When shortages of corn harvest occur, economic forces drive the price of corn and bioethanol higher. In 2012, widespread drought conditions led to corn shortages and increased prices, making bioethanol a less economically feasible biofuel [6]. Researchers are aiming to put an end to this "food vs. fuel" debate by producing biofuels from non-food resources [3]. Cellulosic ethanol, produced from lignocellulosic biomass, and algal biodiesel produced from rapid growing algae are two examples of biofuels that have gained considerable interest among researchers [7-14].

1.3 Biomass to biofuel process

The process of converting solid biomass to liquid biofuel depends on the renewable biomass resource and the desired characteristics of the final biofuel product. Application of heat, pressure, enzymes, and chemicals are needed to convert biomass to biofuel. For example; biodiesel, which is composed of monoalkyl esters of long chain fatty acids is mainly produced by

transesterification of oils and fats via a chemical reaction with methanol or ethanol in presence of acid [15,16]. Biodiesel resources include vegetable oil, seed oil, animal fat and more recently, genetically engineered algae with high lipid content [7].

Production of bioethanol from biomass carbohydrate polymers is more complex and costly than biodiesel production, and requires enzymes to hydrolyze carbohydrate polymers including starch to monomeric sugars. Subsequent fermentation of monosaccharides to ethanol is achieved using microorganisms including yeast and bacteria [13]. When using lignocellulosic biomass as the renewable resource, an additional process called pretreatment is necessary to remove structural and compositional impediments to enzyme hydrolysis [17]. Typical pretreatments involve applying acid or base under elevated pressure/temperature to disrupt the structure of lignin, leading to improved release of carbohydrate resources that can be subjected to subsequent hydrolysis and fermentation to form liquid fuels.

1.4 Lignocellulosic biomass

Non-food crop materials being considered as biofuel feedstocks include rapid-growing plant species that are not readily digestible by humans. Perhaps by definition, these nondigestible materials have inherent resistance to enzymatic processing. Such lignocellulosic biomass sources include corn stover, switchgrass, wheat straw, poplar, and other hardwoods, and are mainly composed of three polymers; cellulose (30-50%), hemicellulose (20-40%) and lignin (15-30%) with 5-30% of other components such as proteins, salts and minerals [18].

Cellulose, which is a polymer of β-(1-4) linked D-glucose units, and hemicellulose (a more heterogeneous polymer derived from 5- and 6-carbon sugars that have undergone various chemical modifications) are abundant potential sources of fermentable sugars. However, in

lignocellulosic biomass, these carbohydrate polymers are protected by physical barriers (a complex array of phenolic substances called lignin) that limit access by hydrolytic enzymes. Thermochemical pretreatments can disrupt the structure of lignin and hemicellulose, leading to releases of carbohydrate polymers, yielding higher efficiencies of enzymatic digestion and microbial fermentation. Figure 1.1 shows a schematic of a lignocellulosic biomass cell wall structure.
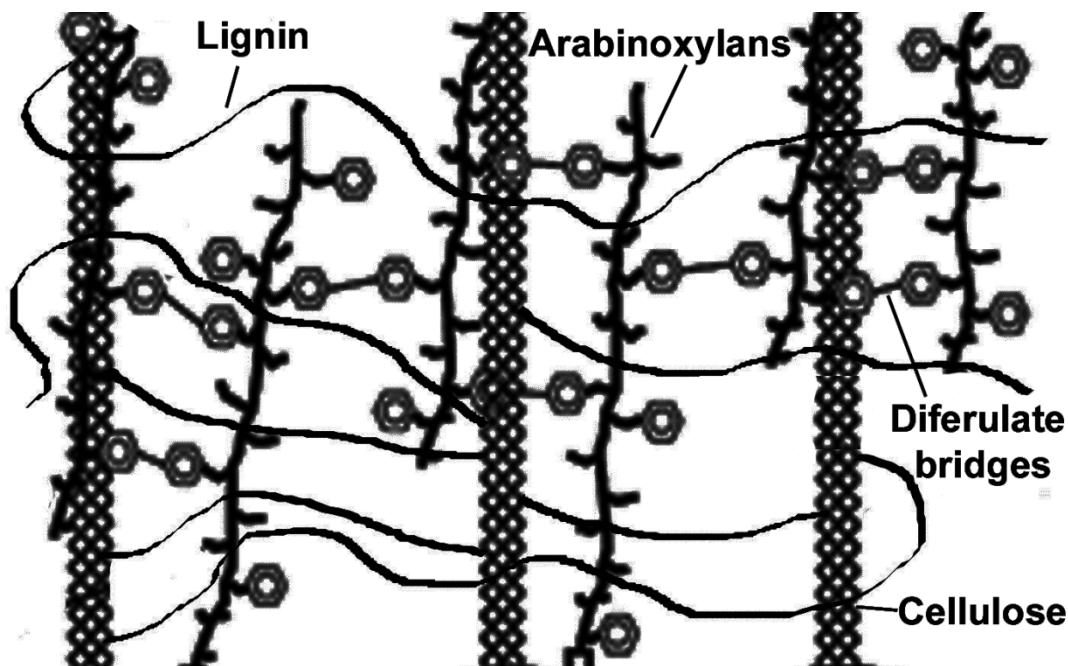
Figure 1.1. Schematic of lignocellulosic plant secondary cell wall structure showing cellulose fibers, hemicellulose arabinoxylans crosslinked by diferulate esters, and lignin. (adapted from Saha, 2003 [19])

1.5 Pretreatment processes

Pretreatment presents one of the most costly steps in production of bioethanol from lignocellulosic biomass, and improvements in pretreatment efficiency offer the potential to

reduce downstream costs of hydrolysis and fermentation [16,20,21]. Pretreatments include physical processes including comminution [22-24] and chemical processes which usually involve application of acids or bases at elevated temperatures to catalyze hydrolytic reactions [15]. The more common pretreatment technologies include steam explosion [25], liquid hot water [26], dilute acid (DA) [27,28], lime pretreatment [29] and ammonia fiber expansion (AFEX) [30,31]. More recently it has been shown that ionic liquids such as 1-butyl-3-methyl- and 1-allyl-3-methylimidazolium chloride can effectively dissolve lignin and therefore have potential for pretreatment applications [32].

Increasing accessible surface area, decrystallizing cellulose, removing hemicellulose and lignin and altering lignin structure are major effects of these pretreatment methods. All mentioned pretreatment technologies affect at least two of these factors [17] and result in higher digestion and fermentation efficiencies, but most pretreatments generate byproducts from lignin and sugar decomposition that inhibit enzymatic and fermentive processing [4,33].

1.6 Effect of biomass degradation products on process streams

Products of cell wall pretreatment can be categorized into five major groups: carboxylic acids, phenolics, furans, inorganic salts and sugars. Extensive research has been conducted on the effect of various degradation products on the fermentation process [4,34-36]. According to Klinke *et al.* the fermentation inhibition may occur due to penetration of degradation products into or through cell membranes (for low molecular weight or inorganic salts) and may influence the expression and activity of sugar and ion transporters (for high molecular weight degradation products) in cell membranes [4]. For instance, based on one suggested mechanism [37], growth inhibition by weak acids has been suggested to arise from their diffusion into cytosol [38] where

the dissociation of acid occurs, leading to a drop in intracellular pH. Plasma membrane ATPase then pumps protons out of the cell (driven by ATP hydrolysis), causing a depletion of ATP in the cells and decreasing rates of critical metabolic transformations [39].

Some mechanisms of inhibition of the yeast *Saccharomyces cerevisiae* by carboxylic acids, furans and phenols have been reviewed by Palmqvist and Hahn-Hagerdal [39]. Findings suggested that phenolic compounds exhibit significant inhibition of the fermentation of lignocellulosic biomass hydrolysates.

In contrast to the extensive research on the influence of degradation products on inhibition of fermentation, less research has focused on the effect of degradation products on enzymatic hydrolysis [40-42], but these factors are recognized as important because inefficient enzyme hydrolysis of cell wall sugar polymers decreases the ethanol yield and hence adds to the total processing costs.

Several phenolics, furan derivatives, and organic acids present in pretreatment process streams are inhibitors of one the most important commercial enzymes used to process cellulose (Spezyme CP, Genencor International, Rochester, NY, USA) with lignin-derived phenolics 4-hydroxybenzaldehyde and vanillin and the furan-derived aldehyde furfural being the most inhibitory [42]. Phenolics are one of the major categories of pretreatment byproducts that show inhibition to both enzymes and fermentation, but many questions about the mechanisms underlying this inhibition remain to be answered.

Monomeric or dimeric sugars have also been considered as potential inhibitors of enzymes important in biomass processing [43]. In the enzymatic hydrolysis process that follows all pretreatments, cellulose is converted by cellulases to the disaccharide cellobiose, and subsequent hydrolysis yields glucose. Glucose, which is the main desired product of cellulase activity, often

reaches process stream concentrations exceeding 6% by weight (~300 mM). At such high concentrations, glucose inhibits β-glucosidase which is one of the enzymes that catalyzes conversion of cellobiose to glucose. Cellobiose also inhibits cellulase, which is used to hydrolyze cellulose in the first place, and so the formation of important intermediates and the desired glucose endproduct slows the enzymatic hydrolysis process [43-45].

1.7 Oligosaccharides released during biomass pretreatment process

Characterization of oligosaccharides released during thermochemical pretreatment of biomass provides information essential for enzymatic hydrolysis optimization. Detailed structural information of oligosaccharides in the biomass hydrolysate aids to the selection cost-effective and efficient enzyme cocktails for converting sugar polymers to fermentable monosaccharides and this information goes beyond destructive conversion of oligosaccharides to monomeric sugars.

Structure of cell wall sugar polymers differ quite significantly among biomass feedstocks. The composition of cellulose is similar in all biomass resources. In contrast, among different plant cell walls, composition seems to vary with regards to hemicellulose (Figure 1.2). For example, hemicellulose in grasses is mainly composed of arabinoxylans, which are complex, highly heterogeneous polysaccharides consisting of a linear β-(1-4) linked xylopyranose backbone to which α-L-arabinofuranose units are attached via α-(1,3) and/or α-(1,2) linkages. The degree of arabinosylation depends on the plant species. Wheat straw contains low degree of arabinose substitution, whereas sorghum xylans are highly branched with arabinopyranosyl groups. Softwood exceptionally contains mannan polymers (polymers based on mannose) including O-acetylated galactoglucomannans as the main hemicellulose sugar polymers. In

9

contrast, in hardwoods, the hemicellulose is mainly composed of arabinoxylans. With this degree of structural diversity, there is a need for fast and improved techniques to characterize oligosaccharides released from treatment of biomass.

Elucidation of oligosaccharide structures often depends on extensive analysis using NMR spectroscopy. However, this technique usually requires milligram quantities of purified material which is challenging and time-consuming to obtain from complex mixtures derived from treated biomass. Relative contents of monosaccharide residues, types and amounts of specific linkages, and anomeric configuration have been obtained for oligosaccharides using NMR, however residue sequence information cannot usually be obtained owing to substructural redundancy. In contrast to NMR, mass spectrometry does not need purified samples particularly if coupled with chromatography. Mass spectrometry has found growing use for characterization of oligosaccharides. Ionizing oligosaccharides using electrospray ionization (ESI) or matrix-assisted laser desorption ionization (MALDI) followed by collision induced dissociation has been one of the main techniques for obtaining sequence, branching and linkage information for a variety of oligosaccharides [46-49].

Efforts to characterize oligosaccharides from different sources of biomass led to identification of various forms of cell wall carbohydrate polymers including galactoglucomannans in spruce [50], arabinoxylans with glucopyranosyl uronic acid branches in wheat bran and corn cobs [51], O-acetylated glucomannans in aspen and birch wood [52] and xyloglucans in the subclass Asteridae plants. The major focus of biomass research on carbohydrate polymers, however, has centered on arabinoxylans because xylose is the main fermentable monosaccharide in biomass after glucose, and both are the most abundant sources of fermentable sugars available in plant cell walls (Figure 1.2). The factors that govern conversion

of cell wall polymers to these monosaccharides are not yet well established, and are the subject

of Chapter Four in this dissertation.
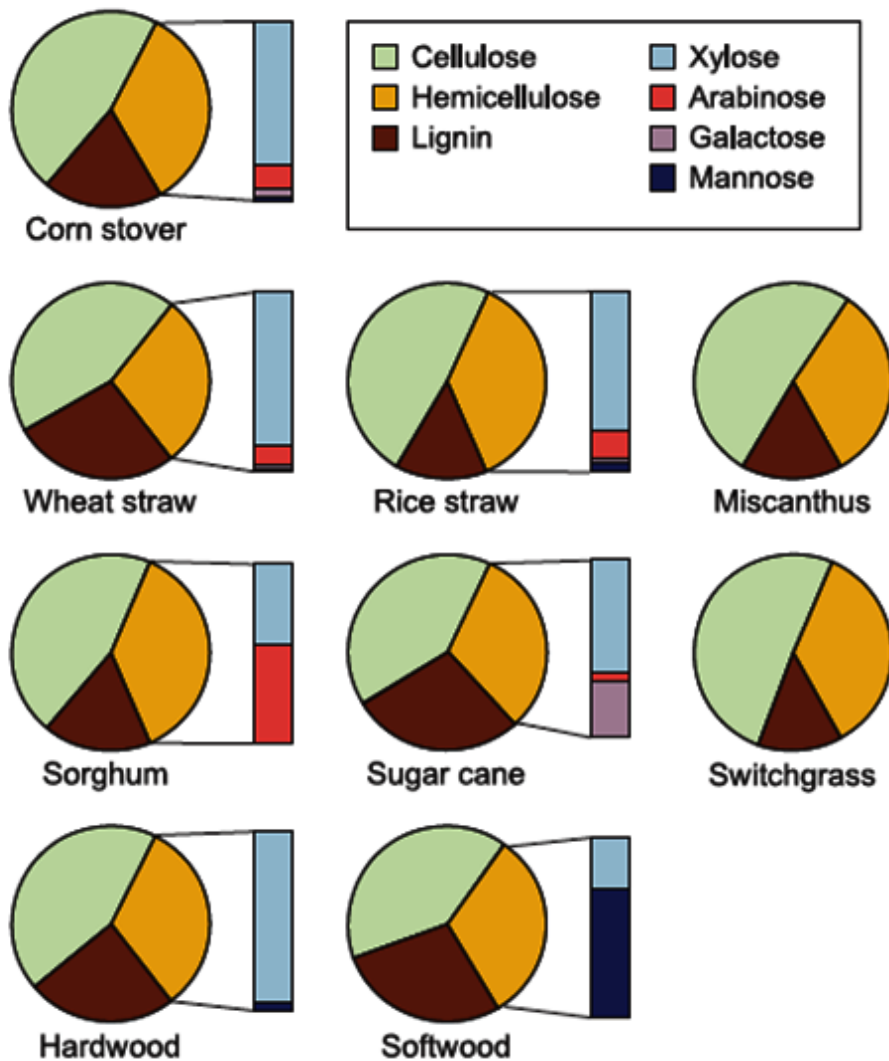
Figure 1.2. Cell-wall polymer composition (percentages of cellulose, hemicellulose and lignin displayed in pie chart) and percentages of components of hemicellulose (bar chart) for a variety of plant materials currently under consideration for use as biofuel feedstocks [18]. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

1.8 Common cell wall degradation products from lignocellulosic material

Due to the complexity of the mixture of products generated by biomass pretreatments, complete characterization of the products in pretreatment process streams requires enormous effort. However, many degradation products can be predicted based on current knowledge of structures of lignin and hemicellulose. Lignin is the second most abundant polymer on earth after cellulose, and may contribute up to 30% (w/w) of lignocellulosic biomass [18,53]. Lignin is an amorphous three-dimensional material based on an irregular assortment of crosslinks between phenylpropanoid monomers, and these traits confer low solubility and molecular heterogeneity that have presented barriers to lignin characterization.

Much current understanding of lignin structure was derived from analysis of lignin structural units that survived an assortment of chemical degradation processes. Numerous studies have confirmed that the principal precursors of lignin are three cinnamyl alcohols varying in aromatic ring substitution including *p*-hydroxycinnamyl alcohol (H), coniferyl alcohol (G) and sinapyl alcohol (S) which will be called H, G and S units from here on (Figure 1.3) [54,55]. Some well-accepted structural models of lignin were proposed by Freudenberg (1968), Nimz (1974), Adler (1977) and Sakakibara (1980) [56-59]. All these models are based on random radical polymerization among the three cinnamyl alcohols, and the simultaneous occurrence of free radicals and atmospheric oxygen can lead to oxidation of the cinnamyl alcohol side-chains. It is therefore not surprising that aldehyde, ketone and carboxylic acid derivatives of lignin H, G and S structural units are present in biomass pretreatment process streams [4]

The polymerization chemistry that generates lignin *in vivo* yields an amorphous material that remains somewhat resistant to heat, particularly in inert atmospheres [53]. Thermogravimetric analyses have demonstrated mass losses from lignin during heating to about 600 ˚C that account for only about one-third to one-half of lignin mass [60], Carbon dioxide

frequently dominates the volatile pyrolysis products, and is often accompanied by an assortment of hydroxy- and methoxy-substituted aromatic compounds [61]. To disrupt lignin's structure for generation of more soluble degradation products, acids or bases are usually applied at elevated temperatures as part of various pretreatment methods. Application of acid, base and heat to lignin forms a variety of degradation products, some of which simply derive from rupture of bonds within lignin whereas others are products result from chemical reactions that occur during pretreatments including hydrolysis and ammonolysis [62].
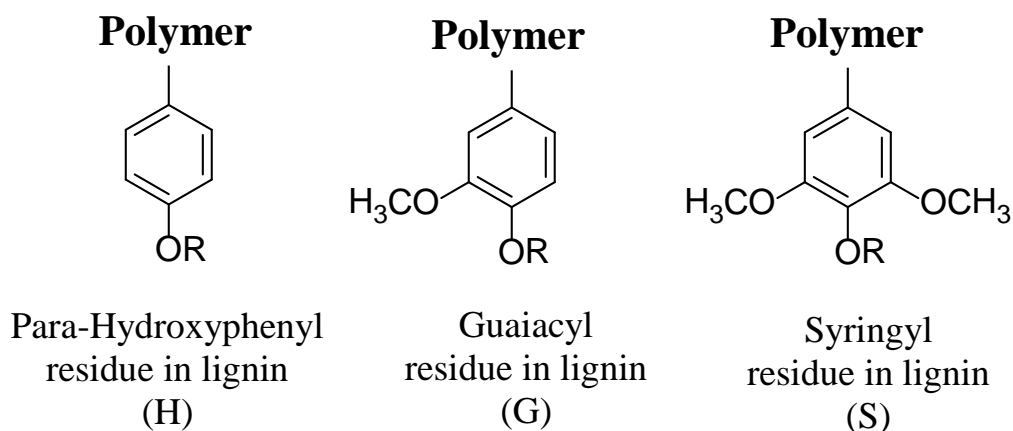
Figure 1.3. Substructures of three main building blocks of lignin, R indicates cross-linking via ether or ester bonds with similar monolignols.

Pretreatment of lignocellulosic biomass generates a complex mixture of products derived from all of the components of the plant cell wall. Figure 1.4 shows categories of degradation products known to form during pretreatment of lignocellulosic biomass. Phenolic monomers derive directly from decomposition of lignin [63]. Oligomeric carbohydrates and monosaccharides come from partial breakdown of cellulose and hemicellulose via hydrolysis or solvolysis [63]. Furan derivatives are products of dehydration of furanose carbohydrates,

including fructose and xylose, which are released from oligosaccharides [4]. Glucose can also rearrange to assorted furan derivatives during pretreatments [42]. Carboxylic acids present in pretreatment process streams derive from oxidation of phenolics or in some cases furans (for example: levulinic and formic acid) [64]. The occurrence of acetic acid has been attributed to hydrolysis of acetyl groups on arabinoxylans in hemicellulose and degradation of HMF [63].

In recent years, growing recognition of the importance of comprehensive understanding of process stream chemistry has driven development of analytical protocols to identify and measure constituents of biomass hydrolysates. The diversity of functional groups and physical properties presents substantial technical challenges, but modern efforts have established a list of 60-70 compounds detected in extracts of pretreated lignocellulosic biomass (Table 1.1). Initial focus has been centered on common organic acids including acetic and formic acids, and the aldehydes furfural and 5-hydroxy methyl furfural (5-HMF) because they are notorious inhibitors of fermentation for many microorganisms [4,65-67].
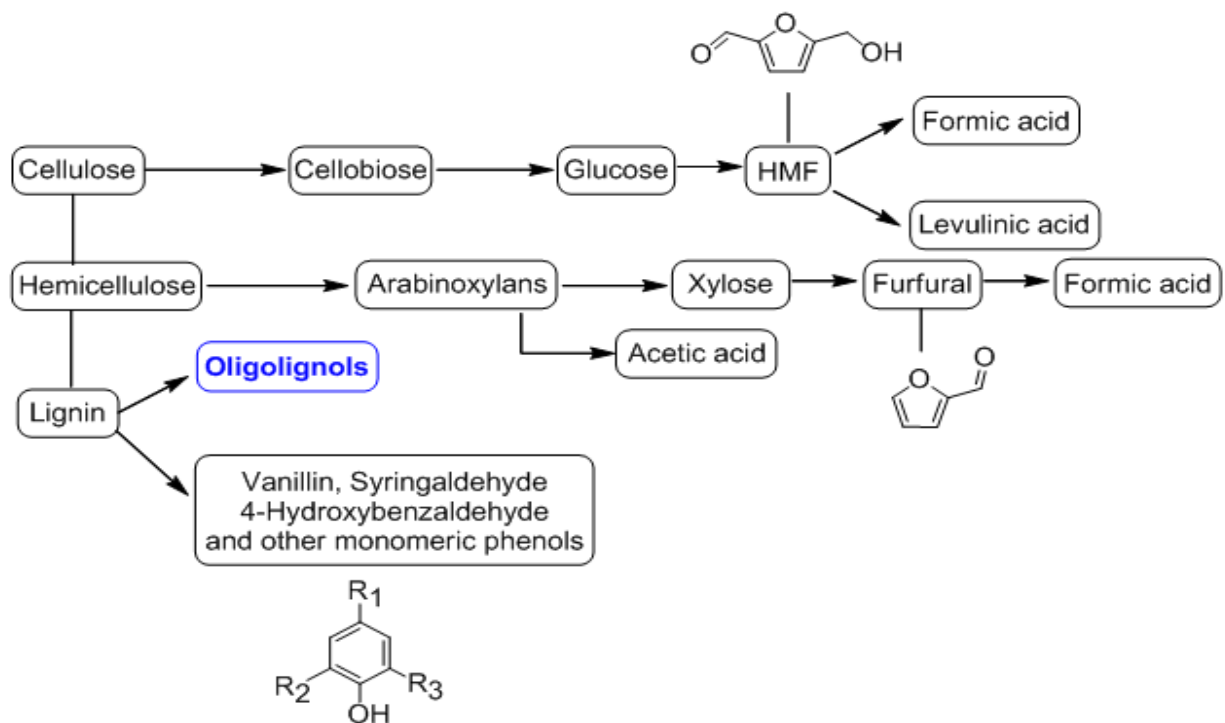
Figure 1.4. Pathways for formation of various degradation products that are generated during processing of lignocellulosic biomass.

1.9 Mass spectrometry in analysis of lignocellulosic biomass hydrolysates: targeted analysis of known degradation products

Attention in the biomass research community has recently turned toward the objective of achieving a new degree of understanding of the degradation products formed during lignocellulosic biomass processing. Analysis of these constituents relies heavily on mass spectrometry mainly due to the complexity of mixtures and the unknown nature of many of these products. Although various targeted analyses using non-mass spectrometry based methods including LC with refractive index detection (LC-RI) or ultraviolet detection (LC-UV) have been used for such purposes, all of these methods focus on known and targeted compounds which are summarized in table 1.1. Mass spectrometry does not require purification of individual

compounds if it is coupled to a chromatographic separation, and use of high-resolution mass spectrometry and CID often leads to prediction of structures of unknown constituents based on accurate molecular and fragment masses. In many cases, an aliquot of process streams can be injected directly into LC/MS for qualitative and quantitative analysis and structure elucidation of unknown compounds.

Table 1.1. Structures of compounds isolated from biomass pretreatment process streams.
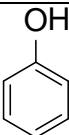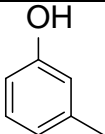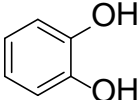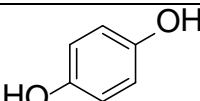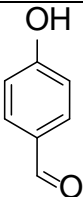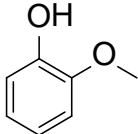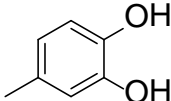
Compound name	Structure	MW	Biomass resource	Reference
Phenolics				
Phenol [a]	OH	94	wheat straw	[33]
2-Methyl phenol (cresol) [a]	OH	108	willow	[68]
1,2 Benzenediol (Catechol) [a]	OH OH	110	willow-spruce	[68,69]
Hydroquinone [a]	OH HO	110	spruce	[69]
4-Hydroxy benzaldehyde [a]	OH O	122	wheat straw-willow	[33,68,70,71]
2-Methoxy phenol (Guaiacol) [a]	OH O	124	wheat straw-willow	[33,68]
Methyl-benzenediol [a]	OH OH	124	willow	[68]

16

Table 1.1 (cont'd)
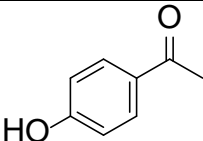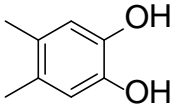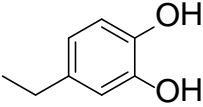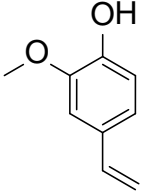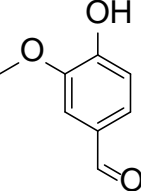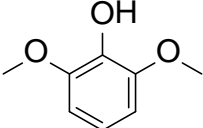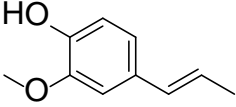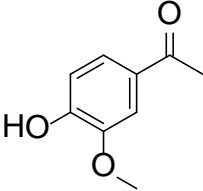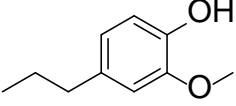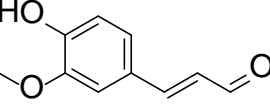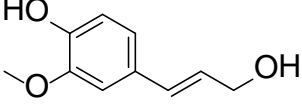
Compound	Structure			
4-Hydroxy acetophenone [a]		136	wheat straw	[33,70,71]
Dimethyl –benzenediol [a]		138	willow	[68]
Ethyl Catechol [a]		138	willow	[68]
2-Methoxy-4-vinyl phenol [a]		150	corn stover	[72]
Vanillin [a]		152	wheat straw- saw dust feedstock -spruce- poplar- corn stover	[33,68-71,73-76]
2,6-Dimethoxyphenol (Syringol) [a]		154	wheat straw- corn stover	[33,72]
2-Methoxy-4-propenylphenol [a]		164	willow	[68]
4-Hydroxy-3-methoxy acetophenone-Acetovanillone (Acetoguaiacone) [a]		166	wheat straw- spruce	[33,69-71,74]
Propylguaiacol [a]		166	willow	[68]
Coniferyl aldehyde [a]		178	spruce	[69,74]
Coniferyl alcohol [a]		180		

Table 1.1 (cont'd)
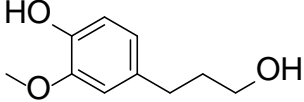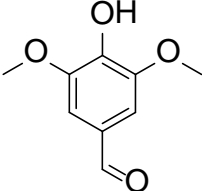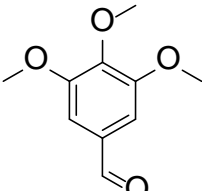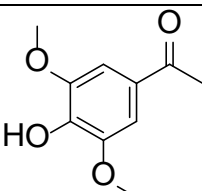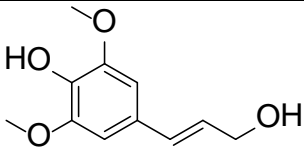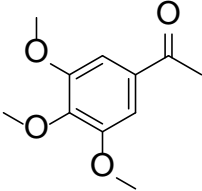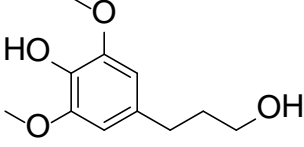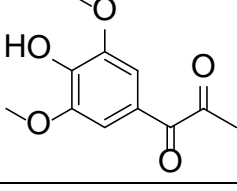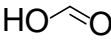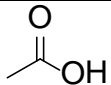
Dihydroconiferyl alcohol [a]		182	red oak wood-spruce	[69,77]
Syringaldehyde [a]		182	Wheat straw-saw dust feedstock-poplar	[33,70,71, 73,75]
3,4,5-Trimethoxy-benzaldehyde [a]		196	wheat straw	[33]
Acetosyringone [a]		196	wheat straw	[33,70,71]
Sinapyl alcohol [a]		210		
3,4,5-Trimethoxy acetophenone [a]		210	wheat straw	[33]
Dihydrosinapyl alcohol [a]		212	red oak wood	[77]
Syringoyl methyl ketone [a]		224	red oak wood	[77]
Carboxylic Acids				
Formic acid [a]		46	wheat straw	[33,78]
Acetic acid [a]		60	red oak – poplar-wheat straw-	[75-78]

18

Table 1.1 (cont'd)
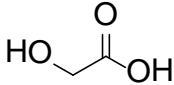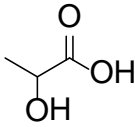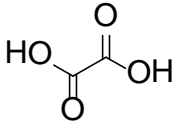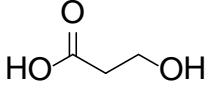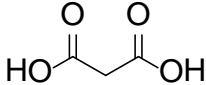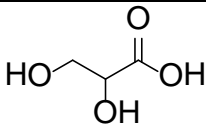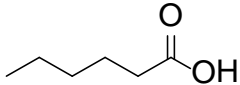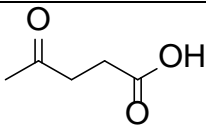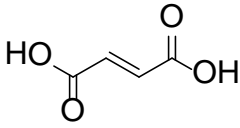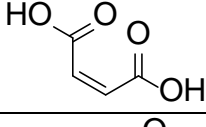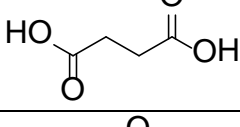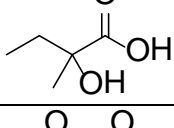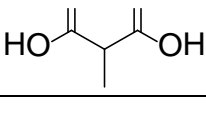
			corn stover	
Hydroxyacetic (glycolic) acid [a]		76		
Lactic acid [a]		90		
Oxalic acid [a]		90		
3-Hydroxy-propanedioic acid [a]		90	poplar	[75]
Propanedioic (malonic) acid [a]		104	corn stover	
2,3-Dihydroxy-propanoic acid [a]		106	poplar	[75]
Caproic (hexanoic) acid [a]		116	red Oak	[77]
4-Oxopentanoic (levulinic) acid [a]		116	poplar	[75]
(E)-Butenedioic (fumaric) acid [b]		116	corn stover	[62]
(Z)-Butenedioic (maleic) acid [b]		116	corn stover	[62]
Succinic acid [a,b]		118	poplar	[75]
2-Methyl-2-hydroxybutanoic acid [a,b]		118	poplar-corn stover	[62,75]
Methylpropanedioic (methyl malonic) acid [a,b]		118	poplar-corn stover	[62,75]

Table 1.1 (cont'd)
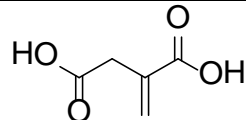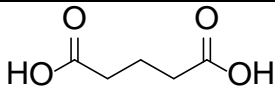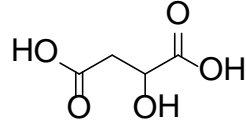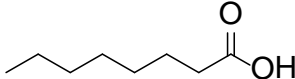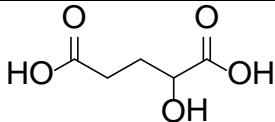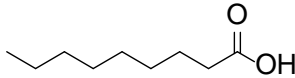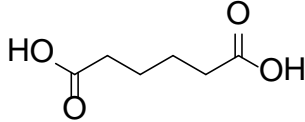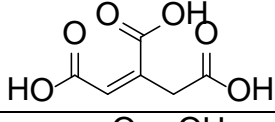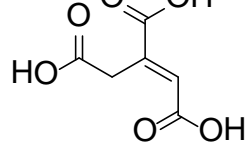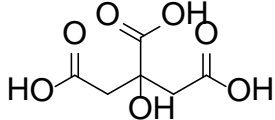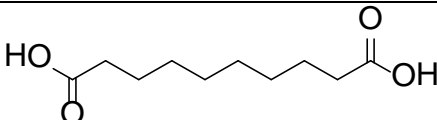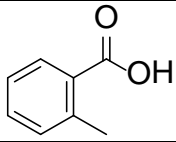
Methylidenebutane-dioic acid (Itaconic) acid [a,b]		130	corn stover	[62]
Pentanedioic (glutaric) acid [b]		132	corn stover	[62]
Malic acid [a]		134	poplar	[75]
Caprilic (Octanoic) acid [a]		144	red Oak Wood	[77]
2-Hydroxypentanedioic acid [a]		148	poplar	[75]
Pelargonic (nonanoic) acid [a]		158	red Oak Wood	[77]
Hexanedioic (adipic)acid [a,b]		146	poplar-corn stover	[62,75]
Cis-aconitic acid [a,b]		174	corn stover	[62]
Trans-aconitic acid [a,b]		174	corn stover	[62]
Citric acid		192		
1,8-Octanedicarboxylic acid [a]		202	poplar	[75]
Phenolic/Aromatic acids				
O-Toluic acid		136		

Table 1.1 (cont'd)
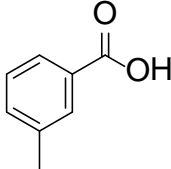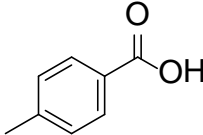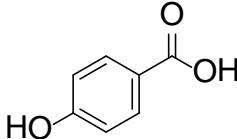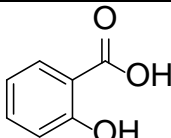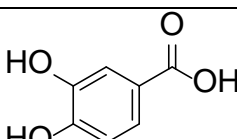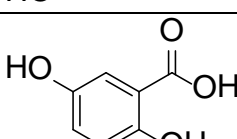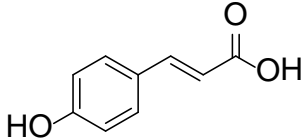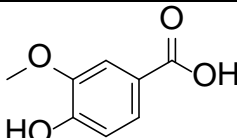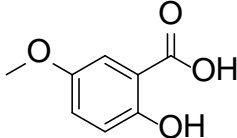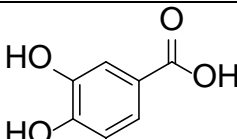
m-Toluic acid		136		
p-Toluic acid		136		
4-Hydroxybenzoic acid [a,b]		138	willow-spruce-wheat straw-poplar-corn stover	[62,68-71,74,75]
2-Hydroxybenzoic (salicylic) acid [b]		138	corn stover	[62]
3,4 Dihydroxybenzoic (Protocatechuic) acid [a,b]		154	saw dust feedstock-willow-poplar	[62,68,73,75]
2,5-Dihydroxybenzoic (gentisic) acid [a,b]		154	willow-poplar	[62,68,75]
Coumaric acid [a,b]		164	corn stover	[62,76]
3-Hydroxy-4-methoxy benzoic acid (Vanillic) acid [a,b]		168	willow-spruce-wheat straw-poplar	[62,68-71,74,75]
2-Hydroxy-5-methoxy benzoic acid [a]		168	poplar	[75]
Gallic acid [a,b]		170	saw dust feedstock-corn stover	[62,73]

Table 1.1 (cont'd)
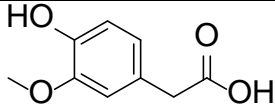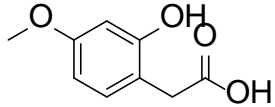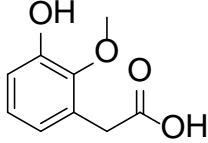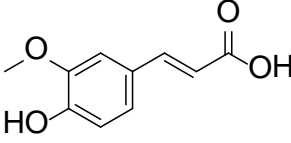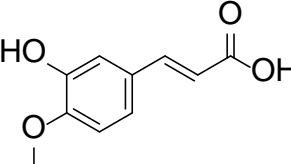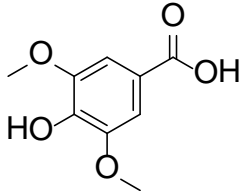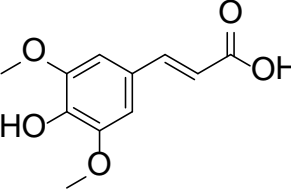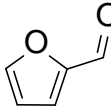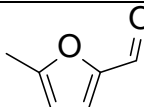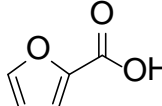
2-(4-Hydroxy-3-methoxyphenol)-acetic acid (homovanillic acid) [a]		182	spruce-poplar	[69,75]
4-Methoxy-α-hydroxy-benzeneacetic acid [a]		182	poplar	[75]
2-Methoxy-α-hydroxy-benzeneacetic acid [a]		182	poplar	[75]
4-Methoxy-3-hydroxy-cinnamic (ferulic) acid [a]		194	poplar-corn stover	[75,76]
3-Methoxy-4-hydroxy-cinnamic (Isoferulic) acid [a]		194	poplar	[75]
4-Hydroxy-3,5-dimethoxybenzoic (Syringic) acid [a]		198	willow-wheat straw-poplar	[68,70,71, 75]
Sinapic acid [b]		224		[62]
Furans				
2-Furfural [a,c]		96	poplar-Saw dust feedstock -wheat straw	[73,75,77, 78]
5-Methyl-2-furfural [a]		110	corn stover	[72]
2-Furoic acid [a]		112	wheat straw	[70,75]

22

Table 1.1 (cont'd)
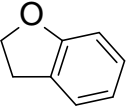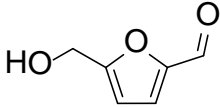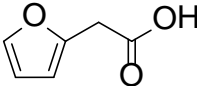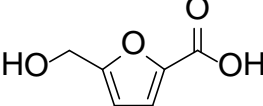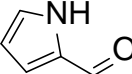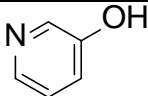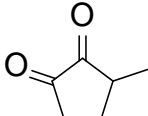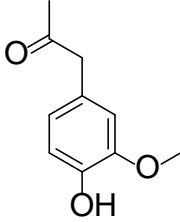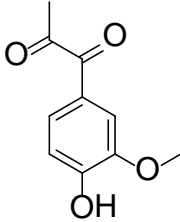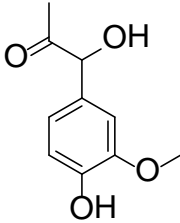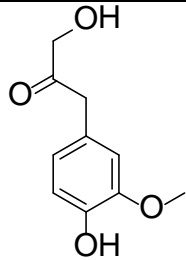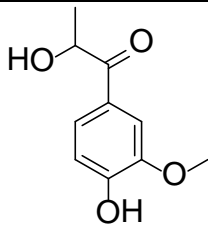
2,3-Dihydrobenzofuran [a]		120	corn stover	[72]
5-Hydroxymethyl-2-furfural (5-HMF) [a]		126	saw dust feedstock -corn stover	[73,76,77]
2-Furanacetic acid [a]		126	poplar	[75]
5-Hydroxy-2- furan carboxylic acid [a]		142	poplar	[75]
Others				
Pyrrole-2-carboxaldehyde [a]		95	corn stover	[72]
3-Hydroxypyridin [a]		95	corn stover	[72]
3-Methyl-1,2-cyclopentanedione [a]		112	corn stover	[72]
Hibbert's Ketones				
1-(4-Hydroxy-3-methoxyphenyl)propan -2-one [a,b]		180	spruce	[69,74]
1-(4-Hydroxy-3-methoxyphenyl)propane-1,2-dione [a,b]		194	spruce	[69,74]
1-Hydroxy-1-(4-hydroxy-3-ethoxyphenyl)propan-2-one [a,b]		196	spruce-poplar	[69,74,75]

Table 1.1 (cont'd)

1-Hydroxy-3-(4-hydroxy-3-methoxyphenyl)propan-2-one [a,b]		196	spruce	[69,74]
2-Hydroxy-1-(4-hydroxy-3-methoxyphenyl)propan-1-one [a,b]		196	spruce-poplar	[74,75]

[a] Measured using GC/MS with or without derivatization

[b] Measured using LC/MS

[c] Measured using APCI-MS

1.10 Identification and quantification of furans

As mentioned above, acid-catalyzed dehydration and condensation of monosaccharides yields oxygen heterocycles, with derivatives of furan being abundant. The primary furans that gained attention as biomass degradation products are furfural and 5-hydroxymethylfurfural (5-HMF) although other acidic furans are also formed during biomass pretreatments. Furfural and 5-HMF exhibit toxicity to fermentation microorganisms including the yeast *Saccharomyces cerevisiae* [67,79-81], ethanologenic *Escherichia coli* [82], *Kluyveromyces marxianus* [36] and *Zymomonas mobilis* CP4 (BZP5) (a recombinant bacterium that can produce ethanol from both xylose and glucose) [73]. Furfural causes a lag-phase in fermentive ethanol formation but does not reduce the final ethanol yield provided that increased time for fermentation is allowed [83]. Removal of furfural increases ethanol productivity in *Saccharomyces cerevisiae* [69]. Kothari *et*

al. showed that furfural and 5-HMF significantly inhibit cellulase and xylanase at concentrations above 5 g/L [63]. So, measuring furfural and 5-HMF is important for assessing the inhibition of fermentation or possible inhibition of oligosaccharide hydrolysis.

Furfural and its derivatives have aldehyde groups that are readily converted to carboxylic acids by oxidation. Examples that have been characterized in biomass hydrolysates include 2-furoic acid, 2-furanacetic acid and 5-hydroxymethyl-2-furancarboxylic acid [75] but their individual effects on fermentation or hydrolysis have not yet been established. Typical furfural derivatives have low molecular weights that make them suitable for GC/MS analysis; however when the aldehyde has been oxidized to a carboxylic acid group, the polar carboxylic acids exhibit poor chromatographic peak shapes, and derivatization such as trimethylsilylation is recommended. Electron ionization generates relatively abundant molecular ions ($M^{\bullet +}$) from furfural and many of its derivatives, and common mass spectrum libraries can be used to aid compound identification. Although GC/MS allows identification and quantification of furans and numerous other degradation products, headspace sampling GC with flame ionization detection (FID) [84] was used for separation and identification of furans including furfural and 5-HMF in biomass hydrolysate. Liquid chromatographic methods including HPLC with UV detection [85] or RI detection [73,86-88] have also been employed, with the latter preferred for measuring compounds that lack UV chromophores. Although LC combined with UV or RI detection has been popular in these applications, the downside to these techniques is the necessity of separation of all degradation products via chromatography. All of these analytes are polar compounds with high aqueous solubility, providing limited opportunities for chromatographic retention and resolution of all compounds. More recently, Davies *et al.* employed headspace sampling and atmospheric pressure chemical ionization (APCI)-MS using a single quadrupole

mass analyzer for direct rapid quantification (2 min/sample) of furfural in wheat straw hydrolysate. In this work, the $[M+H]^+$ ion of furfural (m/z 97) was used for identification and quantification. Although this approach provides for rapid quantitative analysis, in the absence of a physical separation (e.g. chromatography), high-resolution mass spectrometry, or tandem mass spectrometry, the possibility of interference by other substances with the same nominal mass may be significant. Reports of LC/MS methods employing electrospray ionization for detection of furfural and 5-HMF are rare, probably due to the lack of acidic or basic functional groups.

1.11 Identification and quantification of aliphatic carboxylic acids

Hydrolysis of acetyl groups in hemicellulose generates acetic acid, which usually has the highest concentration among carboxylic acids in biomass hydrolysates [4]. Levulinic and formic acids are also present in hydrolysates, and form via degradation of 5-HMF [39,72,89]. Other aliphatic acid degradation products may come either directly from lignin or glycopolymers including cellulose and hemicellulose under acidic and basic conditions of the pretreatment [85]. Under alkaline conditions, carbohydrate rich material [90] including xylans [91] and cellulose [92] will form a variety of carboxylic acids via alkaline peeling reactions that cleave monosaccharides from the reducing end of a glycopolymer. Subsequent keto-enol tautomerizations and retro-aldol reactions form organic acids including lactic acid.

The most common methods for separation and identification of aliphatic carboxylic acids in biomass hydrolysates have employed ion chromatography with conductivity detection [33,69,70,74,93-96] and HPLC-RI [65,73,86-88,97]. As for most identified degradation products in biomass hydrolysates, GC separations coupled to electron ionization mass spectrometry has been the most common approach for analysis of volatile aliphatic carboxylic acids in biomass

hydrolysates. GC/MS has been used with [69,75] and without [76,77,98,99] derivatization for measurement of aliphatic acids in biomass hydrolysates.

A rare but recent effort to quantify small aliphatic acids in biomass hydrolysates using LC/ESI-MS/MS was reported by Sharma *et al.* [100]. In this work, LC separation along with ESI mass spectrometry was used to identify small carboxylic acids using negative ion mode. Eleven small aliphatic acids including malonic, lactic, maleic, *cis*- and *trans*-aconitic, methyl malonic, succinic, fumaric, levulinic, glutaric and adipic acids were measured using LC/MS/MS along with some other degradation products in one analysis. Multiple Reaction Monitoring (MRM) with transitions of $[M-H]^-$ to $[M-H-CO_2]^-$ is used in this work for additional selectivity for detection of aliphatic acids.

Electrospray ionization of aliphatic acids is expected to be efficient in negative ion mode because the carboxylic acids readily release protons. One might expect this to be even more efficient for di- and tri-carboxylic acids including malic and glutaric acids. The challenge lies in chromatographic retention and separation of these polar products using mobile phases that are compatible with ESI. The high polarity and aqueous solubility of these compounds leads to minimal retention on C18 columns unless cationic ion-pairing agents are present in the mobile phase. Only a few reports have described application of reversed phase (RP) LC/MS for characterization of aliphatic acids in biomass hydrolysates [85,100]. However, several papers have shown applicability of LC/MS using hydrophilic interaction chromatography (HILIC) or RP chromatography for measurements of various carboxylic acids in other materials including atmospheric aerosols [101-104] and fermentation broth from antibiotics [105]. In all cases, ESI was used in negative ion mode, and almost exclusively the deprotonated ion $[M-H]^-$ was the dominant ion in the ESI spectra. For larger aliphatic acids (>200 Da), other ion species including

acetate adducts or noncovalent dimer ions including [2M-H]⁻ and [2M-2H+Na]⁻ were also observed [101]. Various types of mass spectrometers were employed including ion trap [101], Q-trap [103,105], time-of-flight (TOF) [102] and quadrupole-time-of-flight (q-TOF) [104] analyzers were successful in these reports, and demonstrated that mass spectrometry based detection of small aliphatic acids can be achieved with most commercial mass spectrometers.

For most aliphatic carboxylic acids, [M-H]⁻ ions are resistant to fragmentation upon collision induced dissociation. This is particularly the case for small aliphatic acids (<150 Da) that lack additional functional groups that promote fragmentation. In one case, a pseudo MRM transition was employed to measure analytes using LC/MS/MS. In this approach, the [M-H]⁻ ion was transmitted by both mass analyzers without formation of a fragment ion [105].

Electrospray ionization of many acids with high aqueous solubility is often inefficient, presumably because these substances do not partition efficiently to the surface of electrosprayed droplets [106]. Oxalic acid was one specific acid reported as difficult to ionize by at least two articles [103,105]. In both cases an AB/Sciex Qtrap 4000 ion source was used for ionization.

Based on various reports of measurements of aliphatic acids using LC/MS, it can be concluded that LC/MS is usually suitable for quantifying aliphatic acids in biomass hydrolysates. Using LC/MS eliminates the need for derivatization. The primary complication in this application lies in the complexity of biomass hydrolysates and the potential for suppression of ionization when other compounds coelute. These issues are particularly acute when chromatographic separations rely on differences in solubility that are minimal, and when formation of compound-selective fragment ions is not feasible.

1.12 Identification and quantification of phenolics and phenolic acids

As mentioned earlier, lignin constitutes a substantial fraction of plant biomass, so the presence of various lignin derived phenolics in biomass hydrolysates is expected. Degradation of lignin and its biosynthetic precursors during pretreatment generates a variety of phenolics in the form of aromatic alcohols, aldehydes, ketones, and acids, with product mixtures depending on pretreatment conditions [33]. The most common phenolics in biomass hydrolysates bear structural similarity to H, G and S monolignols (Figure 1.3). For example the most common phenolic acids characterized in biomass hydrolysates are 4-hydroxybenzoic, vanillic and syringic acids which come from H, G and S monolignols respectively.

Phenolics and phenolic acids have been analyzed in biomass hydrolysates using HPLC with RI [75,107-110] or UV [111-113] detection. Trimethylsilylation (TMS) derivatization with GC/MS analysis has been the primary mass spectrometry-based analysis method for identification of phenolics in biomass hydrolysates [33,69,73,75,79,111]. The principal downside to this approach is the derivatization step. However; this allows for identification of phenolics along with additional degradation products including furans and aliphatic acids in a single analysis. In one case, in the work by Sharma *et al.* that used LC/MS/MS to profile phenolics, a C30 column was used to separate more than 40 degradation products including aliphatic acids, phenolics and phenolic acids. Negative ion mode ESI is often successful for analysis of polyphenolics and phenolic acids because there are several acidic OH or carboxylic acid groups that can release a proton. Sharma *et al.* also reported observation of dominant [M-H]$^-$ ions that later were subjected to CID to generate abundant product ions for multiple reaction monitoring (MRM) quantification. In triple quadrupole mass analyzers, MRM allows for selective detection of analytes based on two ions; a parent ion (usually a pseudomolecular ion) and one fragment that forms upon CID.

One unusual feature of fragments generated from many phenolic substances derives from their propensity to form odd-electron fragment ions. ESI generates even-electron pseudomolecular ions including [M-H]$^-$ and solvent adduct ions including [M+Cl]$^-$, [M+CH$_3$COO]$^-$ and [M+HCOO]$^-$ depending on the compound and ion source and mobile phase conditions. Based on the nitrogen rule [114], organic molecules without nitrogen must have nominal molecular masses that are even integers if they are not free radicals, as is the case for most organic compounds. This rule derives from the unusual property of nitrogen, having an even atomic mass but odd valence. ESI usually forms ions based on attachment or detachment of protons (or other common cations or anions). Compounds that do not contain nitrogen must then form even-electron pseudomolecular ions of monoisotopic m/z that is an odd integer. Collision-induced dissociation of even-electron ions usually proceeds via rearrangements that involve breaking of two bonds, yielding even-electron fragment ions, and as was the case for the pseudomolecular ions, these too will have nominal masses that are odd integers. However, it has been frequently observed that ionized phenolic compounds yield even-mass fragment ions upon CID. These fragment ions must therefore be odd-electron radical ions. In phenolics that contain at least one arylmethoxy group, the most common odd-electron fragments formed upon CID are formed by loss of a methyl radical (CH$_3$$^\bullet$). This was observed for vanillic acid, syringic acid, syringaldehyde and sinapic acid in the report by Sharma *et al.* [100]. Losses of a methyl radical upon CID of phenolics with arylmethoxy groups have been reported before [115-117] and are common in CID spectra of O-methylated flavonoids as well [118,119]. Losses of CO$_2$ (44 Da) also occur in negative CID of phenolic acids, particularly lignin subunits with hydroxyl groups in the 4-position; therefore fragments formed by loss of 44 Da are observed from deprotonated molecular ions of phenolic acids including 4-hydroxybenzoic and *p*-coumaric acids.

Decarboxylation and demethylation may also happen together and so fragments formed by loss of 59 Da will also be common for ferulic acid and similar compounds [100,115,116]. Similar to aliphatic carboxylic acids, phenolics and phenolics acids in biomass hydrolysates have not been analyzed by LC/MS/MS as routinely as LC-RI or GC-MS with TMS derivatization. However, it was shown that LC/MS/MS can be used for measurement of phenolics that have similar structure to those anticipated in biomass hydrolysates [115,116]. ESI in negative mode with MRM transitions with fragments derived from losses of 15, 44 or 59 Da is therefore a promising strategy for analysis of phenolics and phenolic acids in biomass hydrolysates.

1.13 Mass spectrometry in identification of lignin degradation products: compounds larger than monolignols

Because most analyses of lignin degradation have been based upon GC and GC/MS, the range of known compounds has been largely restricted to lignin monomers and slight modifications thereof. Larger oligomers are less volatile, even after derivatization, and authentic reference standards of these substances have been scarce. Characterization of lignin degradation products larger than monolignols requires application of approaches capable of detecting and supporting structure elucidation of larger and less volatile oligolignols. Mass spectrometry has played a crucial role in elucidation of lignin structure since it has provided information essential for identification of monolignol units and for assessing molecular masses of intact lignin molecules using the soft ionization techniques ESI and MALDI [120]. Most research in structural elucidation of lignin with mass spectrometry has been focused on identification of products after thermal (pyrolysis) or chemical (CuO oxidation) degradation, and has been thoroughly reviewed by Reale *et al.* [121].

Combination of pyrolysis with EI mass spectrometry has been a major technique for lignin structure elucidation [121]. Pyrolysis under controlled conditions degrades lignin polymers into volatile thermal dissociation products that can be analyzed by GC-MS. Major degradation products that obtained during lignin pyrolysis are guaiacol, syringol, vanillin, syringaldehyde, phenol, vinyl syringol and vinyl guaiacol [122]. One of the main limitations of GC-MS for characterization of lignin pyrolysis products is that this method works for volatile low molecular weight compounds, but larger degradation products such as oligolignols will be unaccounted for in the analysis [123].

Pyrolysis in combination with soft ionization techniques such as chemical ionization [124], photoionization [125] and fast atom bombardment [126] has extended analysis of lignin degradation products to larger molecules. Van der Hage *et al.* combined pyrolysis on line with LC equipped with a FAB-MS interface which allowed for identification of oligomers that are not amenable to GC/MS analysis including dehydro-sinapyl aldehyde-coniferyl aldehyde and syringaresinol with MW of 386 and 418 respectively [126].

Chemical oxidation with cupric oxide [127,128], nitrobenzene [129] and potassium permanganate [130] combined with GC-MS analysis proved to be a valuable strategy to identify lignin degradation products. These oxidation techniques lead to formation of phenyl aldehydes such as *p*-hydroxybenzaldehyde, vanillin and syringaldehyde, and these oxidation products have been used to classify lignin based on its H, G and S content. Decades ago it was shown that nitrobenzene oxidation of lignin can be used to manufacture vanillin for commercial use [127].

One of the advantages of soft ionization techniques including MALDI and ESI is the ability of these ionization techniques to generate molecular mass information for large nonvolatile molecules, and this knowledge aids characterization of large pieces of lignin

including linkage chemistry between monomeric units. Metzger *et al.* applied MALDI for molecular weight determination of birchwood lignin [131]. The obtained mass spectrum showed a wide range of molecular weights from several hundred to 16000 Da. The center of mass of the peak was about 2600 Da which agrees with the average MW of birchwood lignin determined by gel-permeation chromatography. MALDI has also been applied to polymers similar to lignin produced from polymerization of lignin monomeric units such as coniferyl alcohol. De Angelis et al., took MALDI mass spectra of a reaction mixture of coniferyl alcohol polymerized by horseradish H_2O_2, and showed polymers ranging from 200-1800 Da with mass increments of 178 Da indicating an oxidative radical mechanism coupling the monomers [132]. The same group repeated this work three years later by using ESI to show the applicability of this soft ionization technique for characterization of large lignin molecules [133]. ESI in negative ion mode was also employed by Evtuguin *et al.* to determine the molecular masses of lignins from *Eucalyptus globules* wood [134]. Molecular weights of several detected oligomers such as 811,841; 1069,1099; 1085,1015 and 1270,1300 differed only by the mass difference between a hydrogen and a methoxy group, suggesting presence of structures having different proportions of guaiacyl and syringyl units.

Although pyrolysis and oxidative degradation of lignin coupled with GC/MS analysis have played important roles in characterization of lignin, but GC/MS is usually limited to analysis of substances < 1000 Da. In contrast, the soft ionization MALDI and ESI have successfully ionized large pieces of lignin polymers (up to 16000 MW) to provide molecular mass information.

1.14 Soft ionization coupled to fragmentation for characterization of lignin degradation products

Coupling CID to soft ionization mass spectrometry generates fragment ions whose masses

provide structural information. This powerful tool has aided characterization of lignin degradation products in several recent reports. Time-of-flight secondary ion mass spectrometry (TOF-SIMS) was successfully applied to characterize common lignin interunit linkages in 2005 [135]. Saito *et al.* showed that in positive mode, ions with *m/z* 137 and 151 were prominent fragment ions from guaiacyl (G) units, and ions with *m/z* 167 and 181 are prominent fragment ions from syringyl (S) units in lignin. They generated TOF-SIMS spectra of pine and beech milled wood lignin and observed *m/z* 151 and 137 in spectra from pine lignin, and relative abundances of ions at *m/z* 137, 151, 167 and 181 from beech lignin confirmed the higher ratio of S/G unit in beech lignin relative to pine. Tentative structures of these ions were also presented (Figure 1.5) and these assignments were supported by comparisons to mass spectra from synthetic lignin polymer with deuterium labeled coniferyl alcohol. This group repeated experiments on lignin model compounds synthesized with 8—O—4, 8—1, 8—5 and 5—5 linkages. These designations indicate the positions on the phenylpropanoid unit that are involved in the linkage, and the designation of "O" illustrates that the linkage occurs through an ether rather than a carbon-carbon bond. It turned out that the ions with *m/z* 137 and 151 (illustrated in Figure 1.5) result from rupture of most common interunit linkages except 5—5 and not just 8—O—4 which is the most abundant linkage type in lignin [136], thus demonstrating that SIMS generates aromatic substitution-characteristic fragment ions from various linkage types. Although this work is useful for assessing lignin types based on S/G ratios, the results are not helpful to assign linkage types since no characteristic fragment ions were observed that could distinguish different linkages.

Figure 1.5. Characteristic ions of S and G monolignols generated in TOF-SIMS mass spectra of intact lignin in positive-ion mode [135].

In a recent related report, Morreel *et al.* performed MS/MS experiments on a series of model oligolignols varying in linkages between the two phenyl propanoids including 8—O—4, 8—5 and 8—8 [117]. Model compounds were composed of various H, G and S monomeric units. Some of the model structures are shown in Figure 1.6. CID was performed on pseudomolecular ions from each model compound to draw conclusions about characteristic fragments for each linkage type. For example, neutral losses of 18 Da (water), 30 Da (formaldehyde), and 48 Da (water + formaldehyde) were characteristic product ions of 8—O—4 linkages for both β-aryl ethers and benzodioxanes (Figure 1.6). But the latter product ion (loss of 48) which is usually the base peak in β-aryl ethers was barely detectable for benzodioxanes. Losses of water, formaldehyde and methyl radical (for methyoxylated units) varied in relative

abundance for different linkage types. Prominent fragments in some cases helped to distinguish the aromatic units via cleavage of the ether bonds. These fragmentation schemes laid the foundation for annotation of both linkage type and phenolic units in oligolignols.

Figure 1.6. Model compounds used by Morreel *et al.* to establish characteristic fragmentation of various inter-unit linkages in lignin. R groups are either H, CH_2, CH_2OH, CO or OCH_3.

The findings led to characterization of 134 lignin trimers to hexamers from poplar xylem, of which 36 were completely sequenced [137]. Own et al. also characterized several lignin degradation products based on MS/MS characteristic fragments based on model compounds including: guaiacol, vanillin, coniferyl and sinapyl alcohol, and other phenolics which bear structural similarity [138]. ESI was performed by doping the solvent with NaOH to generate deprotonated molecular ions without in-source fragmentation, and high-resolution multi-stage

tandem mass spectrometry was performed using a Fourier transform-ion cyclotron resonance (FT-ICR) mass analyzer to annotate and confirm assignments of fragment ions generated upon CID. Molecular structures for some dilignols from a mixture of red oak chip lignin were proposed. Doping ESI with NaOH and accurate mass measurements of fragment ions were useful in assigning structures of unknown degradation products.

Electrospray ionization is one of several approaches to ionization that produce molecular adducts with various ions. In some cases, fragment ions form in the ion source via loss of water. When this process is extensive, it becomes challenging to make correct assignments of analyte molecular masses. Haupert *et al.* showed that small lignin degradation products including guaiacol, vanillyl and coniferyl alcohol do not form abundant [M-H]$^-$ ions using atmospheric pressure chemical ionization (APCI) or ESI in negative mode, and in some cases losses of water from pseudomolecular ions occurred in the ion source in positive mode [139]. However, this problem was solved through doping samples with NaCl and NaOH for positive and negative ESI respectively. [M+Na]$^+$ was observed as the base peak with minimal fragmentation, and [M-H]$^-$ ions dominated the mass spectra in negative mode. Since sodium adduct ions have high kinetic barriers to fragmentation relative to many [M-H]$^-$ ions, it was suggested that for CID on lignin degradation products, it is preferable to dope extracts with NaOH, generate deprotonated molecular ions almost exclusively, and perform CID on [M-H]$^-$ ions.

Most recently, Kiyota *et al.* incubated individual lignin monomers H, G and S with peroxidase enzymes to synthesize oligolignols up to tetramers [140]. Structures of more than 15 of these synthesized oligolignols were elucidated using MS/MS spectra in ESI negative mode based on the characteristic fragmentation proposed by Morreel *et al.* [117]. Ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC/MS/MS) analysis was

performed on lignin extracted from sugarcane. Based on the fragmentation of model compounds, several tetramers containing S and G units mainly with 8—O—4, 8—5 and 8—8 linkages were characterized in the lignin extract of sugarcane. This work showed the applicability of UHPLC/MS/MS in determining the ratio of S and G monolignols which is useful for overall lignin recalcitrance of biomass.

Over the last decade, more researchers in the biomass community have explored mass spectrometry behavior of larger lignin degradation products in greater depth [117,135-138]. Their findings regarding the behavior of lignin derivatives in the mass spectrometer promise to guide the analysis of degradation products in biomass hydrolysates. The combination of ultra-high performance separations with mass spectrometry analyses offers great promise that comprehensive identification of biomass hydrolysate constituents will be achieved within a few years.

1.15 Non-targeted metabolomics approaches for profiling of biomass degradation products

The global measurement of metabolites, termed 'metabolomics', aims to provide a global assessment of the entire suite of metabolites produced by a biological system [141,142]. Parallel application of several analytical platforms including MS and NMR is required for comprehensive identification of metabolites, and additional tools including x-ray crystallography may be needed to define absolute stereochemistry. There are two different approaches for metabolomic studies: targeted and non-targeted approaches [143]. Targeted approaches usually focus on measuring a limited number of known metabolites, and their qualitative and quantitative assessment is achievable using mass spectrometry platforms including LC/MS or GC/MS. In both techniques the predicted *m/z* value of each metabolite is used for selective detection in a complex sample

mixture. Yet more selective measurement of pre-selected metabolites is achieved using LC/MS/MS, which provides additional selectivity by requiring that the molecule produce both ions corresponding to the intact molecule and a compound-characteristic fragment ion that forms upon CID. LC/MS/MS also extends the range of compounds that can be quantified to those of high (up to ~ 2000 Da) molecular mass or polar metabolites that are not amenable to GC/MS analysis. In contrast, GC/MS provides more efficient separation and access to libraries of EI mass spectra that can facilitate identification of previously unknown substances.

Non-targeted metabolomics, often called metabolic fingerprinting, aims to provide global measurements of all metabolites. This unbiased global approach classifies samples based on metabolite patterns or fingerprints that may change in response to a treatment [143]. NMR and MS have played a central role in metabolite fingerprinting. One of the notable differences between targeted and non-targeted metabolomics approaches is the amount of data that is generated in the latter. To visualize such high-dimensional data sets, multivariate statistical analysis methods including principal components analysis (PCA) is required to reduce the dimensionality so that most of the variance in metabolite abundances can be represented in two or three dimensions. Supervised statistical methods based on discriminant analysis including partial least squares (PLS-DA) and orthogonal projections to latent structures (OPLS-DA) aid recognition of those metabolites that distinguish sample classes. Thereafter, metabolite identification is performed by application of high-resolution mass spectrometry, tandem mass spectrometry and NMR.

The challenge of analyzing degradation products in biomass hydrolysates is perhaps best approached using techniques and strategies similar to those used for metabolomics. Although post-harvest biomass may be viewed as a dead biological system, the levels of various plant

metabolites in the biomass serve as "chemical fossils" that reflect its chemical history. Therefore, the global profiling of metabolites and their degradation in processed biomass can be viewed as a "post-mortem" metabolomic approach. In this light, non-targeted metabolomic strategies can be used to differentiate treated and untreated biomass extracts or characterize differences in chemical profiles governed by various treatment severities [62]. When coupled with measurements of downstream yields of conversion to desirable products such as bioethanol, candidate "biomarkers" are identified that may serve as surrogates or mechanistic indicators of treatment efficiency.

Our limited understanding of the identities of substances released from biomass during treatment presents a barrier to mechanistic understanding and optimization of such treatments. In principal, all elemental formulas have unique molecular masses, but mass spectrometry technologies do not yet provide the precision and accuracy needed to generate unambiguous formulas from molecular mass alone [144]. This results from the exponential relationship between molecular mass and the number of possible formulas that lie within experimental mass measurement error. One useful piece of information that can be extracted from high resolution mass spectra is termed the relative mass defect (RMD) [145]. Values of RMD are calculated from the mass defect (digits that follow the decimal place) normalized to the measured ion mass, usually expressed in parts-per-million (ppm). Values of RMD show a strong correlation with the percent hydrogen in the molecule by weight, and serves as a measure of the reduced state of a molecular or fragment ion. For example an ion observed at *m/z* 382.1059 has RMD of 277 ppm as calculated below:

$$[(382.1059\text{-}382.0000)/(382.1059)] \times 10^{6} = 277 \text{ ppm}$$

and this relatively low value suggests a hydrogen-depleted substance. For comparison, lipids

have high RMD values (600-900 ppm) because of high hydrogen content and sugars (300-400 ppm) and phenolics (250-300 ppm) have relatively lower RMD values [62]. Since the molecular oxidation state reflects the combined contributions of the molecule's biosynthetic precursors and subsequent metabolic or degradation reactions, RMD values calculated from MS data allow unknown substances to be categorized based on their oxidation state. Characterization of biomass degradation products based in part on their RMD values will be discussed in Chapter Two.

By reviewing published work in "characterization of degradation products in biomass hydrolysates"; it can be concluded that the primary focus of biomass community has been centered on 60-70 compounds (Table 1.1) with molecular weights of less than 250 Da. This list of compounds serves as the defined targets of metabolomics analysis. Although informative and useful in many cases, a significant number of products still remained unidentified and have yet to enter into discussions of biomass processing.

To address the extent to which the composition of biomass treatment process stream remains unknown, the aims of the research described in this dissertation have been centered upon the development and application of mass spectrometry-based methods for comprehensive profiling. Chapter 2 presents the description of a multifaceted analytical approach for profiling substances extracted from corn stover and mushroom spent straw before and after pretreatments. This work has already been published in two papers (Balan, V., L. de Sousa, S. P. S. Chundawat, R. Vismeh, A. D. Jones, and B. E. Dale. Mushroom spent straw: a potential substrate for an ethanol-based biorefinery. *J. Indust. Microbiol. Biotechnol. **2008**, **35**: 293-301; S. P. S. Chundawat, R. Vismeh, L. N. Sharma, L. da Costa Sousa, C. K. Chambliss, A. D. Jones, V. Balan, and B. E. Dale. Multifaceted characterization of cell wall decomposition products formed

during ammonia fiber expansion (AFEX) and dilute-acid based pretreatments, *Bioresource Technol.*, *2010*, **101**: 8429-8438). In Chapter 3, LC/MS/MS methodologies have been developed to distinguish isomeric diferulates, which are present in hemicellulose as crosslinks between oligosaccharide chains. This work has been submitted for publication in Analyst, Vismeh, R., Lu, F., Chundawat, S. P. S., Azarpira, A., Ralph, J., Balan. V., Dale, B. E., Jones, A. D, Profiling of Diferulates (Plant Cell Wall Cross-Linkers) using Ultrahigh-performance Liquid Chromatography-Tandem Mass Spectrometry, *Analyst.,* **2012**.

In Chapter 4, an original approach for LC/time-of-flight MS profiling of oligosaccharides is presented (and is under review by *Carbohydrate Polymers,* Vismeh, R., Humpula, J., Chundawat, S. P. S., J., Balan. V., Dale, B. E., Jones, A. D., Profiling of Soluble Neutral Oligosaccharides from Treated Biomass using Solid Phase Extraction and Liquid Chromatography-Multiplexed Collision Induced Dissociation-Mass Spectrometry, *Carbohydr. Polym.* **2012**)

REFERENCES

REFERENCES

(1)   Kuchler, M.; Linner, B. O. *Food Policy* **2012**, *37*, 581-588.

(2)   Tenenbaum, D. J. *Environmental Health Perspectives* **2008**, *116*, A254-A257.

(3)   Valentine, J.; Clifton-Brown, J.; Hastings, A.; Robson, P.; Allison, G.; Smith, P. *Global Change Biology Bioenergy* **2012**, *4*, 1-19.

(4)   Klinke, H. B.; Thomsen, A. B.; Ahring, B. K. *Applied Microbiology and Biotechnology* **2004**, *66*, 10-26.

(5)   Limayem, A.; Ricke, S. C. *Progress in Energy and Combustion Science* **2012**, *38*, 449-467.

(6)   Lowrey, A.; Nixon, R. *The New York Times* **2012**, *25*, 91-98.

(7)   Chisti, Y. *Biotechnology Advances* **2007**, *25*, 294-306.

(8)   Fukuda, H.; Kondo, A.; Noda, H. *Journal of Bioscience and Bioengineering* **2001**, *92*, 405-416.

(9)   Hill, J.; Nelson, E.; Tilman, D.; Polasky, S.; Tiffany, D. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 11206-11210.

(10)   Ma, F. R.; Hanna, M. A. *Bioresource Technology* **1999**, *70*, 1-15.

(11)   Meher, L. C.; Sagar, D. V.; Naik, S. N. *Renewable & Sustainable Energy Reviews* **2006**, *10*, 248-268.

(12)   Balat, M.; Balat, H.; Oz, C. *Progress in Energy and Combustion Science* **2008**, *34*, 551-573.

(13)   Gray, K. A.; Zhao, L. S.; Emptage, M. *Current Opinion in Chemical Biology* **2006**, *10*, 141-146.

(14)   Kim, S.; Dale, B. E. *Biomass & Bioenergy* **2004**, *26*, 361-375.

(15)   Hendriks, A.; Zeeman, G. *Bioresource Technology* **2009**, *100*, 10-18.

(16)   Wooley, R.; Ruth, M.; Glassner, D.; Sheehan, J. *Biotechnology Progress* **1999**, *15*, 794-803.

(17)   Mosier, N.; Wyman, C.; Dale, B.; Elander, R.; Lee, Y. Y.; Holtzapple, M.; Ladisch, M. *Bioresource Technology* **2005**, *96*, 673-686.

(18)   Pauly, M.; Keegstra, K. *Plant Journal* **2008**, *54*, 559-568.

(19)   Saha, B. C. *Journal of Industrial Microbiology & Biotechnology* **2003**, *30*, 279-291.

(20)   Wyman, C. E.; Dale, B. E.; Elander, R. T.; Holtzapple, M.; Ladisch, M. R.; Lee, Y. Y. *Bioresource Technology* **2005**, *96*, 1959-1966.

(21)   Yang, B.; Wyman, C. E. *Biofuels Bioproducts & Biorefining-Biofpr* **2008**, *2*, 26-40.

(22)   Rivers, D. B.; Emert, G. H. *Biotechnology Letters* **1987**, *9*, 365-368.

(23)   Vidal, B. C.; Dien, B. S.; Ting, K. C.; Singh, V. *Applied Biochemistry and Biotechnology* **2011**, *164*, 1405-1421.

(24)   Tassinari, T. H.; Macy, C. F.; Spano, L. A. *Biotechnology and Bioengineering* **1982**, *24*, 1495-1505.

(25)   Brownell, H. H.; Saddler, J. N. *Biotechnology and Bioengineering* **1987**, *29*, 228-235.

(26)   Bobleter, O. *Progress in Polymer Science* **1994**, *19*, 797-841.

(27)   Esteghlalian, A.; Hashimoto, A. G.; Fenske, J. J.; Penner, M. H. *Bioresource Technology* **1997**, *59*, 129-136.

(28)   Nguyen, Q. A.; Tucker, M. P.; Keller, F. A.; Eddy, F. P. *Applied Biochemistry and Biotechnology* **2000**, *84-6*, 561-576.

(29)   Sierra, R.; Granda, C. B.; Holtzapple, M. T. In *Biofuels: Methods and Protocols*; Mielenz, J. R., Ed. 2009; Vol. 581, p 115-124.

(30)   Teymouri, F.; Laureano-Perez, L.; Alizadeh, H.; Dale, B. E. *Bioresource Technology* **2005**, *96*, 2014-2018.

(31)   Holtzapple, M. T.; Jun, J. H.; Ashok, G.; Patibandla, S. L.; Dale, B. E. *Applied Biochemistry and Biotechnology* **1991**, *28-9*, 59-74.

(32)   Kilpelainen, I.; Xie, H.; King, A.; Granstrom, M.; Heikkinen, S.; Argyropoulos, D. S. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 9142-9148.

(33)   Klinke, H. B.; Ahring, B. K.; Schmidt, A. S.; Thomsen, A. B. *Bioresource Technology* **2002**, *82*, 15-26.

(34)   Sanchez, B.; Bautista, J. *Enzyme and Microbial Technology* **1988**, *10*, 315-318.

(35)   Delgenes, J. P.; Moletta, R.; Navarro, J. M. *Enzyme and Microbial Technology* **1996**, *19*, 220-225.

(36)   Oliva, J. M.; Negro, M. J.; Saez, F.; Ballesteros, I.; Manzanares, P.; Gonzalez, A.; Ballesteros, M. *Process Biochemistry* **2006**, *41*, 1223-1228.

(37)   Russell, J. B. *Journal of Applied Bacteriology* **1992**, *73*, 363-370.

(38)   Axe, D. D.; Bailey, J. E. *Biotechnology and Bioengineering* **1995**, *47*, 8-19.

(39)   Palmqvist, E.; Hahn-Hagerdal, B. *Bioresource Technology* **2000**, *74*, 25-33.

(40)   Cantarella, M.; Cantarella, L.; Gallifuoco, A.; Spera, A.; Alfani, F. *Biotechnology Progress* **2004**, *20*, 200-206.

(41)   Hodge, D. B.; Karim, M. N.; Schell, D. J.; McMillan, J. D. *Bioresource Technology* **2008**, *99*, 8940-8948.

(42)   Jing, X. Y.; Zhang, X. X.; Bao, J. *Applied Biochemistry and Biotechnology* **2009**, *159*, 696-707.

(43)   Hong, J.; Ladisch, M. R.; Gong, C. S.; Wankat, P. C.; Tsao, G. T. *Biotechnology and Bioengineering* **1981**, *23*, 2779-2788.

(44)   Alfani, F.; Cantarella, L.; Gallifuoco, A.; Cantarella, M. *Journal of Membrane Science* **1990**, *52*, 339-350.

(45)   Philippidis, G. P.; Smith, T. K.; Wyman, C. E. *Biotechnology and Bioengineering* **1993**, *41*, 846-853.

(46)   Kabel, M. A.; Schols, H. A.; Voragen, A. G. J. *Carbohydrate Polymers* **2002**, *50*, 191-200.

(47)   Appeldoorn, M. M.; Kabel, M. A.; Van Eylen, D.; Gruppen, H.; Schols, H. A. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 11294-11301.

(48)   Harrison, S.; Fraser, K.; Lane, G.; Hughes, D.; Villas-Boas, S.; Rasmussen, S. *Analytical and Bioanalytical Chemistry* **2011**, *401*, 2955-2963.

(49)   Van Dongen, F. E. M.; Van Eylen, D.; Kabel, M. A. *Carbohydrate Polymers* **2011**, *86*, 722-731.

(50)   Lundqvist, J.; Teleman, A.; Junel, L.; Zacchi, G.; Dahlman, O.; Tjerneld, F.; Stalbrand, H. *Carbohydrate Polymers* **2002**, *48*, 29-39.

(51)   Kabel, M. A.; Carvalheiro, F.; Garrote, G.; Avgerinos, E.; Koukios, E.; Parajo, J. C.; Girio, F. M.; Schols, H. A.; Voragen, A. G. J. *Carbohydrate Polymers* **2002**, *50*, 47-56.

(52)   Teleman, A.; Nordstrom, M.; Tenkanen, M.; Jacobs, A.; Dahlman, O. *Carbohydrate Research* **2003**, *338*, 525-534.

(53)   Boerjan, W.; Ralph, J.; Baucher, M. *Annual Review of Plant Biology* **2003**, *54*, 519-546.

(54)   Freudenb.K *Science* **1965**, *148*, 595-&.

(55)   Freudenberg, K. *Nature* **1959**, *183*, 1152-1155.

(56)   Adler, E. *Wood Science and Technology* **1977**, *11*, 169-218.

(57)   Nimz, H. *Angewandte Chemie-International Edition in English* **1974**, *13*, 313-321.

(58)   Sakakibara, A. *Wood Science and Technology* **1980**, *14*, 89-100.

(59)   Freudenberg, K.; Neish, A. In *molecular biology, biochemistry, and biophysiscs*; Springer-Verlag.: New York, 1968.

(60)   Ramiah, M. V. *Journal of Applied Polymer Science* **1970**, *14*, 1323-&.

(61)   Yang, Q.; Wu, S. B.; Lou, R.; Lv, G. J. *Journal of Analytical and Applied Pyrolysis* **2010**, *87*, 291-291.

(62)   Chundawat, S. P. S.; Vismeh, R.; Sharma, L. N.; Humpula, J. F.; Sousa, L. D.; Chambliss, C. K.; Jones, A. D.; Balan, V.; Dale, B. E. *Bioresource Technology* **2010**, *101*, 8429-8438.

(63)   Kothari, U. D.; Lee, Y. Y. *Applied Biochemistry and Biotechnology* **2011**, *165*, 1391-1405.

(64)   Larsson, S.; Palmqvist, E.; Hahn-Hagerdal, B.; Tengborg, C.; Stenberg, K.; Zacchi, G.; Nilvebrant, N. O. *Enzyme and Microbial Technology* **1999**, *24*, 151-159.

(65)   Taherzadeh, M. J.; Eklund, R.; Gustafsson, L.; Niklasson, C.; Liden, G. *Industrial & Engineering Chemistry Research* **1997**, *36*, 4659-4665.

(66)   Almeida, J. R. M.; Bertilsson, M.; Gorwa-Grauslund, M. F.; Gorsich, S.; Liden, G. *Applied Microbiology and Biotechnology* **2009**, *82*, 625-638.

(67)   Boyer, L. J.; Vega, J. L.; Klasson, K. T.; Clausen, E. C.; Gaddy, J. L. *Biomass & Bioenergy* **1992**, *3*, 41-48.

(68)   Jonsson, L. J.; Palmqvist, E.; Nilvebrant, N. O.; Hahn-Hagerdal, B. *Applied Microbiology and Biotechnology* **1998**, *49*, 691-697.

(69)   Larsson, S.; Reimann, A.; Nilvebrant, N. O.; Jonsson, L. J. *Applied Biochemistry and Biotechnology* **1999**, *77-9*, 91-103.

(70)   Klinke, H. B.; Thomsen, A. B.; Ahring, B. K. *Applied Microbiology and Biotechnology* **2001**, *57*, 631-638.

(71)   Klinke, H. B.; Olsson, L.; Thomsen, A. B.; Ahring, B. K. *Biotechnology and Bioengineering* **2003**, *81*, 738-747.

(72)   Zheng, R. P.; Zhang, H. M.; Zhao, J.; Lei, M. L.; Huang, H. *Journal of Chromatography A* **2011**, *1218*, 5319-5327.

(73)   Ranatunga, T. D.; Jervis, J.; Helm, R. F.; McMillan, J. D.; Hatzis, C. *Applied Biochemistry and Biotechnology* **1997**, *67*, 185-198.

(74)   Persson, P.; Larsson, S.; Jonsson, L. J.; Nilvebrant, N. O.; Sivik, B.; Munteanu, F.; Thorneby, L.; Gorton, L. *Biotechnology and Bioengineering* **2002**, *79*, 694-700.

(75)   Luo, C. D.; Brink, D. L.; Blanch, H. W. *Biomass & Bioenergy* **2002**, *22*, 125-138.

(76)   Humpula, J. F.; Chundawat, S. P. S.; Vismeh, R.; Jones, A. D.; Balan, V.; Dale, B. E. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2011**, *879*, 1018-1022.

(77)   Tran, A. V.; Chambers, R. P. *Biotechnology Letters* **1985**, *7*, 841-845.

(78)   Davies, S. M.; Linforth, R. S.; Wilkinson, S. J.; Smart, K. A.; Cook, D. J. *Biotechnology for Biofuels* **2011**, *4*.

(79)   Clark, T. A.; Mackie, K. L. *Journal of Chemical Technology and Biotechnology B-Biotechnology* **1984**, *34*, 101-110.

(80)   Palmqvist, E.; Almeida, J. S.; Hahn-Hagerdal, B. *Biotechnology and Bioengineering* **1999**, *62*, 447-454.

(81)   Palmqvist, E.; Grage, H.; Meinander, N. Q.; Hahn-Hagerdal, B. *Biotechnology and Bioengineering* **1999**, *63*, 46-55.

(82)   Zaldivar, J.; Martinez, A.; Ingram, L. O. *Biotechnology and Bioengineering* **2000**, *68*, 524-530.

(83)   Chung, I. S.; Lee, Y. Y. *Biotechnology and Bioengineering* **1985**, *27*, 308-315.

(84)   Li, H.; Chai, X.-S.; Zhan, H.; Fu, S. *Journal of Chromatography A* **2010**, *1217*, 7616-7619.

(85)   Du, B. W.; Sharma, L. N.; Becker, C.; Chen, S. F.; Mowery, R. A.; van Walsum, G. P.; Chambliss, C. K. *Biotechnology and Bioengineering* **2010**, *107*, 430-440.

(86)   Scarlata, C. J.; Hyman, D. A. *Journal of Chromatography A* **2010**, *1217*, 2082-2087.

(87)   Xie, R.; Tu, M. B.; Wu, Y. N.; Adhikari, S. *Bioresource Technology* **2011**, *102*, 4938-4942.

(88)   Gurram, R. N.; Datta, S.; Lin, Y. J.; Snyder, S. W.; Menkhaus, T. J. *Bioresource Technology* **2011**, *102*, 7850-7859.

(89)   Dunlop, A. P. *Industrial and Engineering Chemistry* **1948**, *40*, 204-209.

(90)   Sjostrom, E. *Biomass & Bioenergy* **1991**, *1*, 61-64.

(91)   Niemelä, K. *Carbohydrate Research* **1990**, *204*, 37-49.

(92)   Knill, C. J.; Kennedy, J. F. *Carbohydrate Polymers* **2003**, *51*, 281-300.

(93)   Persson, P.; Andersson, J.; Gorton, L.; Larsson, S.; Nilvebrant, N. O.; Jonsson, L. J. *Journal of Agricultural and Food Chemistry* **2002**, *50*, 5318-5325.

(94)   Bonn, G.; Oefner, P. J.; Bobleter, O. *Fresenius Zeitschrift Fur Analytische Chemie* **1988**, *331*, 46-50.

(95)   Chen, S. F.; Mowery, R. A.; Scarlata, C. J.; Chambliss, C. K. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 5912-5918.

(96)   Chen, S. F.; Mowery, R. A.; Sevcik, R. S.; Scarlata, C. J.; Chambliss, C. K. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 3251-3258.

(97)   Huang, H. Z.; Guo, X. Y.; Li, D. M.; Liu, M. M.; Wu, J. F.; Ren, H. Y. *Bioresource Technology* **2011**, *102*, 7486-7493.

(98)   Chundawat, S. P. S.; Donohoe, B. S.; Sousa, L. D.; Elder, T.; Agarwal, U. P.; Lu, F. C.; Ralph, J.; Himmel, M. E.; Balan, V.; Dale, B. E. *Energy & Environmental Science* **2011**, *4*, 973-984.

(99)   Thomsen, M. H.; Thygesen, A.; Thomsen, A. B. *Applied Microbiology and Biotechnology* **2009**, *83*, 447-455.

(100)   Sharma, L. N.; Becker, C.; Chambliss, C. K. In *Biofuels: Methods and Protocols*; Mielenz, J. R., Ed. 2009; Vol. 581, p 125-143.

(101)   Anttila, P.; Hyotylainen, T.; Heikkila, A.; Jussila, M.; Finell, J.; Kulmala, M.; Riekkola, M. L. *Journal of Separation Science* **2005**, *28*, 337-346.

(102)   Mirivel, G.; Riffault, V.; Galloo, J. C. *Analytical Methods* **2011**, *3*, 1172-1179.

(103)   Kitanovski, Z.; Grgic, I.; Veber, M. *Journal of Chromatography A* **2011**, *1218*, 4417-4425.

(104)   Zhang, Y. Y.; Muller, L.; Winterhalter, R.; Moortgat, G. K.; Hoffmann, T.; Poschl, U. *Atmospheric Chemistry and Physics* **2010**, *10*, 7859-7873.

(105)   Preinerstorfer, B.; Schiesel, S.; Lammerhofer, M. L.; Lindner, W. *Journal of Chromatography A* **2010**, *1217*, 312-328.

(106)   Cech, N. B.; Enke, C. G. *Analytical Chemistry* **2000**, *72*, 2717-2723.

(107)   Burtscher, E.; Bobleter, O.; Schwald, W.; Concin, R.; Binder, H. *Journal of Chromatography* **1987**, *390*, 401-412.

(108)   Yu, Y.; Lou, X.; Wu, H. W. *Energy & Fuels* **2008**, *22*, 46-60.

(109)   Guo, G. L.; Chen, W. H.; Men, L. C.; Hwang, W. S. *Bioresource Technology* **2008**, *99*, 6046-6053.

(110)   Kootstra, A. M. J.; Mosier, N. S.; Scott, E. L.; Beeftink, H. H.; Sanders, J. P. M. *Biochemical Engineering Journal* **2009**, *43*, 92-97.

(111)   Pecina, R.; Burtscher, P.; Bonn, G.; Bobleter, O. *Fresenius Zeitschrift Fur Analytische Chemie* **1986**, *325*, 461-465.

(112)   Lapierre, C.; Rolando, C.; Monties, B. *Holzforschung* **1983**, *37*, 189-198.

(113)   Bonn, G.; Pecina, R.; Burtscher, E.; Bobleter, O. *Journal of Chromatography* **1984**, *287*, 215-221.

(114)   McLafferty, F. W. *Analytical Chemistry* **1959**, *31*, 82-87.

(115)   Li, W. K.; Sun, Y. K.; Liang, W. Z.; Fitzloff, J. F.; van Breemen, R. B. *Rapid Communications in Mass Spectrometry* **2003**, *17*, 978-982.

(116)   Fang, N. B.; Yu, S. G.; Prior, R. L. *Journal of Agricultural and Food Chemistry* **2002**, *50*, 3579-3585.

(117)   Morreel, K.; Kim, H.; Lu, F. C.; Dima, O.; Akiyama, T.; Vanholme, R.; Niculaes, C.; Goeminne, G.; Inze, D.; Messens, E.; Ralph, J.; Boerjan, W. *Analytical Chemistry* **2010**, *82*, 8095-8105.

(118)   Cuyckens, F.; Claeys, M. *Journal of Mass Spectrometry* **2004**, *39*, 1-15.

(119)   Fabre, N.; Rustan, I.; de Hoffmann, E.; Quetin-Leclercq, J. *Journal of the American Society for Mass Spectrometry* **2001**, *12*, 707-715.

(120)   Evtuguin, D. V.; Amado, F. M. L. *Macromolecular Bioscience* **2003**, *3*, 339-343.

(121)   Reale, S.; Di Tullio, A.; Spreti, N.; De Angelis, F. *Mass Spectrometry Reviews* **2004**, *23*, 87-126.

(122)   Pandey, M. P.; Kim, C. S. *Chemical Engineering & Technology* **2011**, *34*, 29-41.

(123)   Boon, J. J. *International Journal of Mass Spectrometry and Ion Processes* **1992**, *118*, 755-787.

(124)   Vanderhage, E. R. E.; Mulder, M. M.; Boon, J. J. *Journal of Analytical and Applied Pyrolysis* **1993**, *25*, 149-183.

(125)   Genuit, W.; Boon, J. J.; Faix, O. *Analytical Chemistry* **1987**, *59*, 508-513.

(126)   vanderHage, E. R. E.; Boon, J. J. *Journal of Chromatography A* **1996**, *736*, 61-75.

(127)   Hedges, J. I.; Ertel, J. R. *Analytical Chemistry* **1982**, *54*, 174-178.

(128)   Kaiser, K.; Benner, R. *Analytical Chemistry* **2012**, *84*, 459-464.

(129)   Tai, D. S.; Chen, C. L.; Gratzl, J. S. *Journal of Wood Chemistry and Technology* **1990**, *10*, 75-99.

(130)   Neto, C. P.; Seca, A.; Fradinho, D.; Coimbra, M. A.; Domingues, F.; Evtuguin, D.; Silvestre, A.; Cavaleiro, J. A. S. *Industrial Crops and Products* **1996**, *5*, 189-196.

(131)   Metzger, J. O.; Bicke, C.; Faix, O.; Tuszynski, W.; Angermann, R.; Karas, M.; Strupat, K. *Angewandte Chemie-International Edition in English* **1992**, *31*, 762-764.

(132)   DeAngelis, F.; Fregonese, P.; Veri, F. *Rapid Communications in Mass Spectrometry* **1996**, *10*, 1304-1308.

(133)   De Angelis, F.; Nicoletti, R.; Spreti, N.; Veri, F. *Angewandte Chemie-International Edition* **1999**, *38*, 1283-1285.

(134)   Evtuguin, D. V.; Domingues, P.; Amado, F. L.; Neto, C. P.; Correia, A. J. F. *Holzforschung* **1999**, *53*, 525-528.

(135)   Saito, K.; Kato, T.; Tsuji, Y.; Fukushima, K. *Biomacromolecules* **2005**, *6*, 678-683.

(136)   Saito, K.; Kato, T.; Takamori, H.; Kishimoto, T.; Fukushima, K. *Biomacromolecules* **2005**, *6*, 2688-2696.

(137)   Morreel, K.; Dima, O.; Kim, H.; Lu, F. C.; Niculaes, C.; Vanholme, R.; Dauwe, R.; Goeminne, G.; Inze, D.; Messens, E.; Ralph, J.; Boerjan, W. *Plant Physiology* **2010**, *153*, 1464-

1478.

(138)   Owen, B. C.; Haupert, L. J.; Jarrell, T. M.; Marcum, C. L.; Parsell, T. H.; Abu-Omar, M. M.; Bozell, J. J.; Black, S. K.; Kenttamaa, H. I. *Analytical Chemistry* **2012**, *84*, 6000-6007.

(139)   Haupert, L. J.; Owen, B. C.; Marcum, C. L.; Jarrell, T. M.; Pulliam, C. J.; Amundson, L. M.; Narra, P.; Aqueel, M. S.; Parsell, T. H.; Abu-Omar, M. M.; Kenttamaa, H. I. *Fuel* **2012**, *95*, 634-641.

(140)   Kiyota, E.; Mazzafera, P.; Sawaya, A. *Analytical Chemistry* **2012**, *84*, 7015-7020.

(141)   Fiehn, O. *Plant Molecular Biology* **2002**, *48*, 155-171.

(142)   Sumner, L. W.; Mendes, P.; Dixon, R. A. *Phytochemistry* **2003**, *62*, 817-836.

(143)   Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrometry Reviews* **2007**, *26*, 51-78.

(144)   Kind, T.; Fiehn, O. *Bmc Bioinformatics* **2007**, *8*.

(145)   Stagliano, M. C.; DeKeyser, J. G.; Omiecinski, C. J.; Jones, A. D. *Rapid Communications in Mass Spectrometry* **2010**, *24*, 3578-3584.

Chapter Two: Profiling of Corn Stover Cell Wall Nitrogenous Decomposition Products formed during Ammonia Fiber Expansion (AFEX) Treatment of Corn Stover using Liquid and Gas Chromatography and Mass Spectrometry

2.1 Introduction

Plant lignocellulosic biomass resists enzymatic digestion needed for conversion to liquid biofuels, and this resistance provides a substantial barrier to development of these resources for renewable energy. An assortment of pretreatment methods increase cell wall digestibility, but they also produce various products from lignin and hemicellulose decomposition that inhibit the downstream hydrolysis and fermentation [1,2]. Types and amounts of degradation products depend on the pretreatment methods and on biomass cell wall composition. These complex mixtures demand more comprehensive analytical strategies to provide detailed characterization of degradation products after biomass pretreatments. Such information aids understanding of the chemical reactions that occur during pretreatment, and guides optimization of pretreatment parameters for higher process efficiencies.

Acid pretreatments are notorious for producing significant amount of furan derivatives including the heterocyclic aldehydes furfural and 5-HMF which cause substantial inhibition of enzymatic hydrolysis and microbial activity [3-5]. Since furan derivatives are formed from sugar monomers in acidic environments [4,6], substantial losses of fermentable sugars are unavoidable in acid pretreatments. Acid pretreatments also release substantial amounts of carboxylic and phenolic acids and aromatic aldehydes from decomposition of lignin [1]. Of these, phenolics have shown the most inhibition to various fermentive microorganisms [4].

Alkaline-based pretreatments are less likely to form furan derivatives, but other common degradation products, including an assortment of phenolics, are formed during the process. Phenolics comprise a major category of pretreatment byproducts from acid- or ammonia-based pretreatments that inhibit hydrolytic enzymes and fermentative microorganisms; hence characterization of these products in treated biomass is essential for process optimization.

The presence of inhibitors of enzymatic processing and microbial fermentation in process

streams of ammonia-based biomass pretreatments has been reported by Chundawat et al. [7] and Mes-Hartree and co-workers [8]. However, AFEX-pretreated biomass gives higher yields of enzymatic digestion and fermentation compared to acid-pretreated biomass without need for detoxification or nutrient supplementation processes [7,9,10]. AFEX is one of the leading biomass pretreatment technologies and offers the potential for industrial-scale utilization for cellulosic ethanol production [11]. To better understand the chemistry of AFEX as needed to improve the economic viability of converting cellulosic biomass to ethanol, characterization of those decomposition products that decrease ethanol yields in AFEX-treated biomass extract is crucial. Comparisons of products in the AFEX-treated and untreated biomass extract and characterization of key compounds including cell-wall cross-linkers that bridge hemicellulose polysaccharides reveal information about the mechanisms underlying how AFEX improves digestion and fermentation yields.

AFEX pretreatment is an alkaline process performed in the presence of moisture and ammonia, and therefore hydrolysis and ammonolysis of ester bonds are likely to occur (Figure 2.1) [12]. The focus of the research described in this chapter has centered upon identifying and quantifying products formed during AFEX pretreatment of corn stover, specifically those formed via ammonolysis reactions. These reactions represent a sink for ammonia, which constitutes an important production cost. Initial efforts to characterize decomposition products in AFEX-treated corn stover (AFEXTCS) extract led to detection of various carboxylic acid and phenolics along with a suite of nitrogenous compounds that were not present in untreated corn stover (UTCS) extracts, suggesting incorporation of nitrogen into cell wall components during AFEX. Characterization of two categories of nitrogenous compounds including Maillard reaction products and phenolic amides is the subject of research described in this chapter.
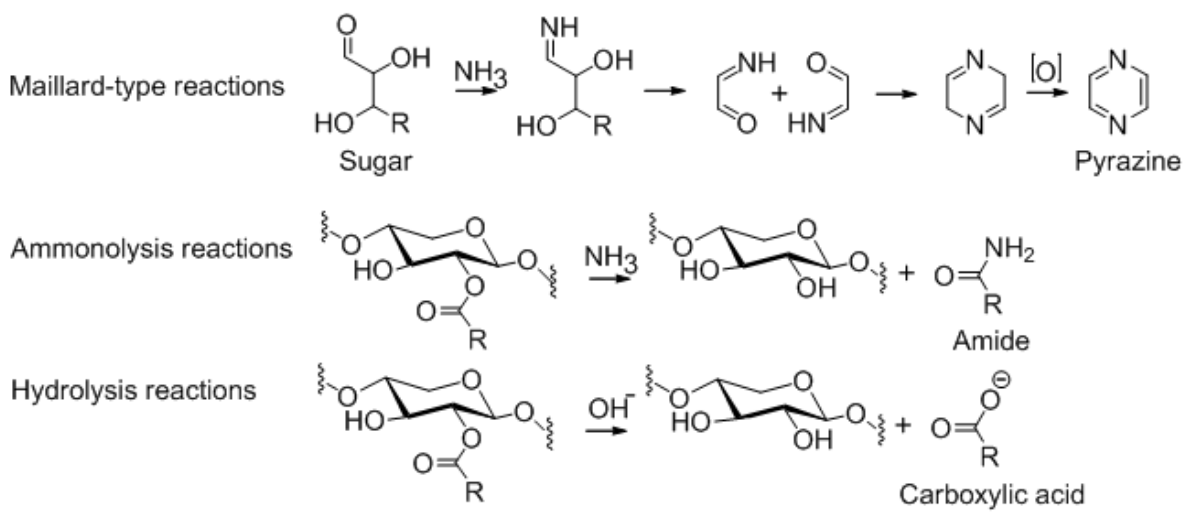
Figure 2.1. Schematic representation of Maillard-type reactions and ammonolysis and hydrolysis of lignin-carbohydrate ester linkages that take place during ammonia based pretreatments.

Previous research indicated the presence of several nitrogen-containing compounds in AFEXTCS extracts including proteins and amino acids and residual ammonia, and these may be utilized to provide and important nutrient sources for ethanologenic microbes [10]. One potential application of AFEX-treated corn stover after production of ethanol is usage of the remaining residues as animal feed, as has long been the practice with ammoniated molasses [13,14]. Hence characterization of nitrogen-containing products in AFEXTCS is important because anticipated nitrogenous compounds including imidazoles and pyrazines have shown toxicity in animals [15,16].

Although research has focused on characterization of carboxylic acids, phenolics and furans in biomass extracts after various pretreatments [17-20] including ammonia-based

pretreatments [7] characterization of decomposition products, most notably nitrogenous compounds in AFEX treated biomass, had not been reported prior to the results described herein. This chapter reports characterization of nitrogenous compounds including phenolic amides, pyrazines, and imidazoles that are formed during AFEX. Several compounds in each category were detected, characterized and quantified using a multifaceted strategy based on LC/MS and GC/MS analyses. Quantification of nitrogenous species was used to calculate the ammonia loss during the AFEX process.

2.2 Materials and methods

2.2.1 Chemicals

Ferulic acid, *p*-coumaric acid, pyrazine, 2-methylpyrazine, 2,5-dimethylpyrazine, 2,6-dimethylpyrazine, 4-methylimidazole, 1-methylimidazole and 2,4-dimethylimidazole were purchased from Sigma. Feruloyl amide and coumaroyl amide were synthesized by treating feruloyl ethyl ester and coumaroyl methyl ester with AFEX following purification.

2.2.2 AFEX and DA pretreatments

Corn stover was harvested in 2002 from Kramer farm in Wray, CO, USA. Acid treated corn stover (ATCS) was prepared using a pilot scale continuous reactor [21] at 190 ^{o}C at a biomass solids and sulfuric acid loading of 30% (w/w) and 0.048 g/g dry corn stover respectively and whole slurry was used for analysis.

AFEX pretreatment was carried out in a sealed reactor capable of maintaining elevated pressures based on a previously published protocol [10]. Experimental conditions were as follows: 130 ^{o}C, 60% moisture and 1:1 NH_3 to biomass loading (w/w) for total of 15 min residence time. This was considered as regular AFEX condition. High severity AFEX (H-AFEX)

57

was performed at 130 ^{o}C, 60% moisture and 3:1 NH$_3$ to biomass loading (w/w) for a total of 45 min residence time (higher ammonia to biomass load and residence time compared to regular AFEX) and low severity AFEX (L-AFEX) was performed at similar conditions as regular AFEX at 100 ^{o}C.

2.2.3 GC-MS analysis

Acetic acid, acetamide, furfural and 5-HMF were measured using a high-throughput GC/MS method which we published earlier [22]. ATCS and AFEXTCS aqueous extracts were injected directly for analysis without derivatization. Analyses were performed using an Agilent 6890N GC, equipped with an Agilent 7683 autosampler that was coupled to an Agilent 5973 quadrupole mass spectrometer. Separation was performed using a 30 m J&W DBWAX column (0.25 mm ID, 0.25 µm film thickness). The injection volume was 1 µL using splitless mode with a 240 ^{o}C injector temperature. Helium was used as the carrier gas at a flow-controlled rate of 1.5 ml/min. The temperature gradient was as follows: start at 50 ^{o}C and hold for 1 min, increase to 100 ^{o}C at a rate of 30 ^{o}C/min and hold for 1 min, increase to 240 ^{o}C at a rate of 38 ^{o}C/min and hold for 1 min with a solvent delay of 5 min. Data were collected using 70 eV electron ionization (EI) and selected ion monitoring (SIM), measuring signals for primary and secondary masses for compounds of interest. Products of Maillard reactions (pyrazines and imidazoles) were also separated and quantified using the same DBWAX column using the following temperature program: initial 40 ^{o}C (2 min), followed by an increase at 6 ^{o}C/min to 150 ^{o}C (1 min hold), 20 ^{o}C/min to 200 ^{o}C (2 min hold), then 20 ^{o}C/min to 240 ^{o}C (4 min hold). Spectra were generated using 70 eV electron ionization in full scan mode, and the NIST05 mass spectrum database was used to support identification of pyrazines and imidazoles. Pyrazine, 2-methylpyrazine, 2,5-dimethylpyrazine, 2,6-dimethylpyrazine, 4-methylimidazole and 2,4-dimethylimidazole standard

solutions were prepared at 1, 5, 10, 25 and 50 mg/L in methanol and used for external standard quantification. 1-Methyl imidazole was used as internal standard and was added to sample extracts and standards at 10 mg/L. Relative response factors were calculated based on the peak area of the base peak extracted ion chromatogram (molecular ion; $M^{\bullet+}$ in most cases) of each compound relative to the area of the internal standard peak. With the assumption that all other pyrazines have the same response factor as 2,5-dimethylpyrazine, pyrazines (that their authentic standards were not available) were quantified based on calibration curves made with 2,5-dimethylpyrazine standard. 2-Methyl-1H-imidazole was also quantified based on the calibration curve of 4-methyl-1H-imidazole assuming that they have identical response factors.

2.2.4 LC-MS analysis

The LC–MS system used for quantification of phenolic amides consisted of an Acquity ultra-performance liquid chromatograph (UPLC) connected to a Waters Quattro Premier XE triple quadrupole mass spectrometer that was operated using electrospray ionization (ESI). AFEXTCS, UTCS and ATCS extracts were directly analyzed by LC–MS using an Ascentis Express C18 (50 × 2.1 mm, 2.7 µm particles) fused core silica reversed phase column (Supelco, USA). A 30-min gradient was used for profiling and non-targeted analysis of AFEXTCS and UTCS extracts and a 5-min gradient was used for large-scale quantification of targeted products. Gradient elution was carried out using aqueous 0.1% formic acid (v/v; solvent A) and methanol (solvent B) under the following conditions. For the 30-min gradient: (1) 95% A from 0-1 min, (2) linear gradient to 70% A at 10 min and then to 50% A at 18 min with a hold at 50% B until 25 min (3) return to initial conditions (95% A) at 25.01 followed by a 5-min hold for column equilibration. For the 5-min gradient: (1) 95% A 0-1 min, (2) linear gradient to 70% A at 4 min followed by a sudden change in gradient to 5% A at 4.01 min and a hold at this composition until

4.50 min (3) change to initial condition (95% A) at 4.51 min followed by a 0.5-min hold. Injection volume and column temperature were 3 µL and 50 ^{o}C respectively. Flow rate was 250 µL/min for the 30 min gradient and 350 µL/min for the 5 min gradient. Mass spectra were acquired in full scan mode from 100-1500 Da with scan time of 0.2 seconds.

For quantification purposes, all mass spectra were acquired using negative mode ESI, and extracted ion chromatograms for the deprotonated molecular ions ([M-H]$^-$) were generated. All peak integrations and quantitative calculations were performed using QuanLynx (V4.1) software (Waters Corporation, USA). 3,4,5 Trimethoxycinnamic acid was added as internal standard at a concentration of 5 µM in all standard solutions and extracts. Ferulic and coumaric acids and their amides were quantified based on external calibration curves (relative response factors to the internal standard) generated using authentic standards. Diferuloyl amides were quantified assuming their response factors were the same as determined for feruloyl amide.

The LC–MS system used for exact mass measurements and non-targeted analysis consisted of a Shimadzu HPLC (LC-20AD pump) coupled to a Waters LCT Premier time-of-flight mass spectrometer (TOF/MS). The column and gradient (30 min) detailed above were used for this analysis as well. Detection, integration, and retention time alignment was performed for the entire mass range, and principal component analysis (PCA) for UTCS and AFEXTCS extract data was performed using MarkerLynx (V4.1) software (Waters Corporation, USA).

2.3 Results and discussion

2.3.1 Nitrogen content and ammonia balance in AFEXTCS

A detailed mass balance of nitrogen performed by our collaborators at the Michigan State University Biomass Conversion Research Laboratory (BCRL) showed about 300% increase in the total nitrogen content of AFEXTCS compared to UTCS [12]. Almost 15% of this increase in

nitrogen content was attributed to residual ammonia or ammonium ion from the process that remained bound even after air drying the pretreated corn stover. The remaining 85% was attributed to soluble nitrogenous products formed by ammonolysis of lignin-carbohydrate-complex (LCC) linkages during AFEX.

2.3.2 LC/MS analysis of AFEXTCS and UTCS water extract

LC/TOF-MS total ion chromatograms of aqueous extracts of AFEXTCS and UTCS revealed complex mixtures (Figure 2.2) composed largely of cell wall-derived phenolic substances based on calculations of relative mass defect (RMD) values. More than 70% were in the range of 200-400 ppm, consistent with their annotation as phenolic compounds.

RMD provides a surrogate measurement of the fraction of molecular mass that corresponds to hydrogen, and this value can aid assignments to various classes of compounds based only on accurate measurements of molecular mass [23]. For example, membrane lipids including phospholipids have high RMD (600–1000 ppm) owing to their greater hydrogen content, while sugars (300–400 ppm) and phenolics (200–400 ppm) have a lower RMD owing to their low hydrogen content. Phenolic compounds typically have RMD values of 200-400. For example, the theoretical mass of [M-H]$^-$ of 193.0511 for ferulic acid has RMD of 265 ppm as calculated below:

$$[(193.0511\text{-}193)/(193.0511)] \times 10^6 = 265 \text{ ppm}$$
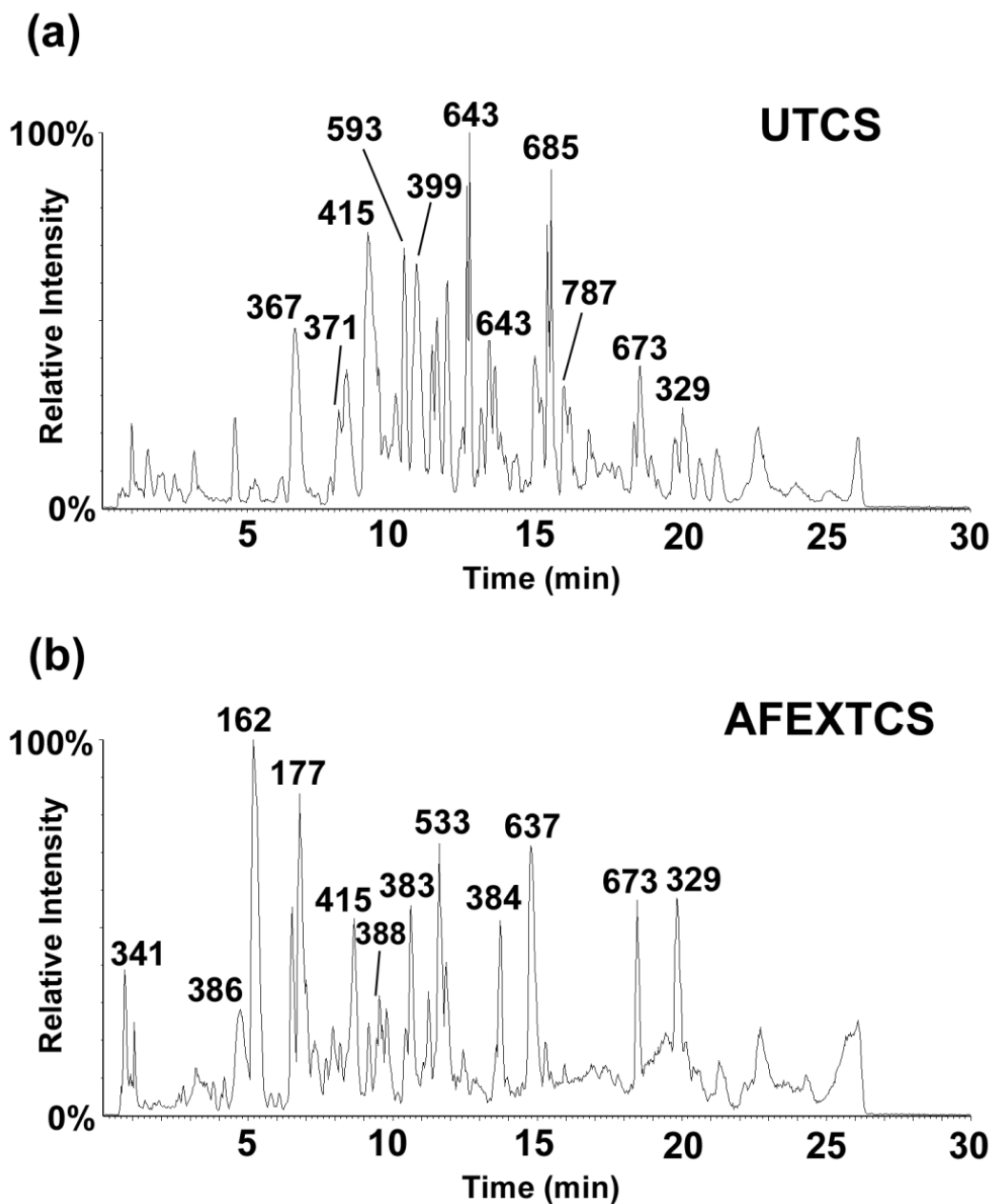
Figure 2.2. LC/TOF-MS total ion chromatograms of aqueous extracts of (a) UTCS and (b) AFEXTCS water extract using ESI negative mode. To highlight differences in composition, several abundant peaks are labeled with nominal *m/z* values.

Automated peak extraction and integration yielded more than 2000 peaks after deisotoping in the AFEXTCS extract. The LC/MS data indicated a range of different molecular masses (*m/z*

values) for the major products as illustrated in Figure 2.2 For example, ions with *m/z* 643, 685, 593 and 399 were among the most abundant ions in the LC/MS chromatogram of UTCS while *m/z* 162, 177, 533 and 637 were among the most abundant ions in LC/MS chromatogram of AFEXTCS.

Principal component analysis (PCA) on the LC–MS data from UTCS and AFEXTCS extracts yielded a loadings plot that revealed several compounds with abundances with substantial contributions to the variance between extracts. In particular, ions with even nominal *m/z* values indicate compounds with odd numbers of nitrogen atoms, a consequence of nitrogen having the unusual properties of an odd valence but even atomic mass. ESI generates closed shell ions (e.g. protonated or deprotonated molecular ions in positive and negative modes, respectively). Therefore, even *m/z* values in electrospray ionization mass spectra reflect odd number of nitrogen atoms in the ion. Based on this even-mass criterion, at least 50 nitrogen-containing ions were detected among the 180 negative-ion signals that contributed most to PCA loadings. Prominent among these were compounds detected as *m/z* 162, 382, 384, 386 and 388 ions. Seventeen uncharacterized ions from the identified 50 nitrogenous compounds with high PCA loading scores are listed in Table 2.1 along with their accurate masses and RMD values. All listed compounds in Table 2.1 were detected only in AFEXTCS extracts. It is anticipated that several nitrogenous compounds listed in Table 2.1 are phenolics based on their RMD values, which lie in the range of 200 to 350 ppm. These are not annotated as simple sugars or compounds of high aqueous solubility because of their relatively high chromatographic retention times using the C-18 column.

Table 2.1. Putative nitrogenous compounds detected only in AFEXTCS extract along with their exact mass, RMD and their LC retention time. Coumaric and ferulic acid retention times, [M-H]$^-$ masses, and RMD values are provided for reference. Compounds are organized based on increasing RMD values. Quantifications of listed compounds were based on feruloyl amide standard assuming response factors equal to feruloyl amide (n=3; the uncertainties are one standard deviation.

Compound	Retention time (min)	Measured mass (*m/z*)	RMD (ppm)	µg analyte/g AFEXTCS
Coumaric acid	8.2	163.0409	251	1085±94
Ferulic acid	9.5	193.0511	265	103±9
1	11.8	442.1212	274	69 ± 5
2	16.3	382.1059	277	102 ± 9
3	14.6	592.1862	314	99 ± 7
4	14.6	610.1975	324	68 ± 10
5	7.1	236.0783	332	152 ± 18
6	8.6	386.1378	357	325 ± 21
7	6.4	356.1274	358	68 ± 9
8	8.3	430.1551	361	394 ± 25
9	12.3	596.2155	362	136 ± 10
10	9.8	412.1507	366	110 ± 9
11	7.3	250.0922	369	120 ± 15
12	8.3	430.1596	371	499 ± 41
13	14.5	580.2174	375	74 ± 8
14	8.1	370.1407	380	86.6 ± 9.5
15	10.5	352.1360	386	74.9 ± 8.8
16	3.4	400.1607	401	118.1 ± 12.9
17	7.2	186.1215	653	138.6 ± 18.6

The fractional hydrogen content is estimated using RMD values, and the distribution of hydrogen content can be assessed from a histogram of RMD values for the 180 ions with highest combined loadings (based on the MarkerLynx score termed 'significance') detected in AFEXTCS or UTCS extracts.

Almost half of the 180 compounds with highest loadings are consistent with their annotation as phenolics based on their RMD values falling in the range of 200 to 350 ppm. About 15% appear to contain sugars, and several of these compounds could be phenolic based on fragment ions generated at elevated CID potentials (for example formation of coumarate or ferulate ions upon CID). Recent analysis has suggested presence of phenolic glucosides within UTCS water extracts [24]. In this report [24], hydrolysis of a red-brown fraction of corn stover extract resulted in formation of various hemicellulose sugars and lignin structural units including coniferyl, sinapyl and coumaroyl alcohols. Based on these findings, it was concluded that about 10-20% of corn stover water extract is composed of phenolic glycosides [24] which could be the case for products shown in the histogram in Figure 2.3 with RMD values higher than those of phenolics (200-350 ppm). Adding sugars to phenolic compounds increases the RMD value. The possibility of some of the compounds with RMD values in the range of 200-300 ppm being flavonoids is not excluded since several flavonol glycosides including quercitrin (quercetin-3-*O*-rhamnoside) and dimethyl-rutin were characterized in the AFEXTCS extract based on the *m/z* values of their [M-H]$^-$ ion (*m/z* 637 and 477 for dimethyl-rutin and quercitrin respectively), comparison of retention time with authentic standards and fragment ions in the MS/MS CID spectra on the [M-H]$^-$ ion. However, flavonoids have distinct CID fragmentation patterns that were not observed at least for compounds listed in Table 2.1.

Almost 25% of compounds in the histogram in Figure 2.3 have RMD values higher than

600 ppm, which means they have high hydrogen percent by weight of the compound and are suggested to be lipid-like substances.

It should be noted that some ions with high PCA loadings including *m/z* 685, 643 and 563 (in negative ESI mode) were detected only in UTCS, suggesting they react to form other products during AFEX. Extracted ion chromatograms for a few of the discriminating ions (*m/z* 643, 388, 563, and 384) illustrate the relative abundances in UTCS and AFEXTCS, and are shown in Figure 2.4.
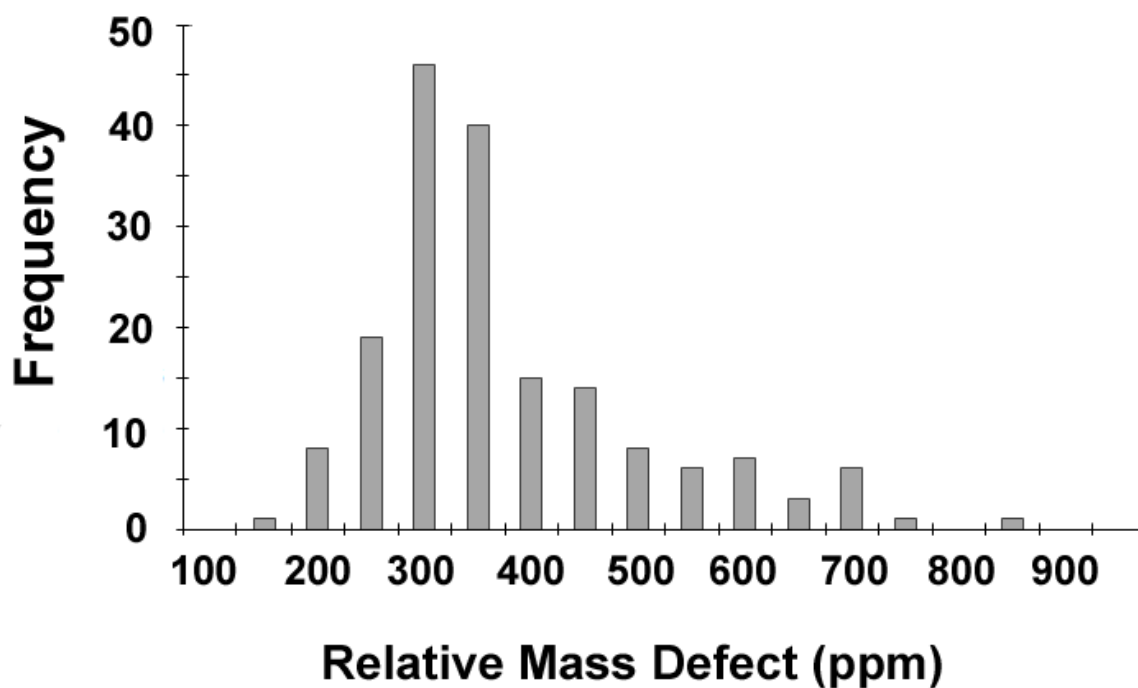
Figure 2.3. Histogram of RMD values of 180 compounds with highest PCA loadings scores detected in negative-ion mode LC/MS analyses from AFEXTCS and UTCS extracts.
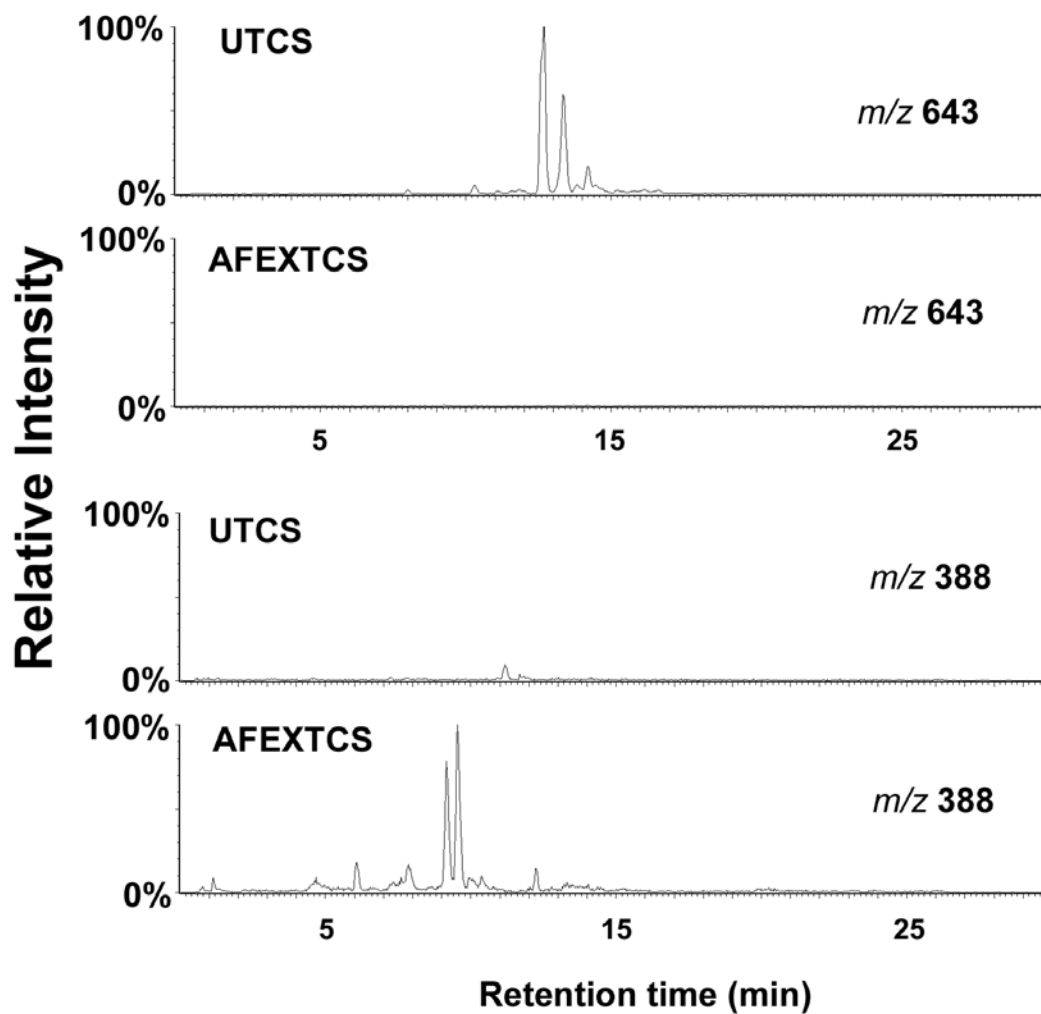
Figure 2.4. Negative-ion mode extracted ion LC/MS chromatograms of ions selected based on high PCA loadings scores from AFEXTCS and UTCS extracts. Ion abundances are normalized to the highest count in each pair to facilitate relative quantitative comparisons.

Figure 2.4 (cont'd)
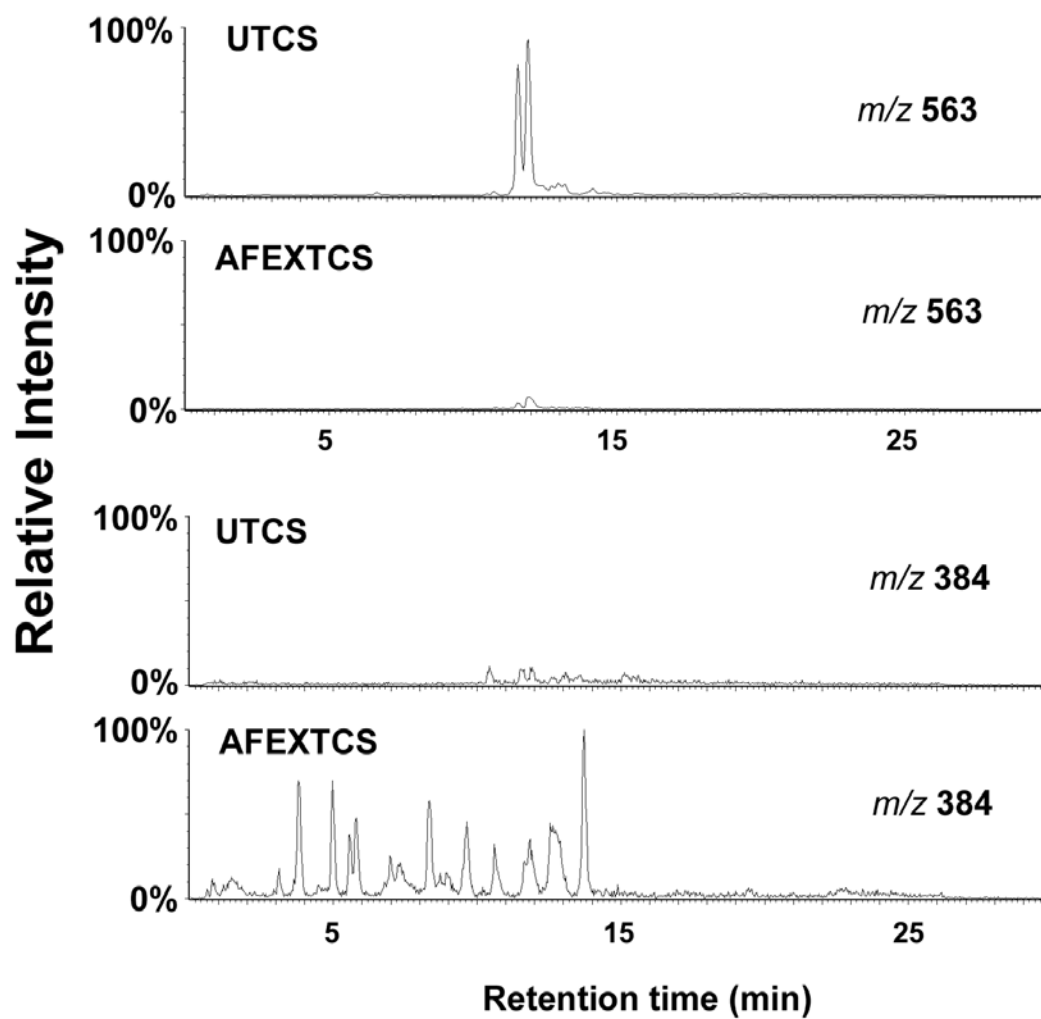
2.3.3 Identification and quantification of important nitrogenous compounds in the AFEXTCS extract

Formation of phenolic amides is an anticipated nitrogen sink during AFEX pretreatment of corn stover, and these compounds are predicted to arise from ammonolysis of phenolic acid esters in plant cell walls. Structures of several nitrogenous compounds present in the AFEXTCS extract can be predicted based on current understanding of secondary cell wall structure and reactions during AFEX. Extracted ion LC/MS chromatograms for several even *m/z* values in AFEXTCS extract (Figure 2.2-b) show evidence of compounds with an odd number of nitrogen atoms. It was mentioned earlier that most of these compounds are annotated as phenolics based on their RMD values.

During the AFEX process, ammonolysis is expected to generate nitrogenous compounds by nucleophilic attack of ammonia at various ester linkages including acetate, coumarate and ferulate esters in corn stover secondary cell walls. Ammonolysis of O-acetyl groups attached to hemicellulose arabinoxylans [25-27] generated about 25 mg acetamide/g AFEXTCS as measured by GC/MS. This reaction turned out to be one of the largest nitrogen sinks in the AFEX process, and generated 58% of the total amount of identified nitrogenous compounds (25 mg of total of 44 mg/g AFEXTCS identified nitrogenous compounds). In acid treatment processing, hydrolysis generated 35 mg/g ATCS of acetic acid, and this value provides an estimate of the amounts of acetate esters present in the cell wall. Phenolic amides generated using AFEX pretreatment of corn stover reflect the second largest nitrogen sink in the AFEX process, as quantification showed they accounted for almost 15 mg/g AFEXTCS. Coumaric and ferulic acids and their derivatives comprise nearly 4-6% of corn bran cell walls [28]. They exist in the form of free acids, and as soluble and insoluble esters in rice, corn and other grains [29]. Coumarate and ferulate ester levels in maize internodes are reported to reach as high as 20 and 7 mg/g cell wall

respectively [30]. In contrast,, the yields of coumaric and ferulic acids formed upon hydrolysis of their corresponding esters during AFEX was about 1.1 and 0.1 mg/g AFEXTCS, respectively. Hence it was concluded that the overwhelming majority of coumarate and ferulate esters were ammonolyzed to their corresponding amides (Figure 2.5).
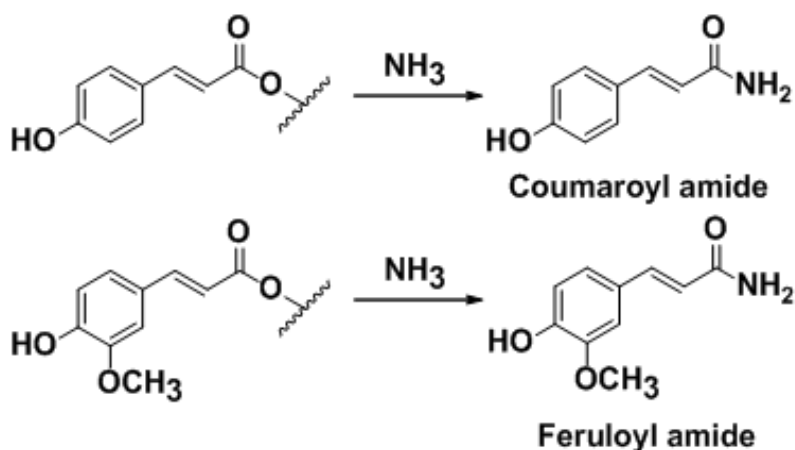
Figure 2.5. Formation of coumaroyl and feruloyl amides from ammonolysis of coumaroyl and feruloyl esters in plant cell walls during AFEX pretreatment of corn stover.

Compounds with nominal pseudomolecular ions of *m/z* 162 and 192 were observed in the ESI negative mode mass spectra generated during LC/MS analysis of AFEXTCS extract (Figure 2.2-b). These compounds eluted at retention times of 5.5 and 8.5 min, and were identified as coumaroyl and feruloyl amides based on the following information. The measured *m/z* values of these ions using TOF-MS analyzer were 162.0530 and 192.0671 which showed 4.9 and 2.6 ppm error from the theoretical mass of [M-H]⁻ ions of coumaroyl (162.0561) and feruloyl (192.0666) amide respectively. Characterization was further confirmed by matching retention times and fragment ions that were generated upon CID of [M-H]⁻ of the authentic standards for both of these ions in the AFEXTCS extracts. Neither compound was detected in either the ATCS nor

70

UTCS extracts.

Ammonolysis and hydrolysis reactions during AFEX are expected to release diferulate cross-linkers that bridge hemicellulose polysaccharides [31-34] since the levels of soluble arabinoxylans after AFEX increase following treatment (discussed in Chapter Four) [12]. Dehydrodimerization of esterified ferulic acids to hemicellulose carbohydrate polymers, notably arabinoxylans, provides cross-linking of polysaccharides in lignocellulosic plant cell walls. From this knowledge, it was anticipated that these ferulate dimers would be released either as di-amide (Di-Am) diferulates via ammonolysis, di-acid (Di-Ac) diferulates via hydrolysis or acid-amide (Ac-Am) diferulates via combination of both ammonolysis and hydrolysis reactions. The complexity of the mixture of products released by AFEX is depicted in Figure 2.6, which shows the total ion LC/MS chromatogram of an AFEXTCS water extract, generated in positive ion mode. In anticipation of detecting diferulates, extracted-ion chromatograms were generated for ions corresponding to $[M+H]^+$ of Di-Am (m/z 385), Ac-Am (m/z 386) and Di-Ac (m/z 387) diferulates, yielding more than 30 chromatographic peaks with these nominal masses. Peaks that exhibited these masses are highlighted in the Figure 2.6. Of these, 15 ions gave masses within 5 ppm of theoretical values for diferulates. Only one of the ions yielded a pseudomolecular ion at m/z 387, which corresponds to a diferulic acid that could arise from hydrolysis, and not ammonolysis, of the two ester bonds. All of the other 14 ions had m/z values of either 386 or 385, corresponding to Ac-Am and Di-Am diferulates respectively. Based on integrated peak areas, over 90% of AFEX-released diferulates contained at least one amide group. These findings demonstrate that ammonolysis is the main reaction in AFEX responsible for releasing arabinoxylans by removing diferulates cross-linkers. All 14 nitrogenous diferulates were quantified based on the assumption that they shared the same relative response factor as feruloyl

amide. Total amount of phenolic amides including feruloyl amide, coumaroyl amide and
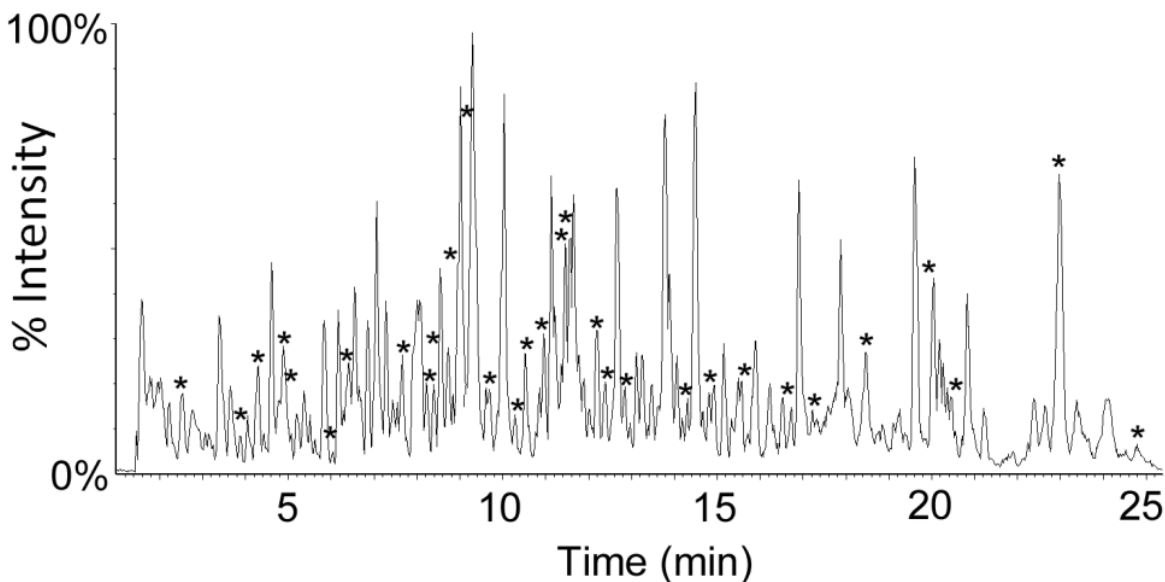
Figure 2.6. LC/TOF-MS total ion chromatogram of AFEXTCS extract generated using ESI in positive ion mode. Asterisks designate compounds with nominal pseudomolecular masses matching anticipated diferulates in Di-Ac Ac-Am, or Di-Am forms.

diferuloyl amides was 14.8 mg/g of AFEXTCS, which were about 35% of the total identified nitrogenous compounds per gram of AFEXTCS. This suggests that similar to ammonolysis of acetyl groups and formation of acetamide, ammonolysis of cell wall ferulate and coumarate esters is the other large nitrogen sink in the AFEX. No amide (monomer and dimer) and no diferulic acid were identified in the ATCS extract and ferulic and coumaric acids together were six-fold lower in the ATCS extract compare to ferulic and coumaric acid in AFEXTCS. This could be due to acid catalyzed decarboxylation of ferulic and coumaric acid during the acid pretreatment of corn stover.

Dehydrodimerization of ferulic acids in cell wall results in formation of various isomers [35-37] that arise from formation of carbon-carbon or carbon-oxygen bonds at various positions.

These isomers share a common chemical formula. The diversity of crosslinking positions combined with differences in ammonolysis or hydrolysis explain why 15 compounds have identified as diferulates in the AFEXTCS extract. Detailed structural characterization of each diferulate isomer will be the subject of discussion in Chapter Three.

The analytical results described above demonstrate that both hydrolysis and ammonolysis reactions occur during AFEX based on the reaction products that were characterized in the AFEXTCS extract. Ammonolysis reactions incorporate nitrogen in cell wall components, causing loss of ammonia during AFEX. Two large nitrogen sinks during AFEX were ammonolysis of arabinoxylan O-acetyl groups to form acetamide and ammonolysis of coumarate, ferulate and diferulate esters which form coumaroyl, feruloyl and various diferuloyl amides. Ammonolysis by-products including acetamide (25 mg/g AFEXTCS) and all phenolic monomer and dimer amides (14.8 mg/g AFEXTCS) account for 36% of the total ammonia loss during AFEX, leaving 64% that must be attributed to other events. Part of the remaining 64% ammonia loss is attributed to formation of other nitrogenous compounds, of which several yet-unidentified compounds are listed in Table 2.1. Maillard reaction products are another category of nitrogenous compounds which form upon reaction of ammonia with monomeric sugars during AFEX, and these are described in more detail in the next section.

2.3.4 Maillard reactions during AFEX

Maillard reactions play important roles in chemical reactions between carbohydrates and amino-containing compounds. In the context of AFEX, such reactions are expected to occur via condensation of sugar aldehyde groups with ammonia during AFEX through α-amino-carbonyl intermediates as shown in Figure 2.1. Pyrazine and related soluble derivatives (e.g. methyl- and dimethylpyrazines) are among the major products of Maillard reactions [38,39], and these are

73

typically analyzed using GC/MS owing to their volatility Figure 2.7 shows a extracted ion chromatogram representing multiple ions indicative of assorted Maillard products from GC/MS analysis of AFEXTCS extract. The total amount of pyrazine and imidazole related derivatives formed via condensation of ammonia with reducing sugars in corn stover was 0.95 mg/g AFEXTCS, accounting for nearly 2% of the ammonia loss during AFEX. Individual amounts and characterized isomers are listed in Table 2.2, and structures of the identified compounds are shown in Figure 2.7-a.

Table 2.2. Maillard reaction products formed during AFEX treatments of corn stover of varying severity, measured using GC/MS (n=3), the uncertainties represent one standard deviation.

		µg analyte/g biomass		
	Maillard reaction products	AFEXTCS	H-AFEXTAV	L-AFEXTAV
1	**Pyrazine**	18.9±3.6	7.1±1.5	24.1±3.5
2	**2-Methylpyrazine**	44.8±6.3	44.3±1.5	63.4±3.0
3	**2,5-Dimethylpyrazine**	5.6±1.4	23.8±3.5	10.8±1.4
4	**2,6-Dimethylpyrazine**	16.3±1.9	32.1±6.1	13.1±2.3
5	**2-Ethylpyrazine**	4.6±0.9	6.9±2.1	7.8±2.1
6	**2,3-Dimethylpyrazine**	5.5±1.1	53.5±8.9	28.4±3.8
7	**2-Ethyl-6-methylpyrazine**	3.6±1.1	9.4±2.3	4.1±0.9
8	**2-Ethyl-5-methylpyrazine**	3.1±0.7	5.4±0.4	2.9±0.3
9	**2,3,5-Trimethylpyrazine**	5.1±1.2	69.1±5.6	12.6±2.1
10	**2-Ethenylpyrazine**	3.8±0.9	12.4±2.3	7.6±2.1
11	**3-Ethyl-2,5-dimethylpyrazine**	1.5±0.1	5.9±1.1	0.9±0.1
12	**2,6-Diethylpyrazine**	3.6±0.8	5.8±1.0	3.1±0.7
13	**Tetramethylpyrazine**	3.1±0.7	5.9±1.0	3.4±1.8
14	**6-Methyl-2-ethenylpyrazine**	2.9±0.6	8.8±2.1	5.6±1.0
15	**Acetylpyrazine**	2.9±0.5	24.1±4.6	12.5±2.1
16	**2-Acetyl-5-methylpyrazine**	2.4±0.8	4.5±1.1	2.1±0.2
17	**3-Methyl-2-pyrazinyl methanol**	14.9±2.4	12.5±2.3	6.9±1.1
18	**6-Methyl-2-pyrazinyl methanol**	265.6±8.5	4.4±1.0	3.5±0.5
19	**2-Methyl-1-H-imidazole**	30.8±2.5	212.5±7.6	29.4±1.2
20	**2,4-Dimethyl-1-H-imidazole**	107.9±8.8	188±8.8	17.4±2.1
21	**4-Methyl-1-H-imidazole**	418±8.9	411±7.5	42±1.7
	Total Maillard reaction products	965±10	1147±7	302±4

Major pyrazine derivatives in AFEXTCS were 2-methylpyrazine, 2,6-dimethylpyrazine and 6-methyl-2-pyrazinylmethanol, and the major imidazole products formed during AFEX were 4-methylimidazole and 2,4-dimethylimidazole.

Maillard reaction products were measured in corn stover extracts that were treated with AFEX at different severities. The extent of Maillard reactions depends not only on the nitrogen source but also on the pretreatment conditions including temperature employed [38-40]. This is apparent for lower severity AFEX treated Avicel (L-AFEXTAV) that gave 3- and 4-fold lower yields of total identified Maillard products compared to regular and high severity AFEX respectively. These findings are consistent with the extent of Maillard reaction during AFEX depending on temperature, residence time, and nitrogen load, with increases in each factor leading to increases in total amounts of Maillard reaction products (Table 2.2). Although this was not tested in the current study, it has been suggested that increasing temperature might result in formation of different Maillard reaction products [40].

Melanoidins are a separate category of Maillard reaction end-products [41] that are larger than pyrazines and imidazoles and are not as volatile; therefore are not amenable to GC/MS analysis [40]. This suggests the possibility of ammonia loss during AFEX due to formation of Maillard reaction products that were not identified in this study. Structures of such compounds are not yet well established, and it cannot yet be ruled out whether some as yet-unidentified products detected using LC/MS are melanoidins.
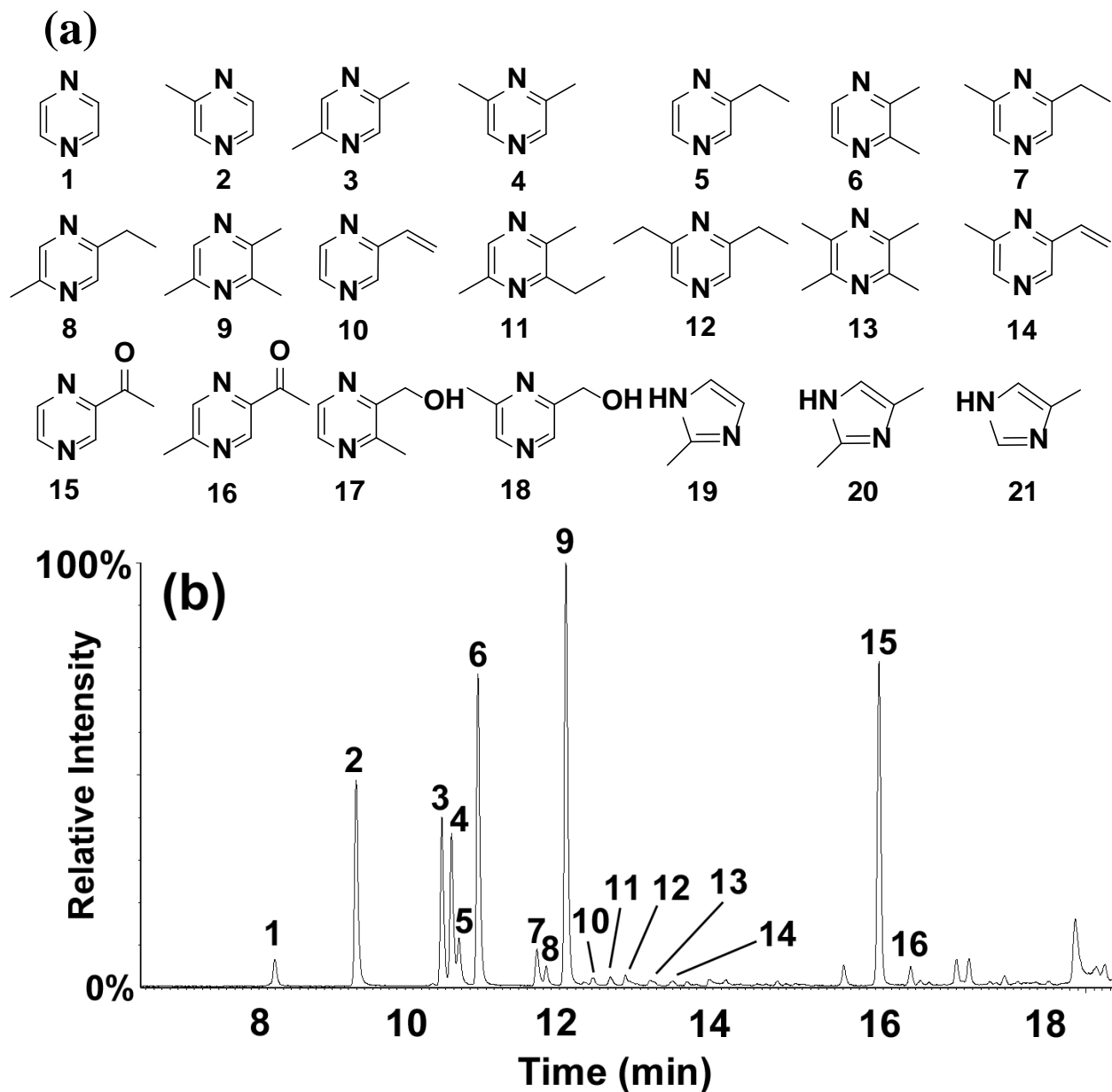
Figure 2.7. Combined extracted ion chromatograms for characterized pyrazines and imidazoles formed during AFEX pretreatment of corn stover (*m/z* 80, 82, 94, 96, 108, 122, 124 and 136; these ions are the base peaks in the EI spectra of each compound, and in several cases are molecular ions). Figure 2.7-b shows combined extracted ion chromatograms of various pyrazines listed in Table 2.2 (pyrazines #1 to 16) and Figure 2.7-c shows pyrazines and imidazoles # 17-21 listed in Table 2.2. Corresponding structures of these compounds are shown in Figure 2.7-a. Compound Identification was based on EI spectra in NIST library and retention time matching with standards for several compounds.

Figure 2.7 (cont'd)
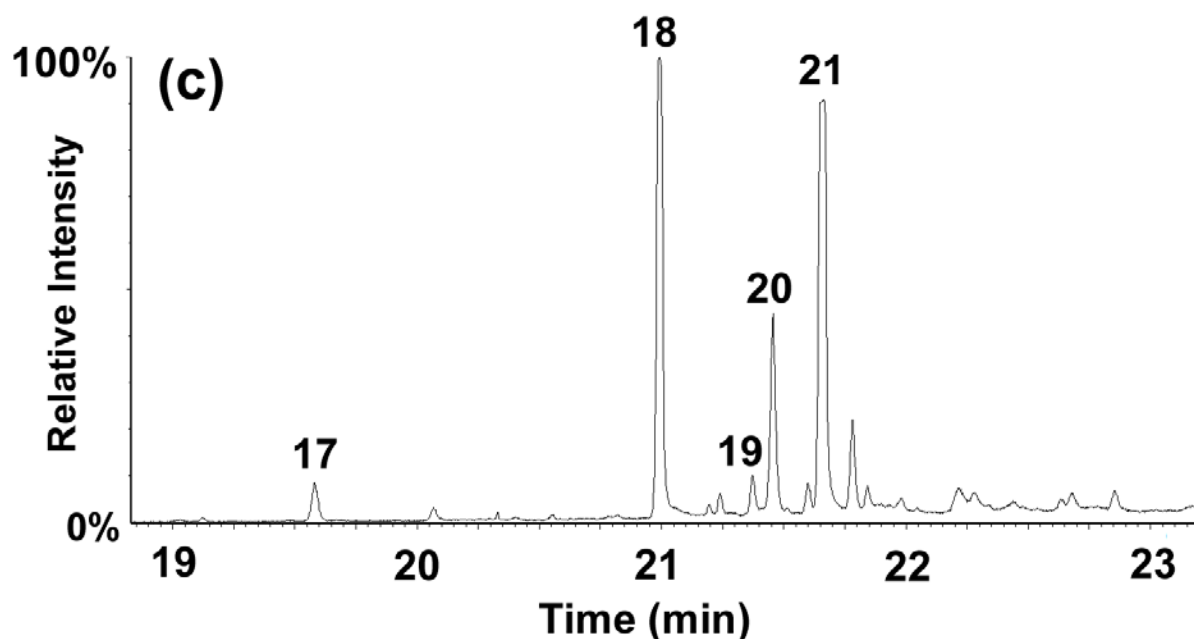
2.3.5 Effect of AFEX severity on amount of degradation products

Experiments were performed to determine the extent to which AFEX severity affected the amounts of degradation products including phenolic acids and amides (Table 2.3). To assess the effect of severity on the amounts of several important hydrolysis and ammonolysis products including coumaric and ferulic acid and their corresponding amide diferulates (Ac-Am and Di-Am diferulates), these products were quantified in corn stover extracts treated with different AFEX severities including regular AFEX, H-AFEX and L-AFEX.

A 5-minute LC gradient was developed and used instead of the 30-min gradient to provide for a cost-effective high-throughput analysis. Figure 2.8 compares the extracted ion chromatograms of ferulic and coumaric acids and their corresponding amides and diferulates from the 5-minute and the 30-minute methods. The 5-minute gradient preserves the baseline resolution of ferulic and coumaric acids and their amides and still separates their *cis-* and *trans-*

isomers. Although several of the diferulate peaks merged after shortening the gradient, unresolved diferulates peaks were still integrated, and their contributions included in the totals.. Increasing ammonia by three-folds and increasing residence time by 30 minutes compared to regular AFEX increased the amount of feruloyl amide, coumaroyl amide and diferulates (mostly Ac-Am and Di-Am) by almost two-fold (Table 2.3). The amount of ferulic acid remained unchanged and coumaric acid was decreased using the higher severity treatment. These findings suggested that ammonolysis increases with increasing ammonia load and residence time in the reactor. L-AFEX treatment, which was performed at 100 $^{\circ}$C (30 $^{\circ}$C lower than regular AFEX) generated almost half the amount of ammonolysis products including feruloyl and coumaroyl amides and diferuloyl amides compared to regular AFEX. Since temperature was the only variable in L-AFEX compare to regular AFEX, results shown that the extent of ammonolysis of cell wall ferulate and coumarate esters that occur during AFEX depends on temperature. Quantification of the same compounds was performed on extracts of AFEXTCS using a residence time of 45 min but keeping all other factors identical to regular AFEX conditions. The analytical results demonstrated that an additional 30 min residence time increased the phenolic amide amounts by 30-40% compared to regular AFEX.

Table 2.3. Important phenolic amides and acids formed upon ammonolysis and hydrolysis during AFEX with varying severities, measured using a fast 5-min LC/MS method. Amounts are µg analyte/g biomass for 5 replicates (n=5), the uncertainties are one standard deviation.

Compound	Regular AFEX 130 °C, 1:1 NH₃/Biomass, 15 min	H-AFEX 130 °C, 3:1 NH₃/Biomass, 45 min	L-AFEX 100 °C, 1:1 NH₃/Biomass, 15 min
Feruloyl amide	4.7 ± 0.40	7.6 ± 0.46	2.4 ± 0.07
Ferulic acid	0.11 ± 0.01	0.12 ± 0.02	0.09 ± 0.01
Coumaroyl amide	3.9 ± 0.14	6.9 ± 0.37	2.2 ± 0.08
Coumaric acid	1.2 ± 0.11	0.86 ± 0.09	1.1 ± 0.12
Diferuloyl amides	5.9 ± 0.52	14.8 ± 1.51	3.1 ± 0.22

H-AFEX also generated 30% higher acetamide and 18% lower acetic acid yield compared to regular AFEX. These findings show that the extent of hydrolysis of the arabinoxylan *O*-acetyl groups increased with AFEX severity. Since the mechanistic basics and reaction kinetics of ammonolysis and hydrolysis during AFEX are not yet clear, conclusions regarding exact correlation between severity conditions and the amount of hydrolysis and ammonolysis products cannot be drawn. Such mechanistic investigations would require quantification of key ammonolysis and hydrolysis products including acetic acid, acetamide and phenolic acids and amides upon changing AFEX severity variables including ammonia load, temperature, moisture content and residence time one at a time and at different levels. With enough statistical replicates (e.g. at least three); such studies generate several hundred samples that need fast and

reproducible analytical techniques for high-throughput screening of desired analytes. The developed methods mentioned in this chapter can be used to investigate relationships between AFEX conditions and yields of nitrogenous products, and can help guide process optimization.
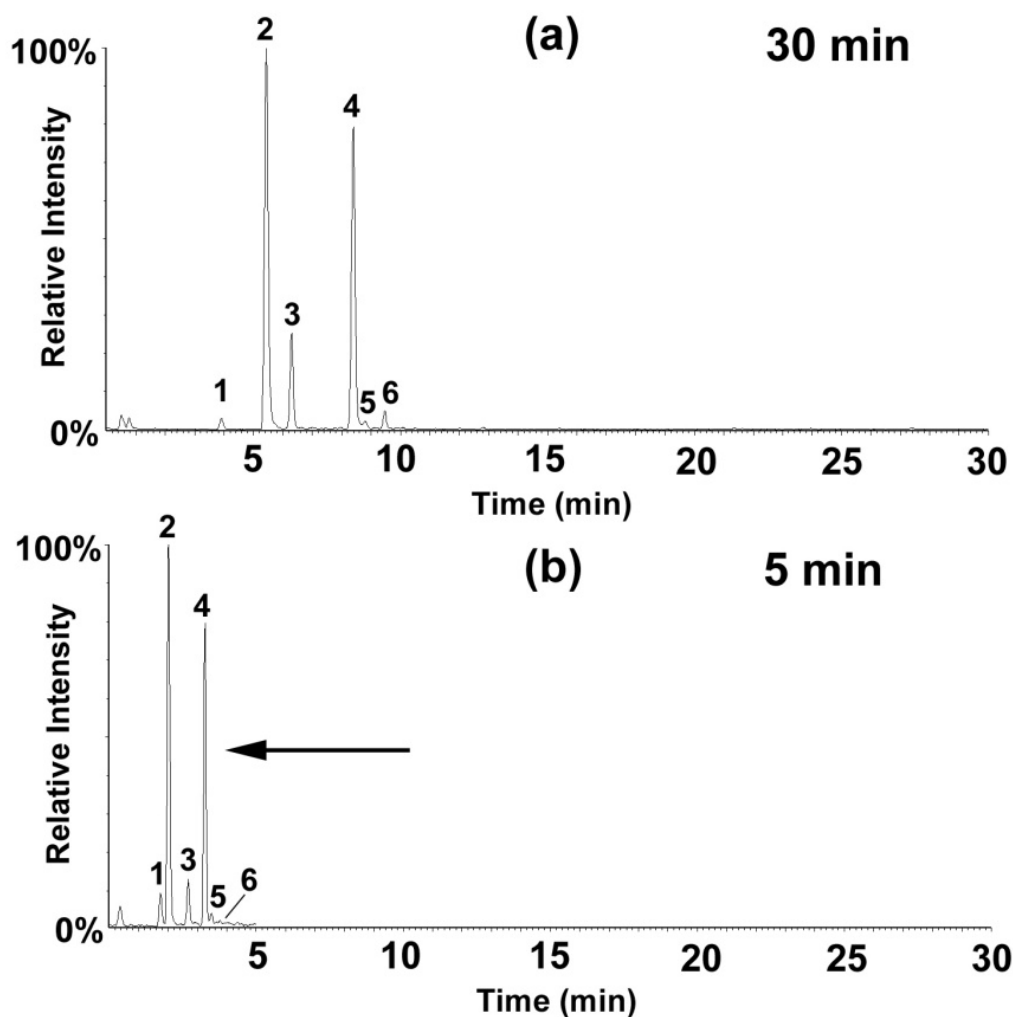
Figure 2.8. Combined extracted ion LC/MS chromatograms in ESI negative mode showing (a) and (b) ferulic (*m/z* 193) and coumaric (*m/z* 163) acids with their corresponding amides (*m/z* 192 and 162) and (c) and (d) Di-Am (*m/z* 383), Ac-Am (*m/z* 384) and Di-Ac (*m/z* 385) diferulates from AFEXTCS extract. Peaks are as follows: 1: *cis*-coumaroyl amide, 2: *trans*-coumaroyl amide, 3: *trans*-coumaric acid, 4: *trans*-feruloyl amide, 5: *cis*-feruloyl amide, 6: ferulic acid. Peaks in (a) and (c) are from 30 minute and peaks in (b) and (d) are from the 5-minute LC gradient

Figure 2.8. (cont'd)
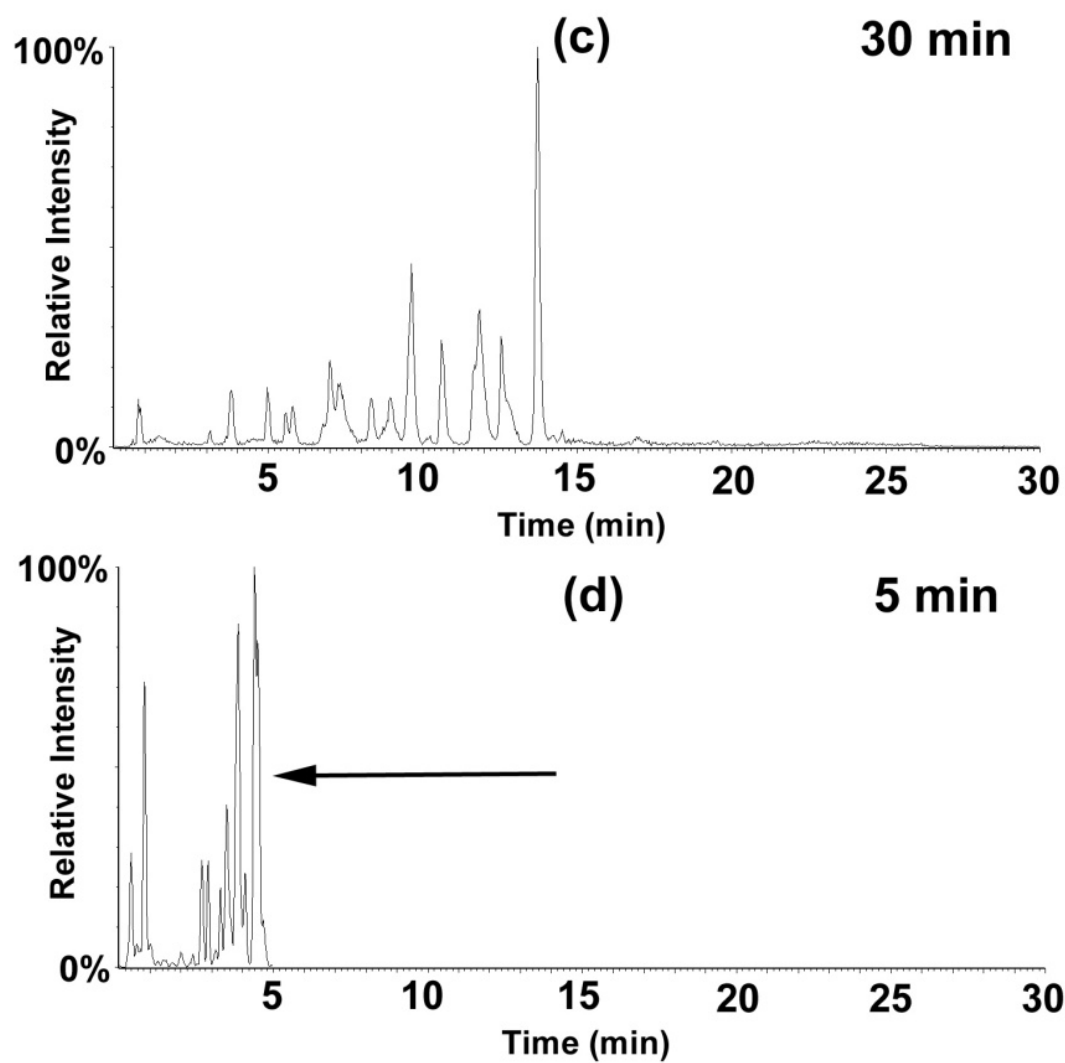
2.4 Conclusions

Comparisons of LC/MS profiles of AFEXTCS and UTCS extract showed that several abundant compounds distinguish treated from untreated corn stover extracts of which some are nitrogenous species present only in AFEXTCS. Many of these compounds are consistent with important roles of ammonolysis reactions during AFEX.

AFEX treatment of corn stover generates numerous water-soluble nitrogenous products totaling about 45 mg/g AFEXTCS. Of these products acetamide (25 mg/g AFEXTCS) derived from ammonolysis of acetate esters represented the largest nitrogen sink in the AFEX process followed by coumaroyl, feruloyl and diferuloyl amides (totaling 14.7 mg/g AFEXTCS). Total uncharacterized phenolic amides were about 2.6 mg/g AFEXTCS of which several are derived from ammonolysis cleavage of lignin carbohydrate complex linkages. Ammonolysis of corn stover cell wall esters and Maillard products formed upon reaction of ammonia with reducing sugars during AFEX accounted for about 45-50% of the total reacted ammonia; hence there are other nitrogenous species in the AFEXTCS extract that need to be characterized for complete mass balance of ammonia.

To establish correlations between amounts of degradation products and AFEX severity conditions, fast and reproducible analytical strategies are needed for high-throughput screening of key ammonolysis and hydrolysis products. A five-minute LC/MS gradient was developed in this work for such analyses, and this approach requires no sample preparation or derivatization unlike the majority of methodologies reported in the literature that are almost entirely based on derivatization and long GC/MS analyses for monitoring similar [2,42,43] analytes. The LC/MS method was shown to be useful for byproduct quantification and monitoring changes of several ammonolysis products generated after varying AFEX severity.

Development of cost-effective, economically feasible low severity thermochemical

pretreatments that minimize formation of biological inhibitors while maximizing cell wall decomposition is the main hurdle in production of bioethanol from lignocellulosic biomass. Fast and inexpensive analytical technologies will play important roles to optimize the pretreatment process and promote this goal.

REFERENCES

REFERENCES

(1)   Klinke, H. B.; Thomsen, A. B.; Ahring, B. K. *Applied Microbiology and Biotechnology* **2004**, *66*, 10-26.

(2)   Klinke, H. B.; Ahring, B. K.; Schmidt, A. S.; Thomsen, A. B. *Bioresource Technology* **2002**, *82*, 15-26.

(3)   Cantarella, M.; Cantarella, L.; Gallifuoco, A.; Spera, A.; Alfani, F. *Biotechnology Progress* **2004**, *20*, 200-206.

(4)   Jing, X. Y.; Zhang, X. X.; Bao, J. *Applied Biochemistry and Biotechnology* **2009**, *159*, 696-707.

(5)   Taherzadeh, M. J.; Eklund, R.; Gustafsson, L.; Niklasson, C.; Liden, G. *Industrial & Engineering Chemistry Research* **1997**, *36*, 4659-4665.

(6)   Assary, R. S.; Redfern, P. C.; Greeley, J.; Curtiss, L. A. *Journal of Physical Chemistry B* **2011**, *115*, 4341-4349.

(7)   Chundawat, S. P. S.; Venkatesh, B.; Dale, B. E. *Biotechnology and Bioengineering* **2007**, *96*, 219-231.

(8)   Meshartree, M.; Dale, B. E.; Craig, W. K. *Applied Microbiology and Biotechnology* **1988**, *29*, 462-468.

(9)   Sousa, L. D.; Chundawat, S. P. S.; Balan, V.; Dale, B. E. *Current Opinion in Biotechnology* **2009**, *20*, 339-347.

(10)   Lau, M. W.; Dale, B. E. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 1368-1373.

(11)   Mosier, N.; Wyman, C.; Dale, B.; Elander, R.; Lee, Y. Y.; Holtzapple, M.; Ladisch, M. *Bioresource Technology* **2005**, *96*, 673-686.

(12)   Chundawat, S. P. S.; Vismeh, R.; Sharma, L. N.; Humpula, J. F.; Sousa, L. D.; Chambliss, C. K.; Jones, A. D.; Balan, V.; Dale, B. E. *Bioresource Technology* **2010**, *101*, 8429-8438.

(13)   Knodt, C. B.; Williams, J. B.; Brumbaugh, J. *Journal of Animal Science* **1950**, *9*, 661-662.

(14)   Knodt, C. B.; Williams, J. B.; Brumbaugh, J. *Journal of Dairy Science* **1951**, *34*, 1042-1046.

(15)   Nishie, K.; Waiss, A. C.; Keyl, A. C. *Toxicology and Applied Pharmacology* **1969**, *14*, 301

(16)   Fairbrother, T. E.; Kerr, L. A.; Essig, H. W. *Veterinary and Human Toxicology* **1987**, *29*, 312-315.

(17)   Huang, H. Z.; Guo, X. Y.; Li, D. M.; Liu, M. M.; Wu, J. F.; Ren, H. Y. *Bioresource Technology* **2011**, *102*, 7486-7493.

(18)   Zheng, R. P.; Zhang, H. M.; Zhao, J.; Lei, M. L.; Huang, H. *Journal of Chromatography A* **2011**, *1218*, 5319-5327.

(19)   Scarlata, C. J.; Hyman, D. A. *Journal of Chromatography A* **2010**, *1217*, 2082-2087.

(20)   Du, B. W.; Sharma, L. N.; Becker, C.; Chen, S. F.; Mowery, R. A.; van Walsum, G. P.; Chambliss, C. K. *Biotechnology and Bioengineering* **2010**, *107*, 430-440.

(21)   Schell, D. J.; Farmer, J.; Newman, M.; McMillan, J. D. *Applied Biochemistry and Biotechnology* **2003**, *105*, 69-85.

(22)   Humpula, J. F.; Chundawat, S. P. S.; Vismeh, R.; Jones, A. D.; Balan, V.; Dale, B. E. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2011**, *879*, 1018-1022.

(23)   Stagliano, M. C.; DeKeyser, J. G.; Omiecinski, C. J.; Jones, A. D. *Rapid Communications in Mass Spectrometry* **2010**, *24*, 3578-3584.

(24)   Chen, S. F.; Mowery, R. A.; Scarlata, C. J.; Chambliss, C. K. *Journal of Agricultural and Food Chemistry* **2007**, *55*, 5912-5918.

(25)   Izydorczyk, M. S.; Biliaderis, C. G. *Carbohydrate Polymers* **1995**, *28*, 33-48.

(26)   Maslen, S. L.; Goubet, F.; Adam, A.; Dupree, P.; Stephens, E. *Carbohydrate Research* **2007**, *342*, 724-735.

(27)   Pauly, M.; Keegstra, K. *Plant Journal* **2008**, *54*, 559-568.

(28)   Saulnier, L.; Thibault, J. F. *Journal of the Science of Food and Agriculture* **1999**, *79*, 396-402.

(29)   Arrieta-Baez, D.; Dorantes-Alvarez, L.; Martinez-Torres, R.; Zepeda-Vallejo, G.; Jaramillo-Flores, M. E.; Ortiz-Moreno, A.; Aparicio-Ozores, G. *Journal of the Science of Food and Agriculture* **2012**, *92*, 2715-2720.

(30)   Morrison, T. A.; Jung, H. G.; Buxton, D. R.; Hatfield, R. D. *Crop Science* **1998**, *38*, 455-460.

(31)   Grabber, J. H.; Hatfield, R. D.; Ralph, J. *Journal of the Science of Food and Agriculture* **1998**, *77*, 193-200.
(32)   Grabber, J. H.; Hatfield, R. D.; Ralph, J.; Zon, J.; Amrhein, N. *Phytochemistry* **1995**, *40*, 1077-1082.

(33)   Grabber, J. H.; Ralph, J.; Hatfield, R. D. *Journal of Agricultural and Food Chemistry* **2000**, *48*, 6106-6113.

(34)   Hatfield, R.; Jones, B.; Grabber, J.; Ralph, J. *Plant Physiology* **1995**, *108*, 125-125.

(35)   Ralph, J. *Phytochemistry Reviews* **2010**, *9*, 65-83.

(36)   Ralph, J.; Bunzel, M.; Marita, J. M.; Hatfield, R. D.; Lu, F.; KIm, H.; Schatz, P. F.; Grabber, J.; Steinhart, H. *Phytochemistry Reviews* **2004**, *3*, 79-96.

(37)   Ralph, J.; Quideau, S.; Grabber, J. H.; Hatfield, R. D. *Journal of the Chemical Society-Perkin Transactions 1* **1994**, 3485-3498.

(38)   Shibamoto, T.; Akiyama, T.; Sakaguchi, M.; Enomoto, Y.; Masuda, H. *Journal of Agricultural and Food Chemistry* **1979**, *27*, 1027-1031.

(39)   Wong, J. M.; Bernhard, R. A. *Journal of Agricultural and Food Chemistry* **1988**, *36*, 123-129.

(40)   Benzingpurdie, L. M.; Ripmeester, J. A.; Ratcliffe, C. I. *Journal of Agricultural and Food Chemistry* **1985**, *33*, 31-33.

(41)   Wang, H. Y.; Qian, H.; Yao, W. R. *Food Chemistry* **2011**, *128*, 573-584.

(42)   Jonsson, L. J.; Palmqvist, E.; Nilvebrant, N. O.; Hahn-Hagerdal, B. *Applied Microbiology and Biotechnology* **1998**, *49*, 691-697.

(43)   Luo, C. D.; Brink, D. L.; Blanch, H. W. *Biomass & Bioenergy* **2002**, *22*, 125-138.

Chapter Three: Profiling of Diferulates (Plant Cell Wall Cross-Linkers) using Ultrahigh-performance Liquid Chromatography-Tandem Mass Spectrometry

3.1 Abstract

Recalcitrance of grasses to enzymatic digestion arises to a significant degree from a complex array of phenolic crosslinks between cell wall glycopolymer chains that inhibit their conversion to biofuels and lower their nutritive value for animal feed applications. Polysaccharide esters of ferulic acid are abundant in plant cell walls. Crosslinks between polysaccharides are formed through oxidative dehydrodimerization of ferulates, producing dehydrodiferulates (henceforth termed diferulates). Such ferulates and diferulates further crosslink plant cell walls by radical coupling cross-reactions during lignification.

Although cell wall digestibility can be improved by cell wall metabolic engineering, or post-harvest by various pretreatment processes, a more comprehensive understanding of the role and impact of ferulate crosslinking on polysaccharide hydrolysis would be accelerated by availability of a routine and universal method for identifying and quantifying diferulates. In this Chapter, an ultrahigh-performance liquid chromatography/tandem mass spectrometry (UHPLC/MS/MS) strategy for comprehensive separation and identification of diferulate isomers is presented. Collision-induced dissociation (CID) mass spectra of [M+H]$^+$ ions distinguished various isomers without requiring derivatization. Characteristic product ions for 8–O–4-, 8–8-non-cyclic, 8–8-cyclic, 8–5-cyclic, 8–5-non-cyclic, and 5–5-linked isomers were identified. All diferulates were identified either as di-acids in extracts of NaOH-hydrolyzed corn stover, or as a diverse group of diferulate mono- and di-amides in extracts of ammonia fiber expansion (AFEX)-treated corn stover. This approach allows for direct analysis of released diferulates with minimal sample preparation, and is suitable for high-throughput profiling and correlating pretreatment conditions with biomass digestibility in a biorefinery producing biofuels and biochemicals.

3.2 Introduction

Grasses serve society in myriad ways, providing food in the form of grain (e.g., rice, wheat, maize, and oats), and nutrition for livestock as grain and/or silage. Society has turned increasing attention to developing sustainable resources to provide biomass for conversion to renewable liquid fuels. In light of increasing worldwide demand for food, and the escalating costs and environmental impacts of liquid fossil fuels, improvement of grasses for biomass production, pest resistance, and conversion to bioenergy products will rely on better understanding the chemistry of grass cell walls and the genes that regulate this chemistry. The exciting developments in plant engineering, DNA sequencing, and metabolic profiling [1], offer prospects for substantial improvements in the production of sustainable renewable plant resources that provide high-value products.

The chemistry of cell walls in grasses governs important functions including the regulation of plant growth, mechanical strength, resistance to pathogens and insects, and cell wall degradation [2]. These cell walls consist of about one-third hemicelluloses (total cell wall mass fraction basis) including various (glucurono-arabino) xylans, glucans, and xyloglucans, roughly one-half cellulose, with the balance coming from other biopolymers (notably lignin), polyphenol derivatives, ash, and extractives [3].

Recalcitrance of cell walls to enzymatic digestion presents the major hurdle to economic development of grass-derived biomass in applications ranging from animal nutrition to biofuels [4]. A substantial part of this recalcitrance arises from crosslinking of cell wall arabinoxylans via dehydrodimerization of ferulate esters that produces stable ether and carbon-carbon bonds [5-8]. Although this crosslinking presents a substantial obstacle to enzymatic digestion [9,10], thermochemical pretreatments (including AFEX) cleave these ester linkages, releasing diferulates (or derivatives) and arabinoxylans, allowing for increased sugar conversion [11,12].

90

As diferulates act as cell wall cross-linkers, identification and quantitative profiling of diferulates released by chemical pretreatment, or by relevant enzymes such as esterases, provides a revealing and important measure of the role of cell wall crosslinking on wall deconstruction efficiency, and serves as a key metric for guiding process optimization.

Cross-linked arabinoxylans are a major polysaccharide constituent of grass hemicelluloses [19,20], as illustrated in Figure 3.1. Their structure is based on β-D-(1,4)-linked xylose monomers that are substituted with arabinosyl groups as side-chains on every 2-3 xylose units [13]. The side-chain arabinosyl C-5 hydroxyl groups are often acylated by the hydroxycinnamates ferulate and *p*-coumarate [13,21]. Feruloylated oligosaccharides have been isolated and identified from cell walls of maize [22,23], barley straw [24], sugar cane bagasse [25], and wheat bran [26]. Oxidative dehydrodimerization (simply termed dimerization from here on) of arabinoxylan ferulate esters forms covalent crosslinks between arabinoxylan chains, unambiguously documented with the isolation of a diferulate from bamboo shoots in which each ferulate moiety acylated an arabinoxylan fragment [27]. A more recent investigation detected pentose-diferulate linkages in corn fibers [28]. Diferulate ester crosslinks are distributed among diverse chemical forms because they are formed by free-radical coupling at the O–4-, 5-, and 8-positions to form dimeric structures annotated as 5–5-, 8–5-, and 8–8- (formed through C-C bonds) and 8–O–4- and 4–O–5- (formed through ether bonds) coupled diferulates as a total of 7 isomers plus additional chemical forms [3,29,30]. It has been suggested that different isomers of diferulate crosslinks may have distinct physiological functions [2], but control over chemical radical coupling reactions makes this an unlikely plant strategy [5]. Regardless, the scarcity of
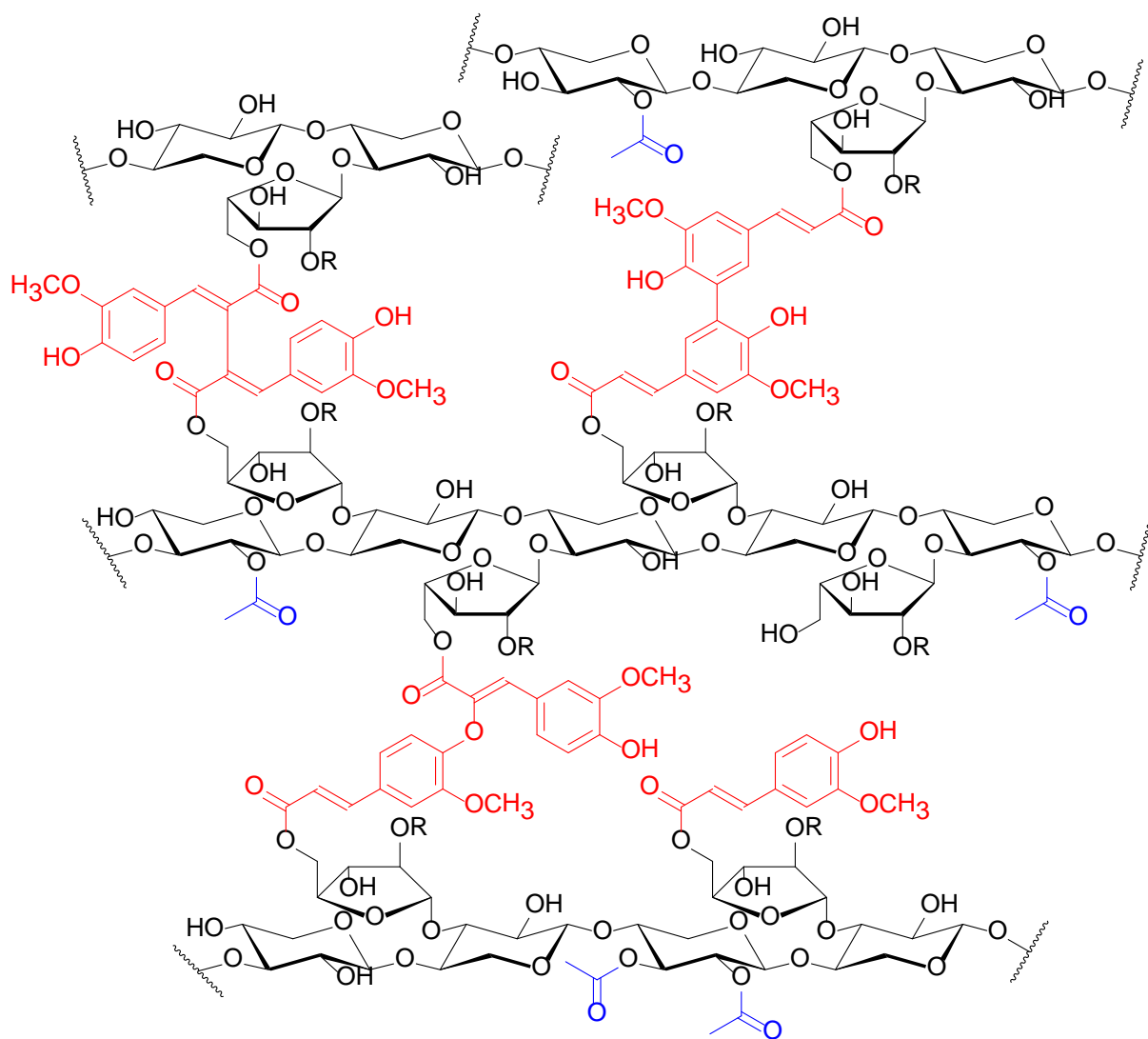
Figure 3.1. A schematic of grass cell wall hemicelluloses, showing crosslinking of arabinoxylans by diferulates. This is a schematic model showing key features of grass cell walls. There are many more arabinosyl (Ara) units in arabinoxylans without ferulate/diferulate substitution; there are also other substitutions such as with glucuronate units not shown. Note that Ara substitution has been shown at the 3-position of xylosyl (Xyl) residues (where it is most frequently found)[13], and that ferulate is invariably on the primary (C5) OH of Ara units. The model shows cross-linking of the arabinoxylan chains by 5–5-, 8–5-, 8–O–4-, and two forms of 8–8-diferulates. Acetyl substitution on the 2-, 3-, and 2,3-positions of xylan units is also shown. Finally, note that Ara units branching off the xylan chain may themselves have xylosyl substitution (usually/invariably at C2) – hence the R = H, Xyl designation here [14-18], R=H or Xyl.

92

accurate quantitative information about abundances of specific diferulate links in various cells and tissues remains underexplored. Recognition and quantitation of diferulate isomers released from cell walls has been non-routine because of limited availability of authentic standards [31], even though syntheses have been reported [6,32]. To date, high-performance liquid chromatography with ultraviolet (HPLC-UV) spectroscopic detection and gas chromatography-mass spectrometry (GC-MS) have been the main analytical tools for identification and quantification of diferulate isomers from hydrolyzed cell walls [31,33]. HPLC with UV detection lacks selectivity, as phenolic compounds may co-elute. GC/MS provides an alternative approach, but requires derivatization to confer sufficient volatility needed for GC separations. The requirement that solvents be removed before derivatization may result in irreproducible and/or poor derivatization yields, and serves as a barrier to high-throughput sample analyses needed for process optimization.

In addition, the 70 eV electron ionization (EI) mass spectra of diferulate trimethylsilyl derivatives are characterized by molecular ions that are frequently of low abundance, and present few abundant fragment ions that distinguish isomers [34]. Avoidance of derivatization provides important practical advantages for diferulate analyses, and is compatible with LC/MS/MS. However, collision-induced dissociation MS^2 spectra have yet to be reported for the entire suite of diferulate isomers. In a related paper, Morreel and co-workers published negative-mode collision-induced dissociation (CID) spectra of dilignols, a class of compounds with close similarity to diferulates, but lacking carboxylic acid groups [35,36]. Although the behavior of some linkage types between monolignols were investigated, the chemical structures are sufficiently different from diferulates that the monolignol CID spectra cannot be extrapolated to distinguish diferulate isomers. Four diferulates were also detected in extracts of alkali-treated

93

maize grain using LC/atmospheric pressure chemical ionization MS in positive-ion mode. This approach distinguished the four isomers largely based on abundances of in-source fragments arising from losses of 1-2 water molecules, but did not present MS/MS spectra [37]. If profiling of the multitude of isomers from biomass hydrolysates was not already sufficiently challenging, pretreatments such as ammonia fiber expansion (AFEX) [38,39], which can cleave diferulate bridges through combinations of hydrolysis and ammonolysis [40], (Figure 3.2) make the task yet more formidable. In ammonia-based biomass treatments, the two diferulate ester groups can be converted to either their amide or carboxylic acid forms, and this multiplies the complexity of released product isomers. In this Chapter, we have assembled a library of tandem mass spectra of protonated diferulic acid (Di-Ac), diamide (Di-Am), and bifunctional acid-amide (Ac-Am) dimers, with the intent that the MS^2 spectra will be helpful for distinguishing released dimers when authentic standards are not available for comparisons. When isomer-specific fragment ions are formed, these can provide the foundation for rapid LC/multiple reaction monitoring profiling of released diferulates. In addition, this report also describes an ultrahigh-performance liquid chromatography (UHPLC) tandem mass spectrometry approach for separation of modified and unmodified diferulates that does not require chemical derivatization. Extracts from corn stover treated with aqueous NaOH or the AFEX process were characterized to demonstrate the application of this methodology.
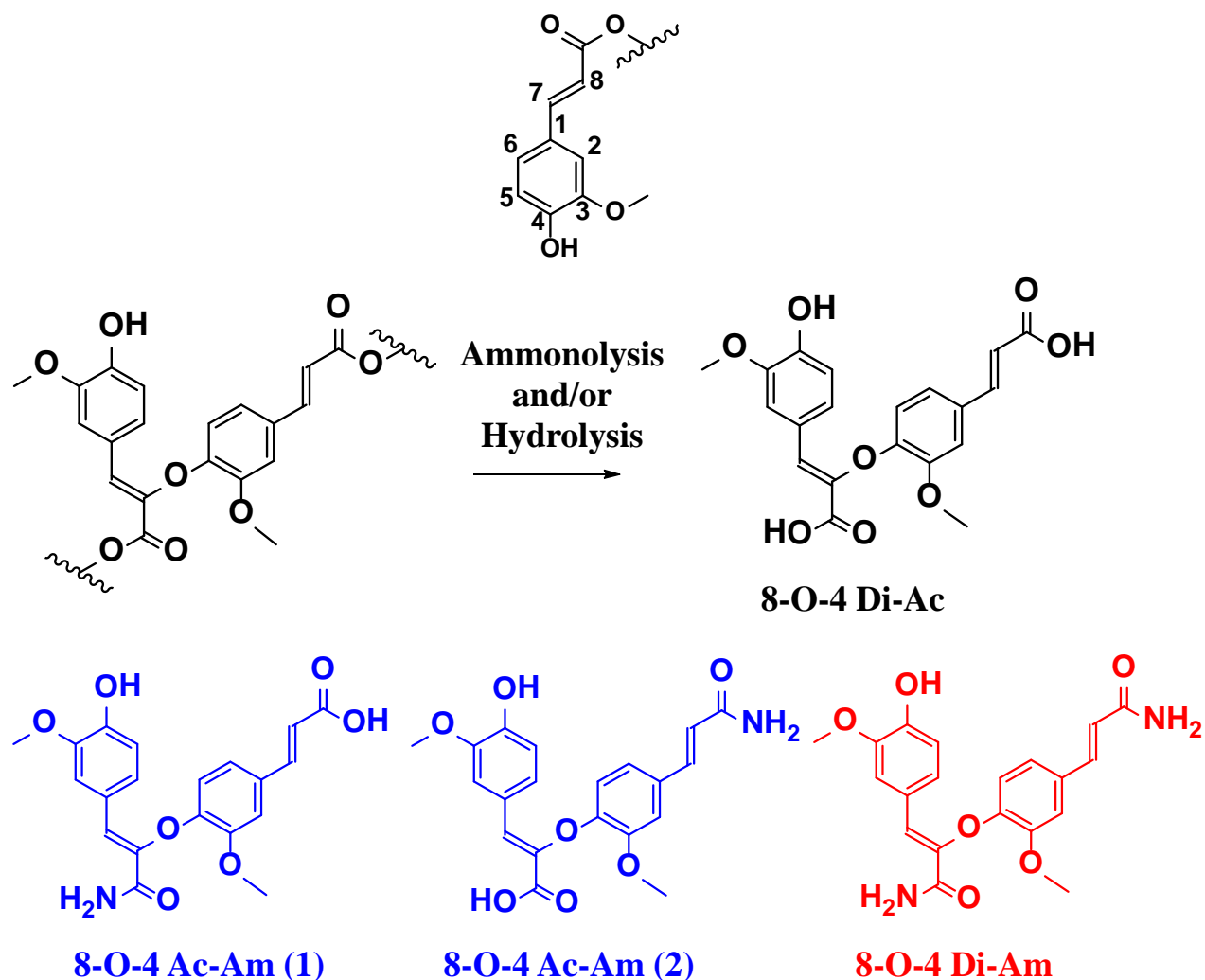
Figure 3.2. Structures of anticipated diacid (Di-Ac, black), acid-amide (Ac-Am, blue), and diamide (Di-Am, red) products of ammonolysis and hydrolysis of 8–O–4-, 8–8NC-, 8–8C-, 8–5-, and 4–O–5-isomers of plant cell wall diferulates during ammonia-based biomass pretreatment using AFEX process. Note that only a single 8–5-diferulate (the cyclic phenylcoumaran structure) is found in the wall, but that various non-cyclic isomers arise following hydrolysis or ammonolysis.

Figure 3.2 (cont'd)

8-8NC Di-Ac

8-8NC Di-Ac

8-8NC Ac-Am

8-8NC Di-Am

8-8C Di-Ac

8-8C Ac-Am (1)

8-8C Ac-Am (2)

8-8C Di-Am

Figure 3.2 (cont'd)

8-5C Di-Ac

8-5C Ac-Am (1)   8-5C Ac-Am (2)   8-5C Di-Am

8-5NC Di-Ac

8-5NC Ac-Am (1)   8-5NC Ac-Am (2)   8-5NC Di-Am

97

Figure 3.2 (cont'd)

Ammonolysis and/or Hydrolysis

5-5 Di-Ac

5-5 Ac-Am

5-5 Di-Am

Ammonolysis and/or Hydrolysis

4-O-5 Di-Ac

4-O-5 Ac-Am (1)

4-O-5 Ac-Am (2)

4-O-5 Di-Am
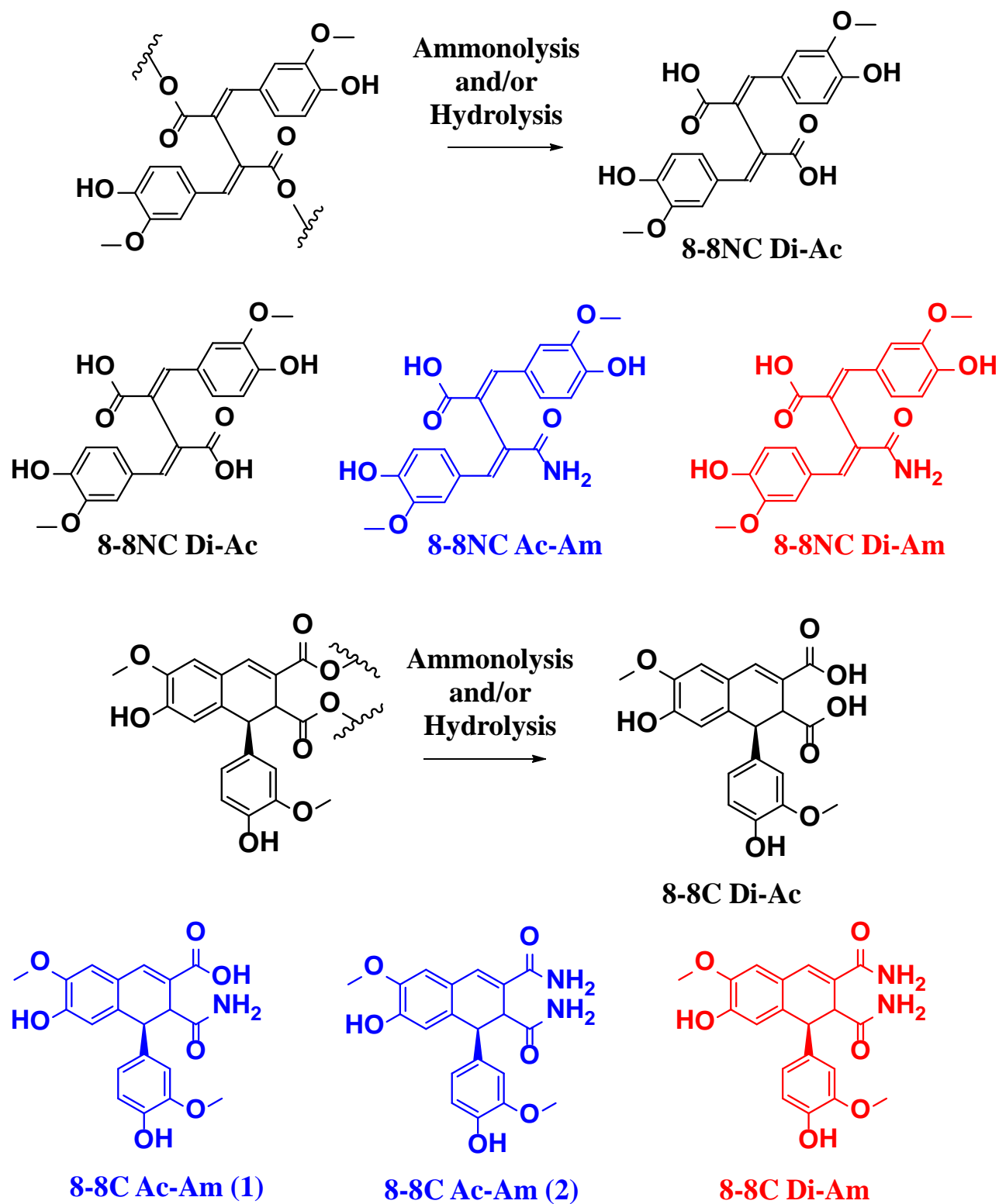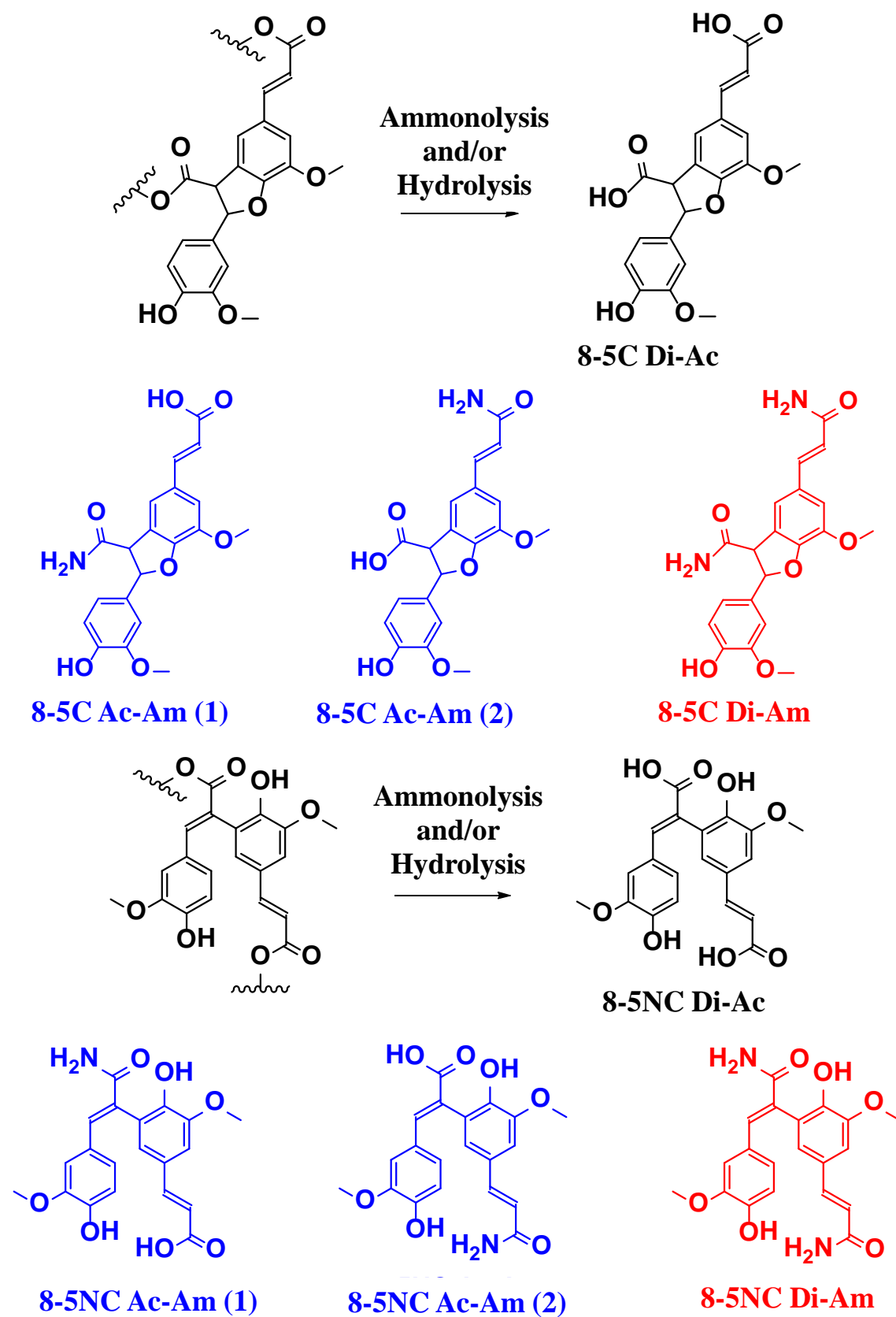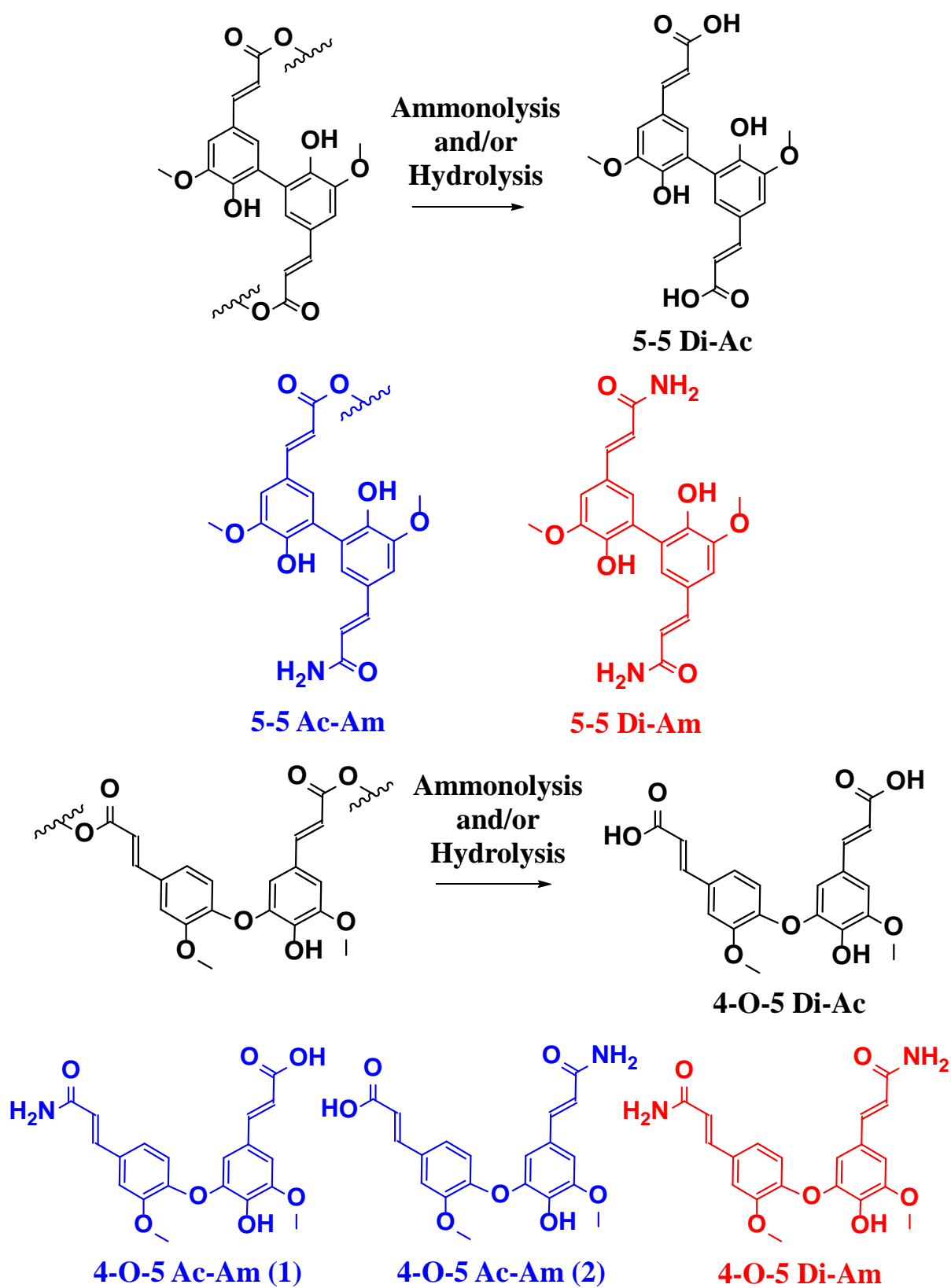
3.3 Experimental

3.3.1 AFEX pretreatment

Corn stover (biomass) was pretreated in a stainless steel high pressure reactor (Parr Instrument Company, Moline, IL) with liquid ammonia at 130 $^\circ$C, 0.6:1 water to biomass loading and 1:1 NH_3 to biomass loading (w/w) for 15 min total residence time. Detailed protocols can be found elsewhere [12].

3.3.2 NaOH pretreatment of corn stover

One gram of untreated corn stover was pretreated with 2 mL of 2 M NaOH at 25 $^\circ$C for 20 h as previously suggested [41]. Afterwards, the reaction mixture was brought to pH 7 with formic acid, and aliquots of the mixture were analyzed without further processing and dilution.

3.3.3 Synthesis of diferulic acid ester standards

Diferulate esters (8–O–4, 8–8NC, 8–8C, 8–5NC, 8–5C, and 5–5 isomers) were synthesized from ethyl ferulate via oxidative coupling catalyzed by copper(II)-tetramethylethylenediamine complex in acetonitrile as recently reported [32].

3.3.4 Synthesis of Ac-Am and Di-Am diferulates

To generate various Di-Am and Ac-Am derivatives, 200 µL of 15 M NH_4OH was added to 1 mg of each of the ethyl diferulate esters in 200 µL of dichloromethane solution, and each mixture was stirred at 25 $^\circ$C for 16 h. Reactions were quenched by addition of 1 mL water and 1 mL dichloromethane, and the organic layer was collected. Aliquots were diluted 100-fold in acetonitrile before UHPLC/MS analysis.

3.3.5 Separation and analysis of synthetic diferulates and pretreatment byproducts using UHPLC/MS and UHPLC/MS/MS

Diferulates were profiled using separation on an Ascentis Express C18 column with either UHPLC/TOF MS and multiplexed collision-induced dissociation [42] on a Waters LCT Premier or MS/MS on an AB/Sciex QTRAP 3200 mass spectrometer, using N_2 as collision gas in the latter, at 3.3×10^{-5} Torr.

The LC/MS system consisted of two Prominence LC-20AD (Shimadzu Corp.) pumps connected to a QTRAP 3200 hybrid-linear ion trap mass analyzer (Applied Biosystems) equipped with an electrospray ionization source (ESI). Diferulic acid standards, synthesized Ac-Am and Di-Am diferulates, and NaOH-hydrolyzed and AFEX-treated corn stover extracts were analyzed directly using a fused core Ascentis Express C18 (50×2.1 mm; 2.7 µm particles) column (Supelco, USA). Binary gradient elution was used with 0.15% aqueous formic acid (A) and methanol (B) under the following conditions and linear gradients: 0-1 min 5% B, linear increase to 30% B at 20 min; linear 40% B at 25 min; sudden increase to 99% B at 25.01 min and held until 27 min; returned to initial conditions at 27.01 min and held until 30 min. Injection volume, flow rate and column temperature were 5.0 µL, 0.4 mL/min and 50 °C.

Fragment ions were generated in both positive and negative ion modes for $[M+H]^+$ and $[M-H]^-$ ions using collision induced dissociation (CID) with nitrogen as collision gas, using enhanced product ion scanning. Major product ions generated at collision energy of 20 eV were subjected to a second CID stage in MS^3 scans. Accurate mass measurements were performed using a Waters LCT Premier orthogonal acceleration time-of-flight (TOF) mass spectrometer with the same chromatography system and multiplexed non-selective collision induced dissociation through three collision potentials (15, 45, and 70 V) with spectrum acquisition of 0.2 s/function.

AFEX treated corn stover was extracted with hot water (20 mL/g at 60 $^{\circ}$C) for 2 h. The mixture was cooled and filtered through Whatman filter paper (grade 40, pore size 8 µm). Filtrates were centrifuged for 10 min at 10,000g, and supernatant was used directly for LC/MS analysis via injection with no further processing. Electrospray ionization in negative mode was employed for analysis of process byproducts. Nontargeted peak detection, integration, and retention time alignment was performed using Waters MarkerLynx (V4.1) software.

3.4 Results and discussion

3.4.1 Complexity of treated biomass extract

UHPLC/TOF MS profiling of constituents in an aqueous extract of AFEX-treated corn stover reveals a rich mixture composed largely of cell wall-derived aromatic substances (Figure 3.3-a). Automated peak extraction and integration yielded more than 2000 features after deisotoping, including numerous phenolic compounds [12], as expected from a plant extract. In anticipation of observing diferulate derivatives, extracted-ion chromatograms were generated for ions corresponding to protonated Di-Ac, Ac-Am and Di-Am diferulates, yielding more than 30 chromatographic peaks with these nominal masses. Of these, 15 compounds gave masses within 5 ppm of theoretical values for diferulate, Di-Ac, Ac-Am or Di-Am forms of diferulates. These information are provided in Supplemental Information (SI)-Table 3-S3 at the end of this Chapter.

3.4.2 Diferulate isomers can be distinguished from their CID MS^2 spectra

To determine whether isomeric diferulates can be distinguished by their CID-generated fragment ions, MS^2 spectra of product ions derived from [M+H]$^+$ of synthetic diferulic acid isomers were generated using 20 eV collision energy (Figure 3.3-b-3.3-g) using the QTRAP mass spectrometer. MS^2 spectra of the diferulate synthetic standards in negative-ion mode (on

[M-H]$^-$) are provided in SI-Figure 3-S6. MS^2 spectra differentiated all isomers in both positive and negative ion modes, often displaying unique fragments from specific isomers. In other cases, relative abundances of fragment ions distinguished isomers. These observations set the basis for isomer differentiation using CID spectra. Although phenolic compounds ionize more efficiently in electrospray negative-ion mode owing to their acidity, fragmentation in positive-ion mode yielded more extensive cleavage of bonds that aids structural elucidation.
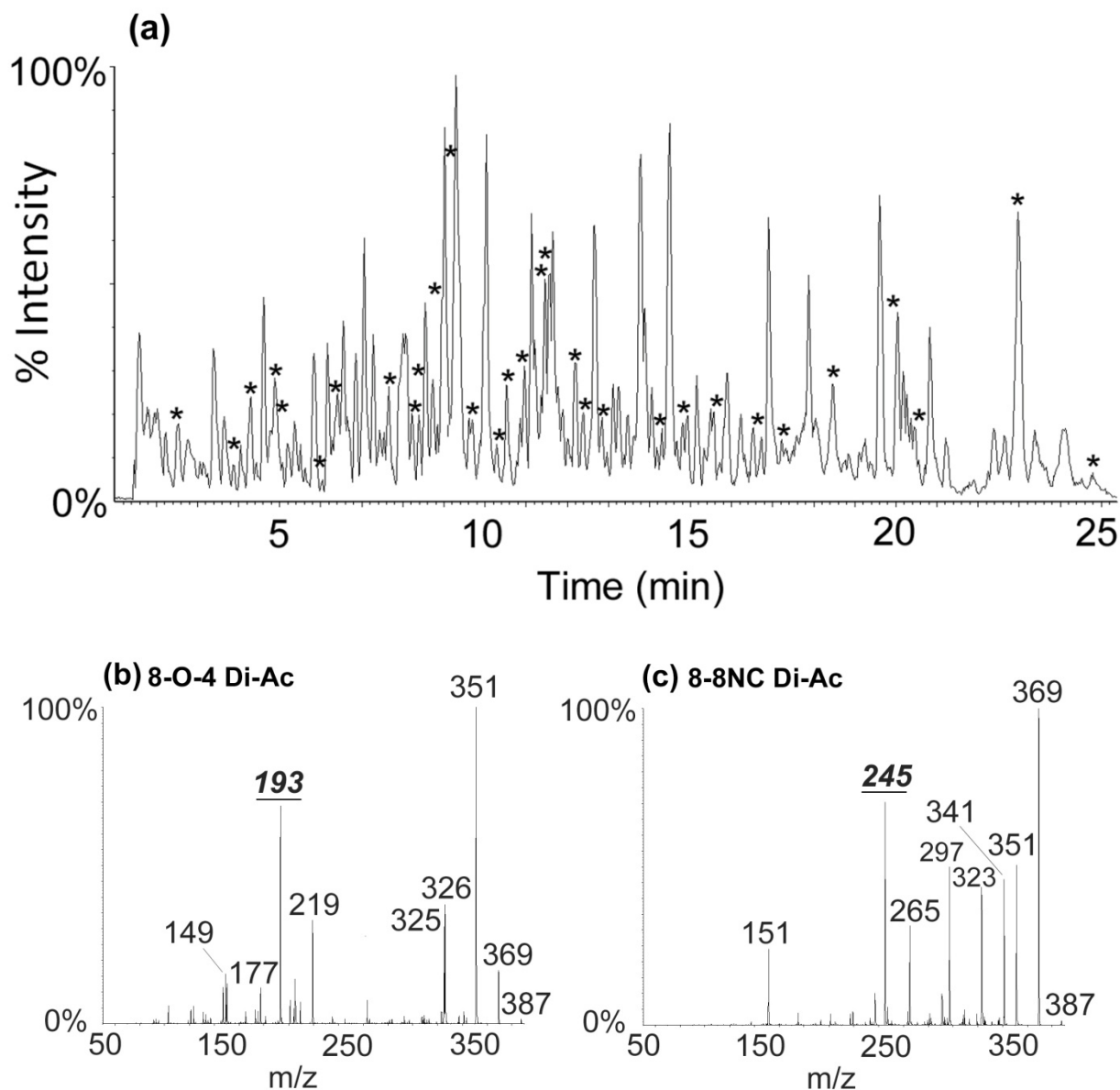
Figure 3.3. (a) UHPLC/TOF-MS total ion chromatogram of aqueous extract of AFEX-treated corn stover using electrospray ionization in positive-ion mode. Peaks labeled with asterisks (*) have the same nominal masses of [M+H]$^+$ ions as one of the three forms of released diferulates (Ac-Am, Di-Ac or Di-Am); (b-g) CID MS^2 spectra of diferulic acid standards and (h-j) CID MS^2 spectra of Di-Ac, Ac-Am and Di-Am synthetic standards for 8–O–4-diferulate. Precursor ions were [M+H]$^+$ at *m/z* 387, 386, and 385 for Di-Ac, Ac-Am, and Di-Am, and all CID mass spectra were generated on a QTRAP 3200 mass spectrometer using a collision energy of 20 eV.

103

Figure 3.3 (cont'd)

(d) 8-8C Di-Ac

(e) 8-5C Di-Ac

(f) 8-5NC Di-Ac

(g) 5-5 Di-Ac

(h) 8-O-4 Di-Ac

(i) 8-O-4 Ac-Am (1)

Figure 3.3 (cont'd)

(j) 8-O-4 Di-Am

8-O-4

[M+H]$^+$
387 → (193) $\xrightarrow{-H_2O}$ 175 $\xrightarrow{-CO}$ 147

(193) $\xrightarrow{-CO_2}$ 149

387 $\xrightarrow{-H_2O}$ 369

351 $\xleftarrow{-H_2O}$ 369 $\xrightarrow{-CO\ -CH_3^{\cdot}}$ 326

369 $\xrightarrow{-CO_2}$ 325

351 → 263 → 219 → 204, 201, 177

8-8 NC

[M+H]$^+$
387

387 $\xrightarrow{-H_2O}$ **369**

323 $\xleftarrow{-CO}$ 351 $\xleftarrow{-H_2O}$ **369** $\xrightarrow{-CO}$ 341

341 $\xrightarrow{-H_2O}$ 323

341 $\xrightarrow{-CO_2}$ 297 $\xrightarrow{-CH_3OH}$ 265

341 → 151

369 $\xrightarrow{-C_7H_8O_2}$ (**245**)

8-8 C

[M+H]$^+$
387

387 $\xrightarrow{-H_2O}$ **369**

323 $\xleftarrow{-CO}$ 351 $\xleftarrow{-H_2O}$ **369** $\xrightarrow{-CO}$ **341**

369 $\xrightarrow{-CO_2}$ 325

341 $\xrightarrow{-H_2O}$ 323

341 $\xrightarrow{-CO_2}$ (**297**) $\xrightarrow{-CH_3OH}$ 265 $\xrightarrow{-CO}$ 237
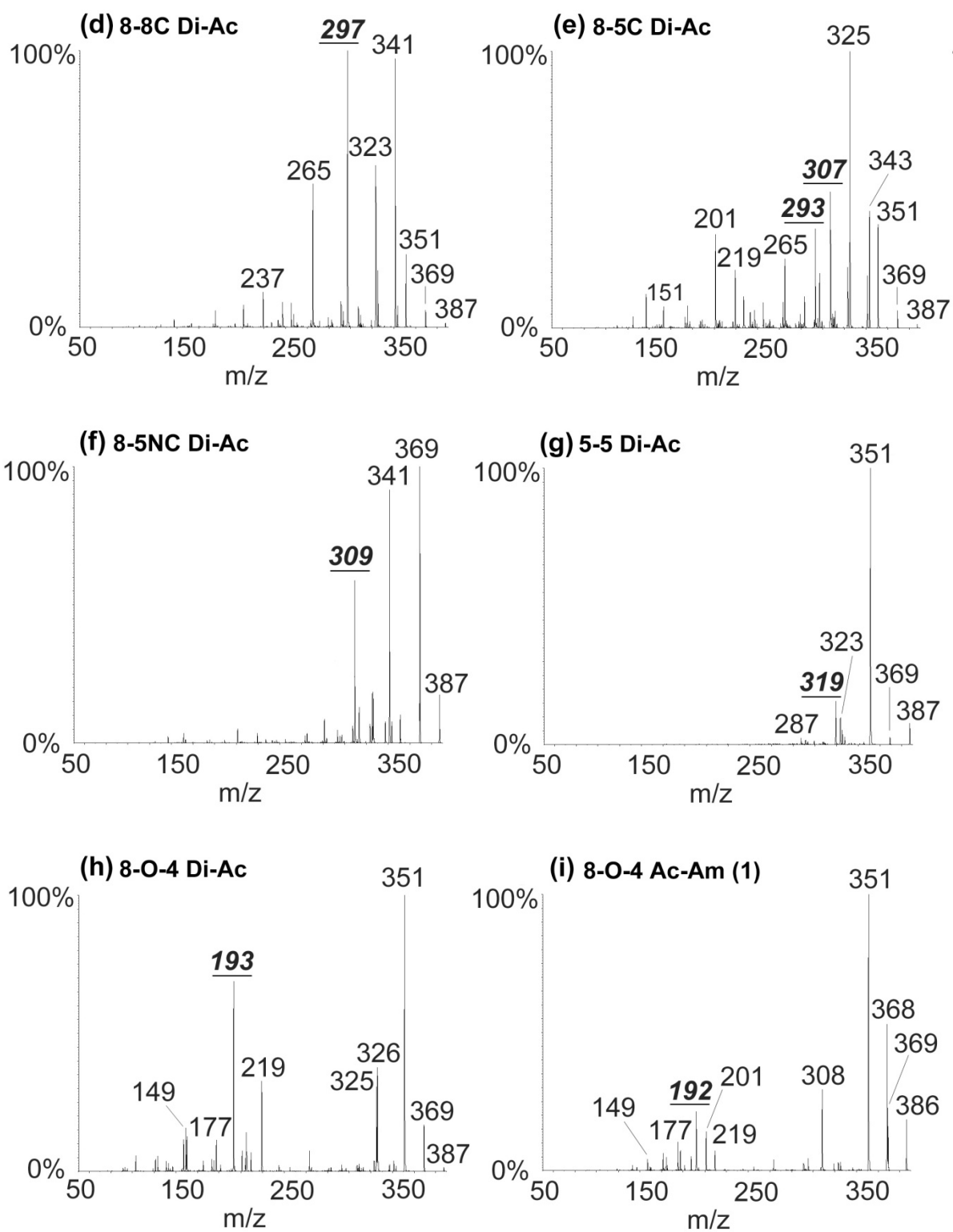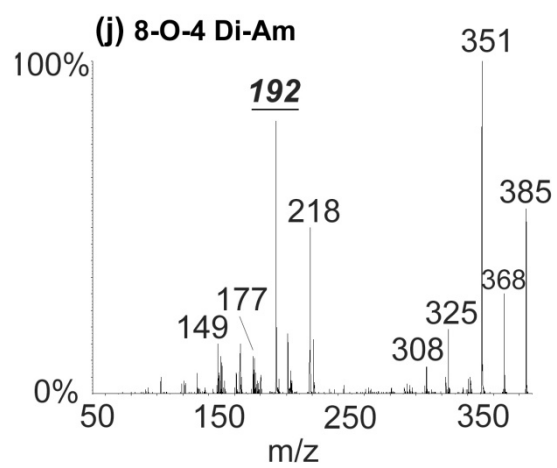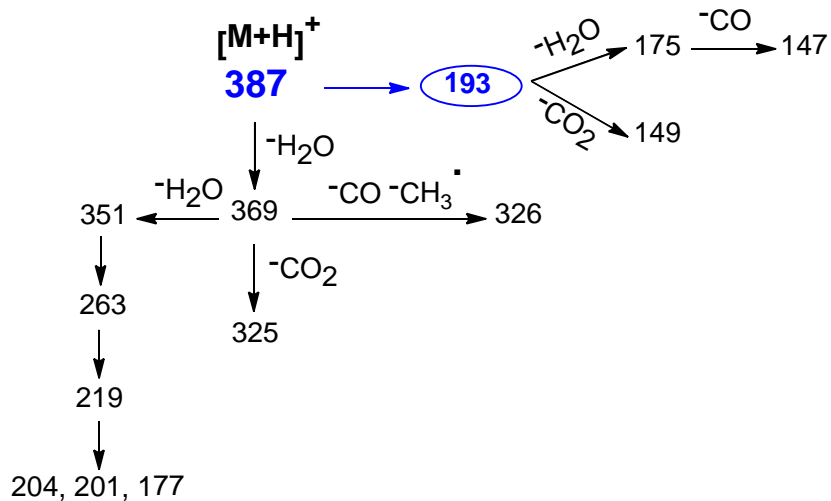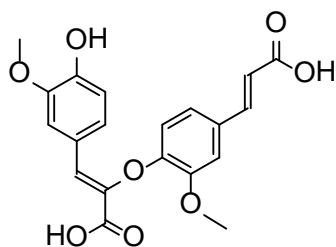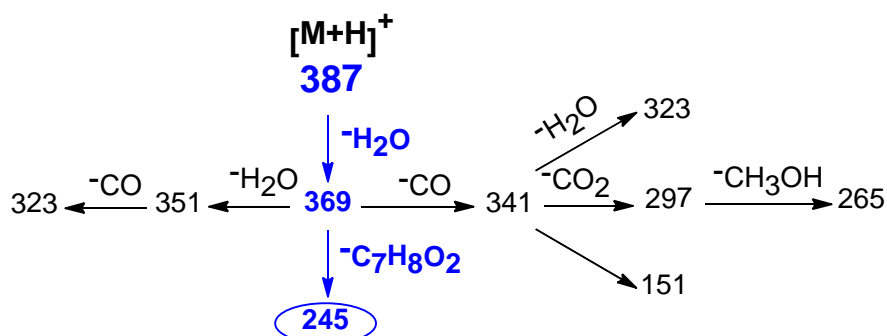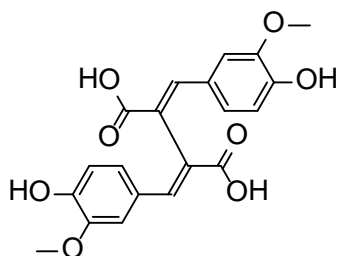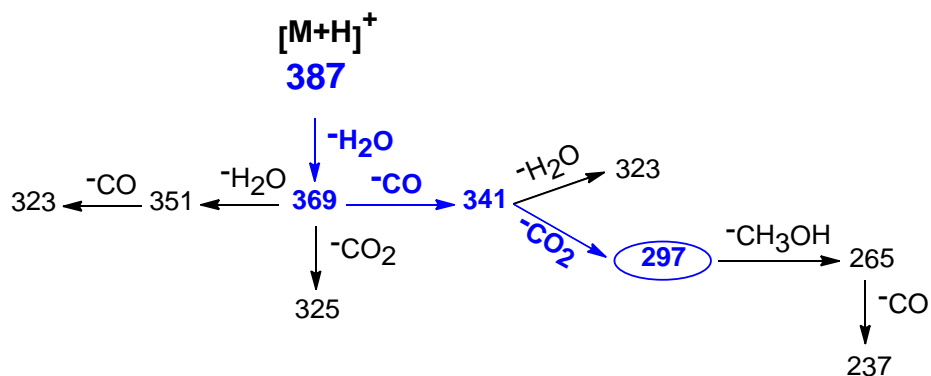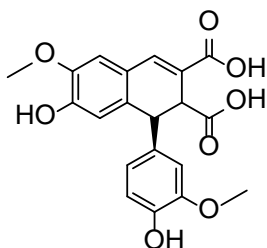
Figure 3.4. Summary of fragment ion annotation for CID MS/MS spectra of diferulic acid isomers. Assignments and proposed fragmentation pathways are based on accurate fragment masses measured using UHPLC/TOF MS with non-selective CID and from MS/MS/MS spectra generated on a QTRAP mass spectrometer. Each isomer-distinguishing fragment ion is highlighted inside an ellipse.

Figure 3.4 (cont'd)

8-5 C
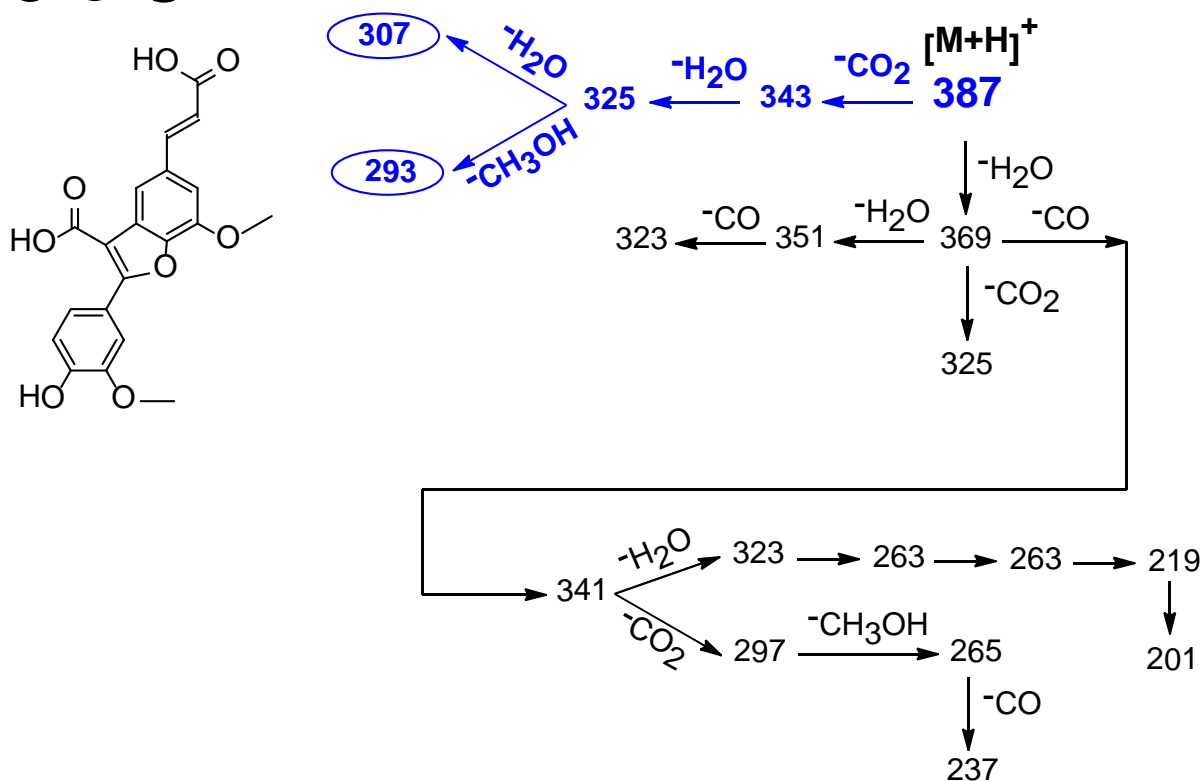
8-5 NC
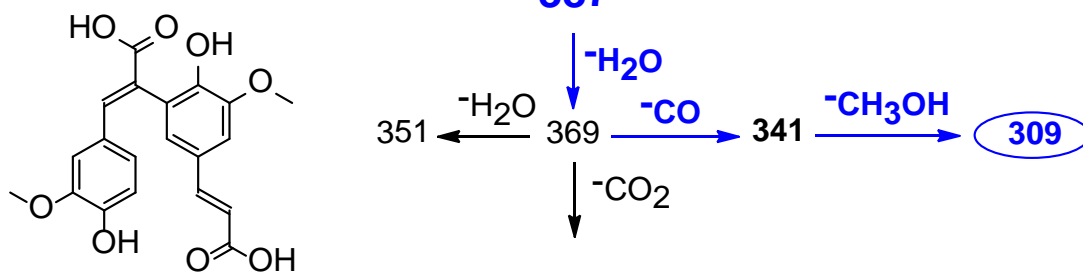
Figure 3.4 (cont'd)

5-5
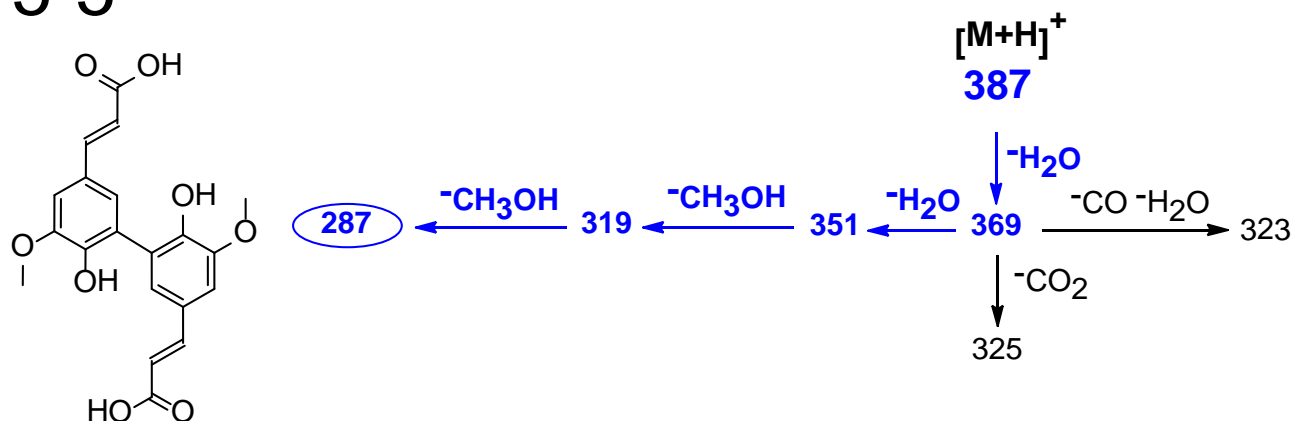
3.4.3 Fragment ions derived from side-chain neutral losses

For the purpose of this discussion, side-chains on diferulates are defined as all substituents attached to rings, such as carboxylic acid groups, methoxyl groups, and oxidized propyl moieties. Positive-mode CID mass spectra of all diferulic acid isomers exhibit fragment ions (Figure 3.3 and SI-Table 3-S1) corresponding to losses of 1 or 2 neutral water molecules (*m/z* 369 and 351). Most isomers underwent subsequent losses of CO, with the 8–O–4 isomer standing out as an exception, exhibiting an even-mass fragment (*m/z* 326) corresponding to losses of neutral CO and a CH_3 radical. This was the only isomer to yield significant amounts of radical fragment ions. Significant loss of CO_2 (-44 Da) from $[M+H]^+$ was only observed for the 8–5C isomer, but combined losses of CO_2 plus other side-chain-derived neutrals were common for all isomers except the 5–5-isomer. The aromatic methoxyl group participated in fragmentation, with fragments corresponding to losses of various side-chain components plus CH_3OH observed in all isomers. Although losses of CH_3OH from aromatic methoxyl groups are unusual in 70 eV EI mass spectra, they have been observed in CID spectra of protonated methyl ethers of the flavonoid myricetin [43].

3.4.4 Characteristic fragment ions that distinguish diferulic acid isomers

Each diferulic acid isomer standard yielded fragments upon CID that allowed for isomer discrimination, and these are highlighted with bold italic labels in the spectra (Figure 3.3b-g). In one case, the 8–O–4-isomer, the fragment at m/z 193 corresponds to cleaving the dimer in half, i.e., from cleavage of the ether link between monomeric units. In contrast, the 5–5-isomer yielded a characteristic fragment at m/z 319, derived from losses of two water molecules from the carboxylic acid groups and CH_3OH from the aromatic methoxyl group. This isomer did not yield significant fragments below m/z 250, demonstrating a resistance to dissociation of its core structure. The two 8–8-isomers were distinguished by two prominent fragments. For isomer 8–8NC, the fragment at m/z 245 is attributed to an unusual elimination of one water and methoxyphenol (guaiacol) that involves rearrangement and cleavage of a bond attached to an aromatic ring. In comparison, the 8–8C isomer yielded a dominant fragment at m/z 297 (loss of H_2O, CO, and CO_2), which was less abundant from isomers 8–8NC and 8–5C. The two 8–5-linked dimers also yielded distinguishing fragments, with m/z 293 (loss of H_2O, CO_2, and CH_3OH) and 307 ($2H_2O$ and CO_2) at 40-50% of the base peak in 8–5C, whereas m/z 309 (loss of H_2O, CO, and CH_3OH) was prominent in 8–5NC. Assignments of structural features to fragment ions can be challenging, but supporting evidence can be derived from MS^3 spectra and from accurate measurements of fragment masses using multiplexed and nonselective CID, which subjects all ions to collisional activation without prior mass filtering [42,44]. All of the fragment ions described above were detected as m/z values measured by UHPLC/TOF MS consistent, within 10 ppm, with the proposed elemental formulas. Further evidence about structures of these fragments came from observations of granddaughter (secondary product) ion masses. A summary of fragment ion annotations is presented in Figure 3.4, along with proposed pathways of fragment ion formation.

3.4.5 Comparison of CID spectra of diferulates with mono and di-amide functionalities

When diferulate esters are cleaved by a combination of ammonolysis and hydrolysis, Di-Am or two isomers of each asymmetrical Ac-Am product (except from the 8–8NC-diferulate that has a center of symmetry) may form (Figure 3.2). The CID spectra of synthetic Ac-Am and Di-Am standards bear great similarity to the corresponding Di-Ac spectra, differing primarily in loss of NH_3 instead of H_2O, and in relative abundances of analogous fragments. In addition, CID spectra of 8–O–4 Ac-Am (the later-eluting isomer) and Di-Am (Figure 3.3h-j) exhibited even-mass fragments at *m/z* 192 corresponding to cleavage of the central ether with charge retention on the amide portion. The two 8–O–4 Ac-Am isomers were readily distinguished from one another by their CID spectra (SI-Figure 3-S1a and 3-S1b) in that the loss of water (*m/z* 368) is more prominent in isomer 1 as is the *m/z* 308 fragment. This fragment ion is consistent with losses of two water molecules plus CO and methyl radical from $[M+H]^+$. Furthermore, isomer 2 yields higher relative abundance of fragments at *m/z* 326 (loss of one water plus CO and CH$_3$•) but minimal amounts of *m/z* 308. In other cases, the two Ac-Am isomers were not distinguished by their CID spectra, as illustrated by 8–5C isomers (SI-Figure 3-S1d and 3-S1e), and other means such as chromatographic retention time are necessary to distinguish the isomers. In the absence of standards of the individual isomers, their structural assignments (Figure 3.2) should be regarded as tentative.

3.4.6 Identification of diferulates released from corn stover cell walls upon NaOH catalyzed hydrolysis

Starting with information from CID mass spectra of authentic diferulates described above, diferulates derived from pretreatment of corn stover with NaOH were profiled using UHPLC/MS; released Di-Ac diferulates were identified by generating an extracted-ion

chromatogram for [M+H]$^+$ (*m/z* 387, Figure 3.5-a). At least twelve chromatographic peaks were identified as diferulates based on accurate mass measurements and MS^2 spectra, which are provided in SI-Figure 3-S3 for comparison with the MS^2 spectra of synthetic standards. Di-Ac 8–O–4-, 8–5C-, 8–5NC-, 8–8NC-, 8–8C- and 5–5-isomers were confirmed based on coelution with standards and equivalent MS^2 spectra. Peaks 1, 7 and 11 gave molecular masses consistent with diferulates and shared some side-chain loss fragments with other diferulates, but lacked unique product ions necessary to distinguish them from other Di-Ac isomers. Peak number 1 remains unauthenticated but it exhibits a fragment at *m/z* 193 suggestive of an ether linkage. Based on this observation and the long retention time similar to 8–O–4-isomers, this product is annotated as the 4–O–5-isomer, the only other anticipated isomer capable of forming *m/z* 193 via cleavage of the ether linkage. It has been reported before that the 4–O–5-isomer is not common in grasses and cereal fibers [33,45], and its low relative abundance is consistent with these earlier findings. The possibility of a configurational isomer (e.g., with one or both of the double bonds as *cis* geometrical isomers) of 8–O–4-diferulate is discounted because this compound was not formed upon photoirradiation of 8–O–4-diferulate, whereas Peak 2 increased during these experiments. Additional peaks are annotated as diferulates but did not match retention times with any standards. It is expected that some of the observed Di-Ac products are *cis*-isomers. For example, peak numbers 5 and 8 are both assigned as 8–5NC because of indistinguishable MS^2 spectra. The UHPLC retention time of compound 8 matched the synthetic standard of 8–5NC-diferulate. Information regarding all diferulate products including accurate masses is represented in SI-Figure 3-S2. Elution order of diferulates in a reversed-phase separation, and occurrence of identified diferulate isomers in corn stover presented here, are consistent with those in previous publication [33].

3.4.7 Identification of diferulates released from corn stover cell walls upon AFEX pretreatment

Hydrolysis and ammonolysis of diferulate esters are the two primary reactions that take place during AFEX pretreatment [12]. As suggested in Figure 3.2, it is expected that AFEX treatment cleaves diferulates via a combination of ammonolysis and hydrolysis, removing the hemicellulose cross-links, and releasing Di-Ac, Ac-Am, and Di-Am diferulates. After biomass pretreatment, the amounts of the individual diferulates released from the cell wall depend upon AFEX process conditions, abundances of various diferulate esters in the cell wall, and relative rates of hydrolysis and ammonolysis. Yields of these products reflect yields of cross-link removal, and quantification of these compounds provides measures of process efficiency in generating more digestible cell wall glycopolymers. Profiles of the three diferulate forms released upon AFEX treatment of corn stover are evident from extracted-ion chromatograms for m/z 385, 386, and 387 ($[M+H]^+$ for Di-Am, Ac-Am, and Di-Ac combined; Figure 3.5-b). At least 12 of the 15 diferulate peaks have been identified in extracts of AFEX treated corn stover. Retention times of the identified peaks were based in part on retention time and MS^2 matches (SI-Figure 3-S5) with synthesized Di-Ac, Ac-Am and Di-Am standard isomers. Information regarding all diferulate products including accurate masses is represented in SI-Figure 3-S4. In both NaOH and AFEX pretreatment, products derived from 8–O–4-diferulate were the most abundant. Ac-Am and Di-Amides of 8–O–4-, 8–8C-, 8–5NC- and 5–5-diferulates were present in the AFEX-treated corn stover extract, and are consistent with observation of the corresponding Di-Ac forms in the products of NaOH pretreatment. Released 8–8NC-diferulate products were observed as Di-Ac products from NaOH pretreatment, but were only observed in AFEX products in the Ac-Am form. Although the 8–5C Di-Ac isomer was the second most abundant product from NaOH hydrolysis, this compound or its other cyclic analogs were not observed in any form

following AFEX pretreatment. AFEX conditions likely converted it to 8–5NC variants that may undergo decarboxylation, but such products were not detected in non-targeted UHPLC/TOF MS analyses of AFEX products. Based on integrated peak areas, over 90% of AFEX-released diferulates contained at least one amide group (i.e., nearly all released identified diferulates are either Ac-Am or Di-Am, with 8–O–4-Di-Ac accounting for about 9% of the total diferulate products). Under these AFEX treatment conditions, rates of ammonolysis were therefore apparently greater than those of hydrolysis. Several of the observed Ac-Am and Di-Am diferulates are identical to products detected in reaction products of diferulate esters with liquid ammonia and water under AFEX-like conditions [40].

Figure 3.5. UHPLC/MS/MS extracted-ion chromatograms of diferulates from (a) NaOH-pretreated corn stover (*m/z* 387 for Di-Ac) and (b) AFEX-pretreated corn stover (*m/z* 385, 386 and 387 combined for Di-Am, Ac-Am and Di-Ac). Compound annotations are provided in Supplemental Information Figure 3-S2 and 3-S3.

3.4.8 Similar CID fragmentation using different mass spectrometers

To explore whether the findings of CID behavior can be extended to other kinds of mass spectrometers, CID spectra were generated for all diferulate forms on a hybrid linear ion trap (QTRAP 3200) using enhanced product-ion scanning, a triple quadrupole (Quattro Premier XE) using product-ion scans, and a quadrupole/time-of-flight hybrid (QToF Ultima API). In addition, nonselective CID spectra were generated on an orthogonal TOF instrument (LCT Premier). A comparison of CID spectra for 8–O–4-Di-Ac is presented in SI-Figure 3-S7. Despite differences in the time frame between ion-molecule collision and ion detection, and collision gases (N_2 in QTrap, Ar in others), all CID spectra share common major fragment ions, suggesting that the identifications and analysis here will be universally useful.

To provide for quantitative analysis of individual Di-Ac forms of diferulates, a multiple reaction monitoring (MRM) method was developed using transitions from $[M+H]^+$ to a characteristic product ion for five isomeric forms. Relative response factors (RRFs) were determined for each isomer (Figure 3-S8). Although it is recognized that response factors will vary across different instrument platforms and with various instrumental conditions, the differences between isomer responses were consistent with a combination of two factors: (1) increasing ionization efficiency as the methanol content of the mobile phase increased, and (2) differences in yields of the detected product ions, as evident from the product ion spectra (Figure 3.3).

3.5 Conclusions

Comprehensive determination of the chemical diversity of diferulate cross-linkers in plant cell walls provides a rich source of information useful for guiding the breeding of grasses for improved digestibility and resistance to pests. Future efforts to engineer plant cell walls for desirable traits will depend on analytical tools for rapidly profiling diferulates released from

plant tissues. The discoveries presented above demonstrate that diferulate structural isomers are distinguished by their MS^2 product-ion spectra when combined with UHPLC retention times, though discrimination of *cis-* and *trans-* isomers may require their chromatographic resolution. In many cases, isomer-specific fragment ions offer the prospect for rapid UHPLC/MS/MS analyses using MRM for quantitative analyses. This approach avoids the need for derivatization and related sample processing steps, allowing direct analysis of crude biomass extracts, and promises to address the need for more robust quantitative methods for ferulate oligomers [46]. The primary barrier to adopting such strategies lies in the limited availability of authentic standards, though most of the diferulate esters themselves are now readily available [32]. Cell walls also accumulate larger ferulate oligomers and lignin cross-coupling products [47-50] that will present a yet more complex array of isomeric products. Our findings also demonstrate the diversity of diferulates from NaOH and AFEX pretreatment of cell walls, showing most of the anticipated diferulate isomers but revealing additional complexity. It is our hope that the CID spectra can be used in MS/MS spectrum libraries to aid diferulate characterization, and that the ion fragmentation chemistry demonstrated in this work will guide future efforts to characterize higher ferulate oligomers [31,36,46], and establish the levels and roles of diferulates in cell wall structures of wild and engineered plant materials.

APPENDIX

Table 3-S1. Fragment ions from CID MS/MS spectra of $[M+H]^+$ from diferulic acid standards using 20 eV collision energy, with suggested consecutive losses of side-chains and relative abundance for each product ion. MS^3 performed on major product ion supports these suggested pathways.

Type of ion	Neutral Loss (Da)	m/z	8-O-4	8-8 NC	8-8 C	8-5 C	8-5 NC	5-5
$[M+H^+-H_2O]^+$	18	369	10	100			100	
$[M+H^+-2H_2O]^+$	36	351	100	48	24	35		100
$[M+H^+-CO_2]^+$	44	343				40		
$[M+H^+-H_2O-CO]^+$	46	341		43	95		90	
$[M+H^+-CO-CH_3^{\bullet}]^+$	61	326	37					
$[M+H^+-H_2O-CO_2]^+$	62	325	32			100	18	
$[M+H^+-2H_2O-CO]^+$	64	323		40	55	20		10
$[M+H^+-2H_2O-CH_3OH]^+$	68	319						18
$[M+H^+-H_2O-CO-CH_3OH]^+$	78	309					55	
$[M+H^+-2H_2O-CO_2]^+$	80	307				50		
$[M+H^+-H_2O-CO-CO_2]^+$	90	297		49	100			
$[M+H^+-H_2O-CO_2-CH_3OH]^+$	94	293				35		
$[M+H^+-2H_2O-2CH_3OH]^+$	100	287						4
$[M+H^+-H_2O-CO-CO_2-CH_3OH]^+$	122	265		30	54	26		
$[M+H^+-H_2O-C_7H_8O_2]^+$	142	245		71				
$[M+H^+-H_2O-2CO-CO_2-$	150	237			12			
$[263-CO_2]^+$	168	219	33					
$[219-H_2O]^+$	186	201	15					
$[M+H^+- C_{10}H_{10}O_4]^+$	194	193	70					
$[219-CH_3OH]^+$	210	177	15					
$[M+H^+-C_{10}H_{10}O_4-CO_2]^+$	238	149	15					
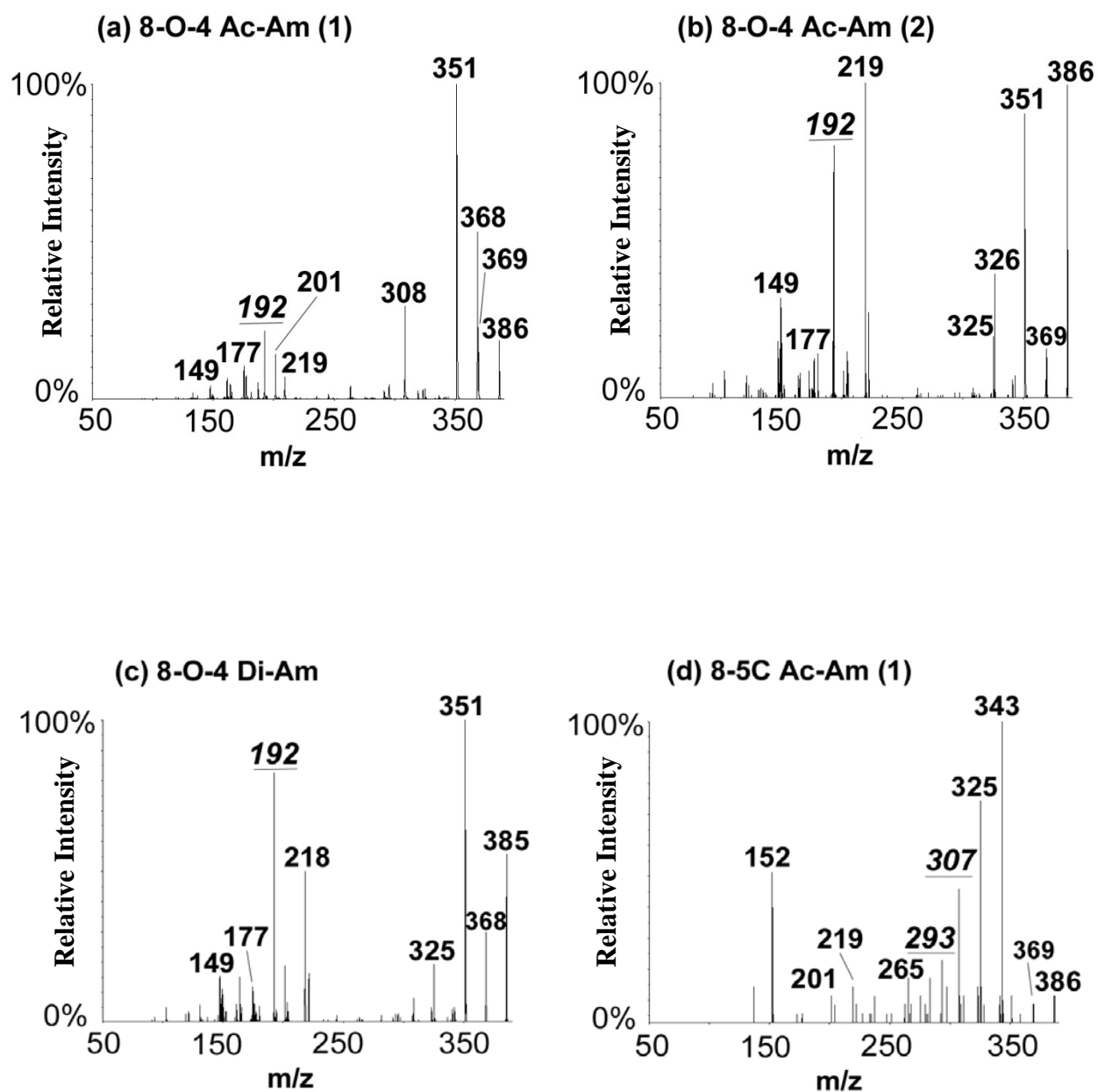
Figure 3-S1. Enhanced product ion (MS/MS) spectra of [M+H]$^{+}$ from synthetic diferulate Ac-Am and Di-Am

Figure 3-S1. (cont'd)

(e) 8-5C Ac-Am (2)

(f) 8-8NC Ac-Am

(g) 5-5 Ac-Am

(h) 5-5 Di-Am

(i) 8-8C Ac-Am

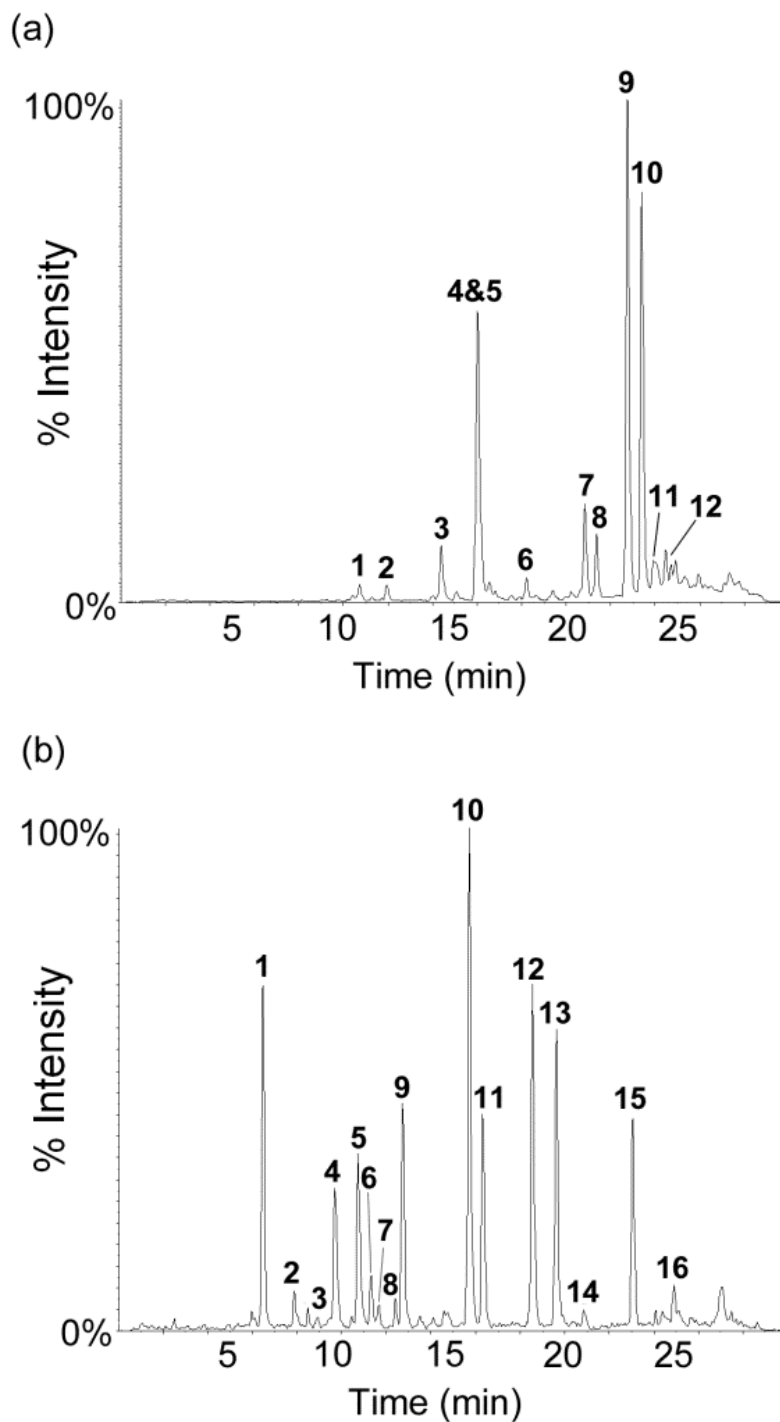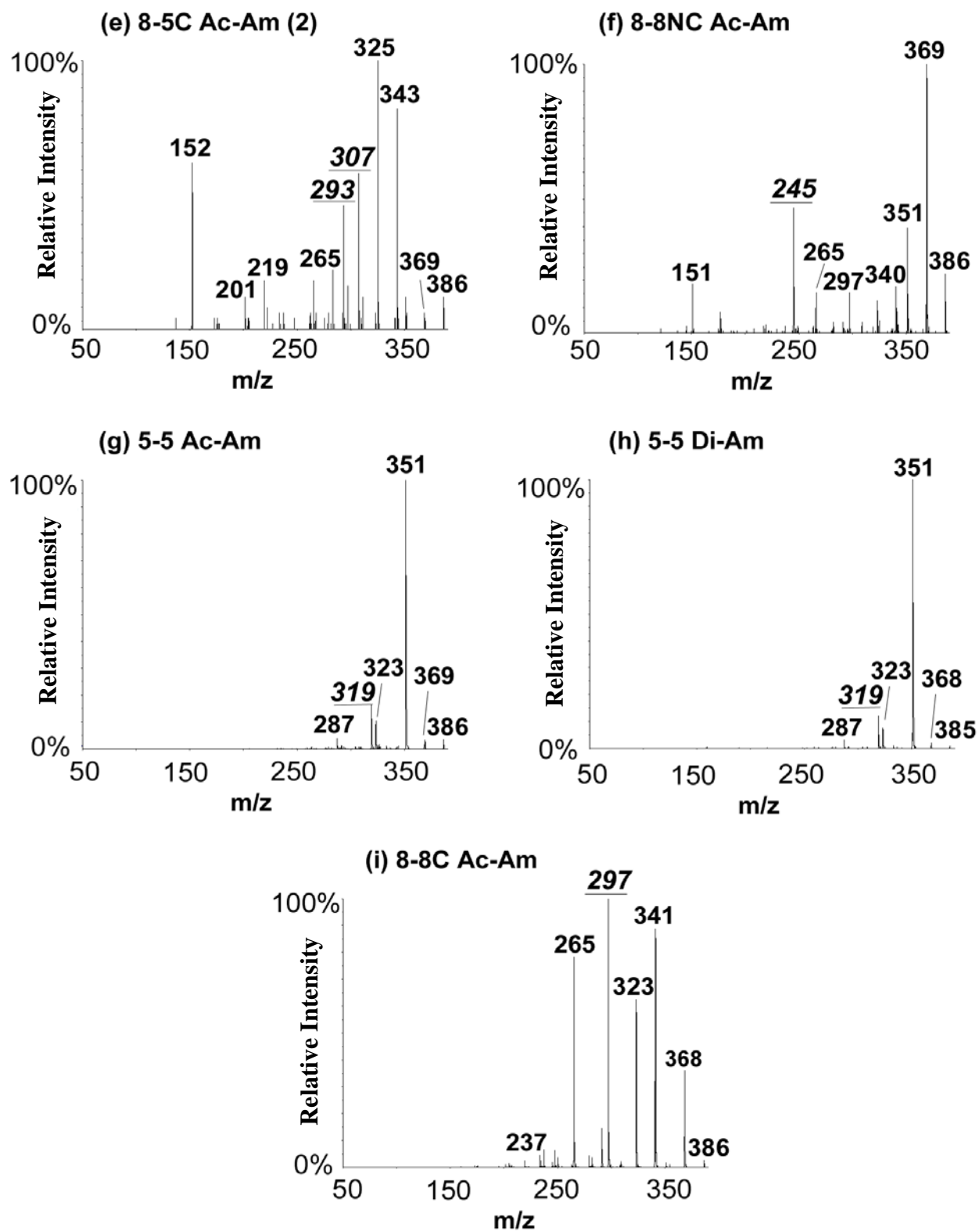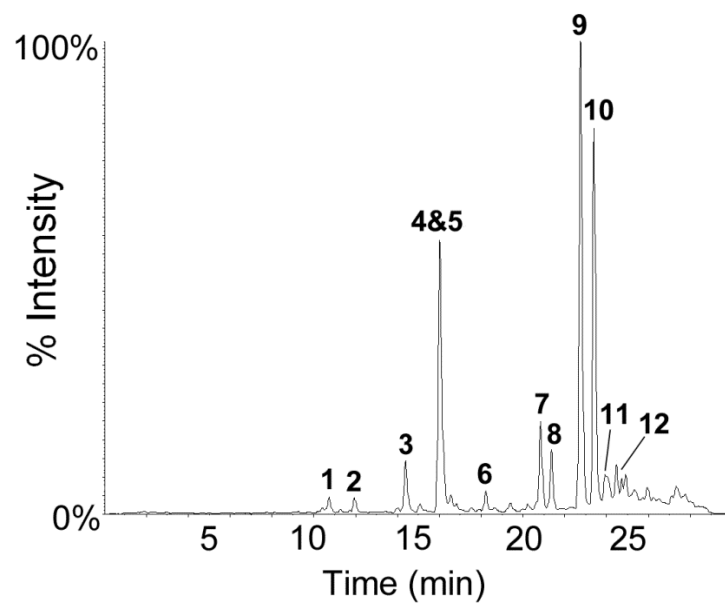**Figure 3-S2.** UHPLC/MS/MS extracted-ion chromatograms of diferulates from NaOH pretreated corn stover (*m/z* 387 for Di-Ac)

Table 3-S2. Retention time and high resolution mass measurements of diferulic acids identified in products from NaOH-hydrolyzed corn stover.

Peak #	Retention time	*m/z*	Theoretical Mass	Δm (ppm)[a]	Compound
1	10.8	387.1062	387.1080	-4.6	8-5C Di-Ac (*cis*)[b]
2	12.0	387.1069	387.1080	-2.8	Unknown Diferulate[c]
3	14.5	387.1074	387.1080	-1.5	8-8C Di-Ac
4	16.2	387.1063	387.1080	-4.4	8-8NC Di-Ac
5	16.1	387.1069	387.1080	-2.8	8-5NC Di-Ac
6	18.4	387.1067	387.1080	-3.4	Unknown Diferulate[c]
7	20.9	387.1068	387.1080	-3.0	5-5 Di-Ac
8	21.6	387.1072	387.1080	-2.1	8-5NC Di-Ac
9	22.9	387.1078	387.1080	-0.5	8-O-4 Di-Ac (*trans*)
10	23.4	387.1075	387.1080	-1.3	8–5C Di-Ac[b]
11	23.8	387.1062	387.1080	-4.7	8–O–4 Di-Ac (*cis*)[d]
12	24.8	387.1065	387.1080	-3.9	4–O–5 Di-Ac

a) Δm (ppm)=(measured mass-theoretical mass/measured mass) $\times 10^6$

b) The huge retention time difference between the two 8–5C isomers is attributed to differences in intramolecular hydrogen bonds

c) These two diferulates have fragmentation patterns similar to 8–8C

d) Annotated as a *cis*-isomer (one or both double bonds) as abundance increases upon photoirradiation.
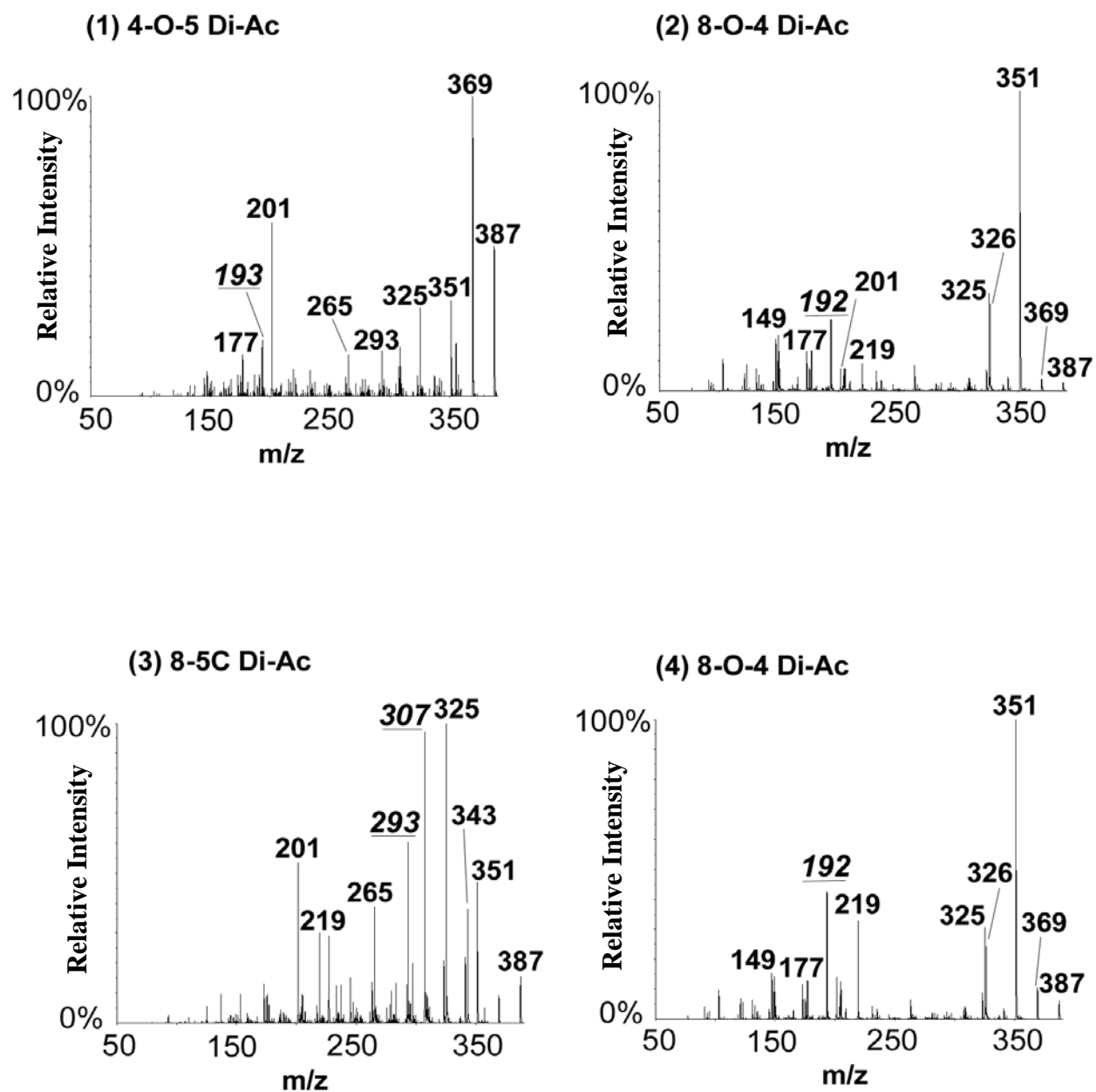
Figure 3-S3. CID MS/MS spectra of diferulic acids from NaOH-hydrolyzed corn stover. Labels 1-12 correspond to the peaks in Table 3-S2

Figure 3-S3. (cont'd)

(5) 8-5NC Di-Ac

(6) 5-5 Di-Ac

(7) Unknown Di-Ac

(8) 8-5NC Di-Ac

(9) 8-8NC Di-Ac

(10) 8-8C Di-Ac

Figure 3-S3. (cont'd)

(11) Unknown Di-Ac

(12) 8-5C Di-Ac

(b)
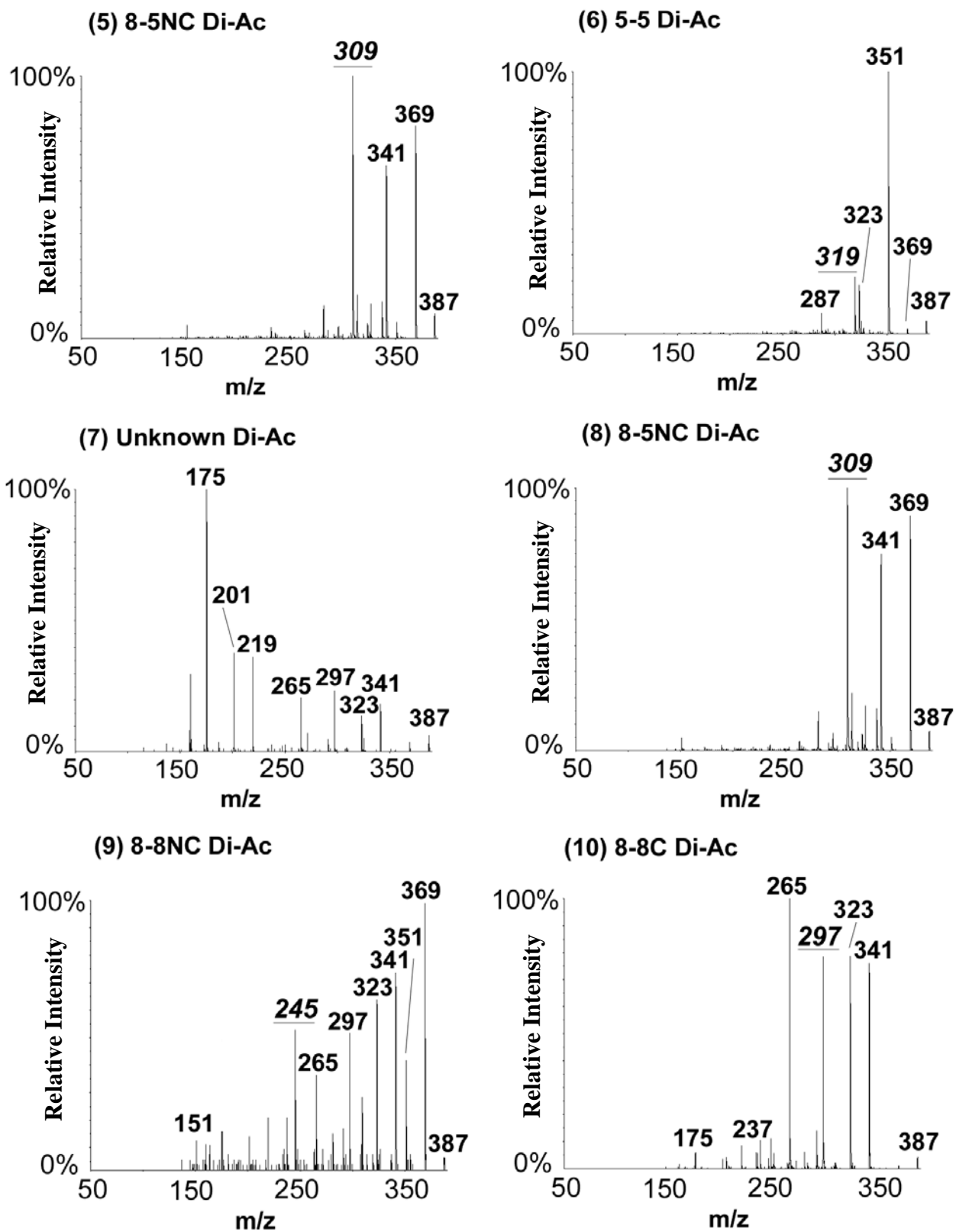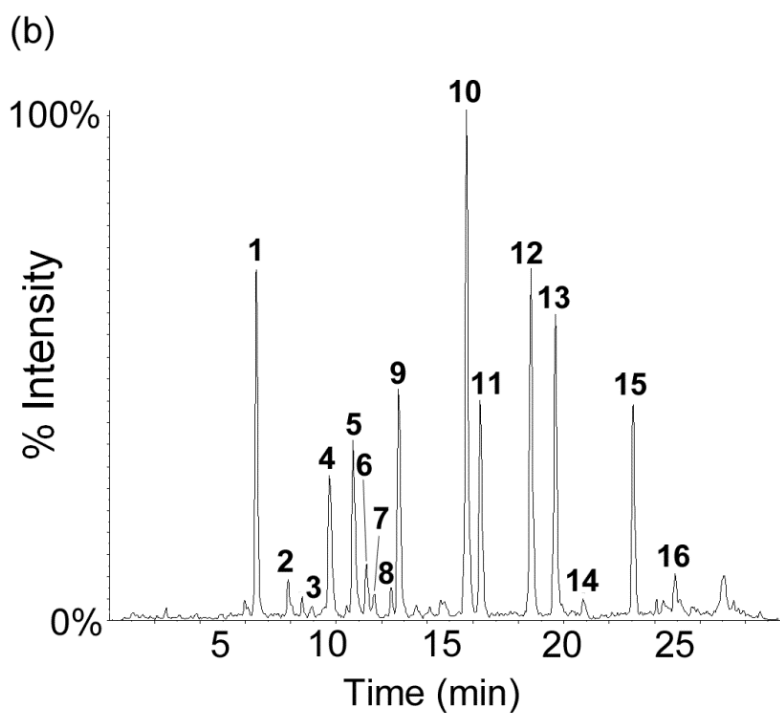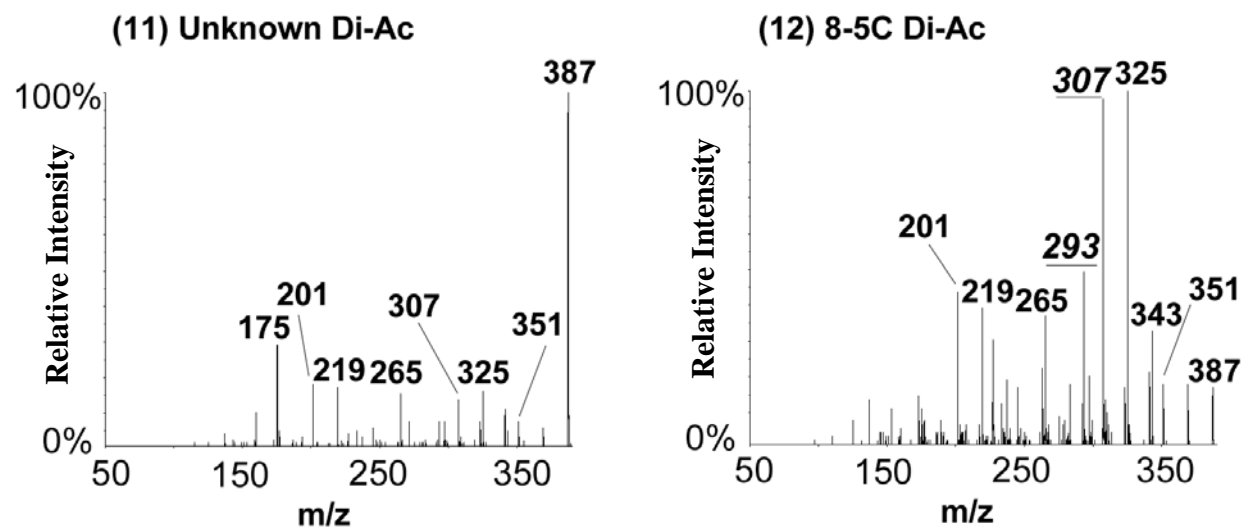
Figure 3-S4. UHPLC/MS/MS extracted-ion chromatograms of diferulates from AFEX-pretreated corn stover

Table 3-S3. Retention time and high resolution mass measurements of Di-Ac, Ac-Am and Di-Am diferulates identified in products from AFEX-treated corn stover.

Peak #	Retention time	*m/z*	Theoretical Mass	Δm (ppm)	Compound
1	6.4	387.1547	387.1556[a]	-2.3	Feruloyl amide
2	7.8	385.1390	385.1400	-2.6	8-8C Di-Am
3	8.5	385.1387	385.1400	-3.4	8-5NC Di-Am
4	9.7	386.1230	386.1240	-2.6	8-8C Ac-Am (2)
5	10.7	386.1235	386.1240	-1.3	8-8C Ac-Am (1)
6	11.3	386.1226	386.1240	-3.6	Unknown Diferulate Ac-Am
7	11.6	386.1242	386.1240	-0.5	8-5NC Ac-Am
8	12.4	385.1387	385.1400	-3.3	5-5 Di-Am
9	12.7	386.1231	386.1240	-2.3	8-8NC Ac-Am
10	15.7	385.1392	385.1400	-2.1	8-O-4 Di-Am
11	16.3	386.1233	386.1240	-1.8	5-5 Ac-Am
12	18.5	386.1237	386.1240	-0.8	8-O-4 Ac-Am (2)
13	19.6	386.1235	386.1240	-1.3	8-O-4 Ac-Am (1)
14	20.8	386.1225	386.1240	-3.9	Unknown Diferulate Ac-Am
15	23.0	387.1069	387.1080	-2.8	8-O-4 Di-Ac
16	24.9	386.1222	386.1240	-4.7	Unknown Diferulate Ac-Am

a) This is the mass of a non-covalent dimer ($[2M+H]^{+}$) of feruloyl amide. The 387 > 193 MRM transition, which was used for 8–O–4-diferulic acid generates signal for this ion.
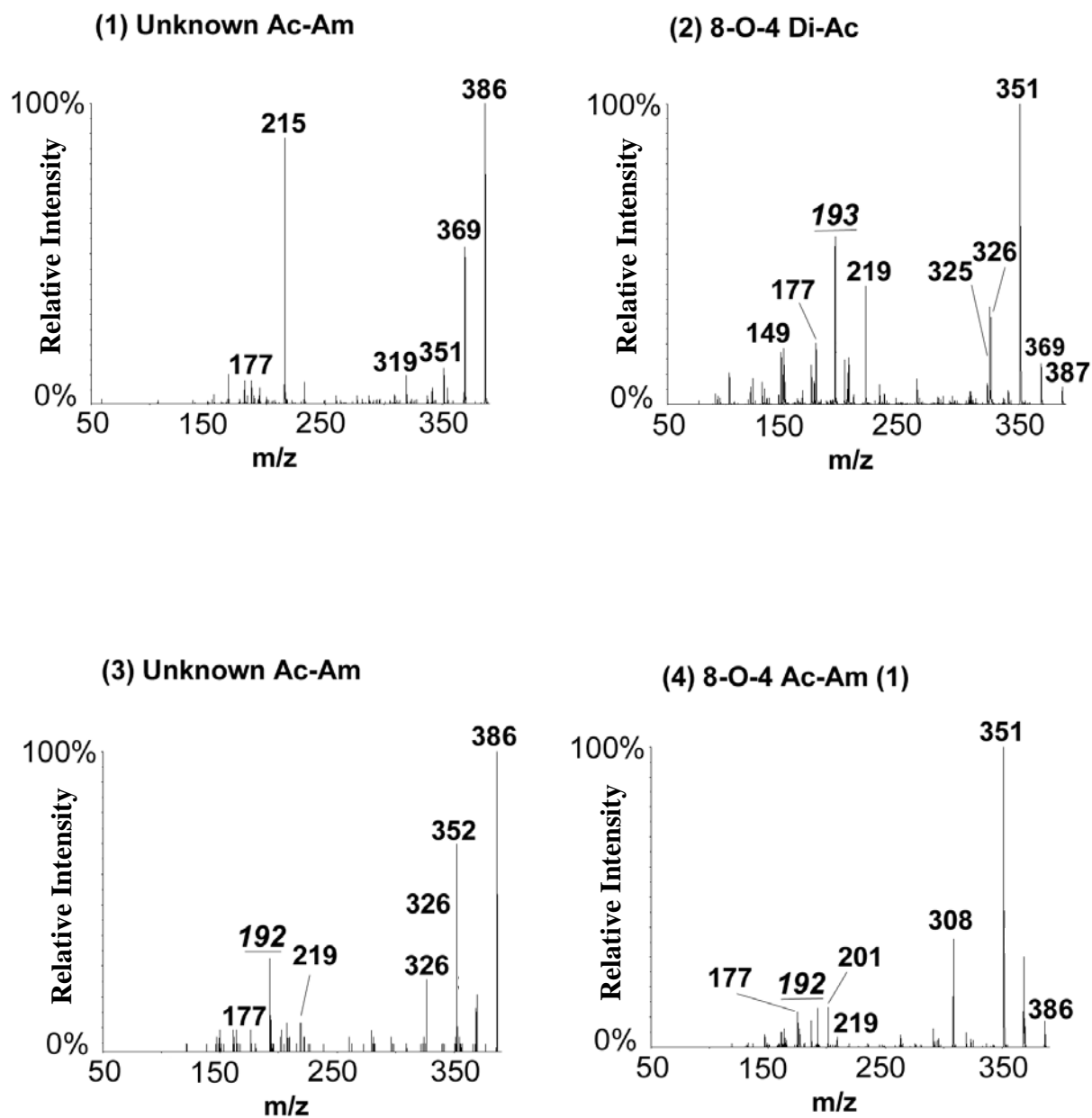
Figure 3-S5. CID MS/MS spectra of Di-Ac, Ac-Am and Di-Am diferulates identified in products from AFEX-treated corn stover. Labels 1-16 correspond to the peaks in Table 3-S3.

Figure 3-S5. (cont'd)

(5) 8-O-4 Ac-Am (2)

(6) 5-5 Ac-Am

(7) 8-O-4 DoAm

(8) 8-8NC Ac-Am

(9) 5-5 Di-Am

(10) 8-5NC Ac-Am

Figure 3-S5. (cont'd)

(11) Unknown

(12) 8-8C Ac-Am (1)

(13) 8-8C Ac-Am (2)

(14) 8-5NC Di-Am

(15) 8-8C Di-Am

(16) Feruloyl amide non-covalent dimer
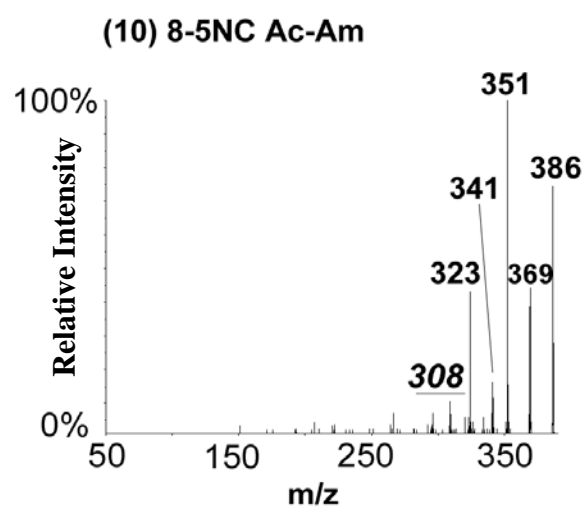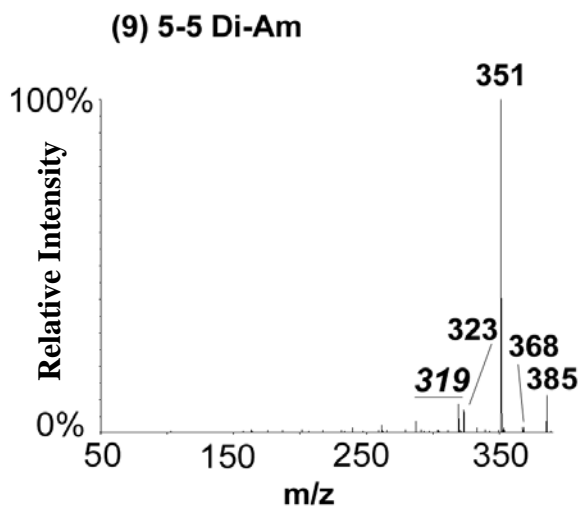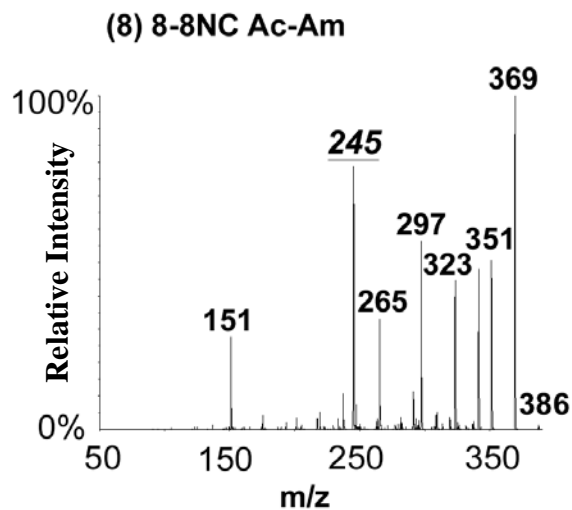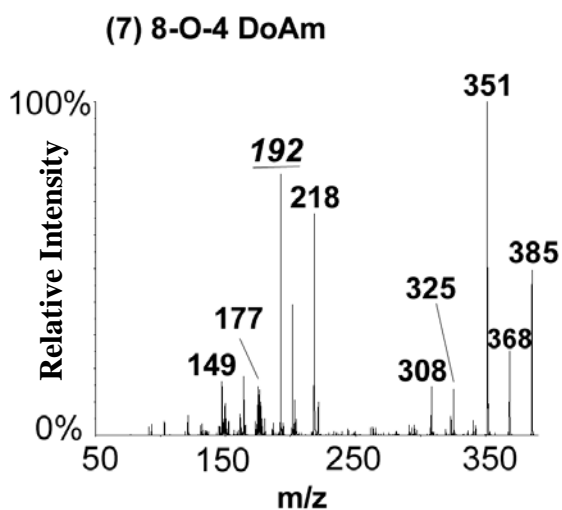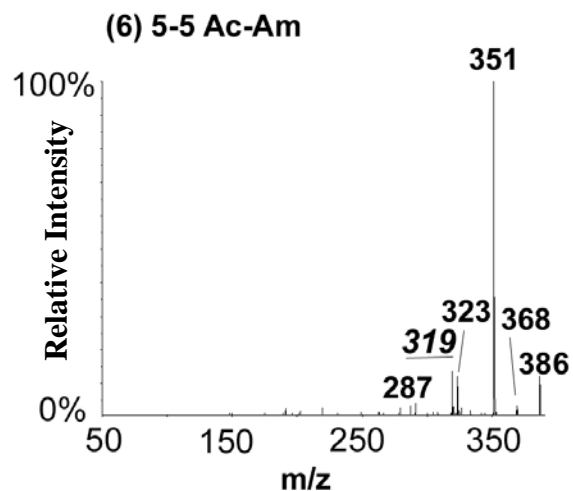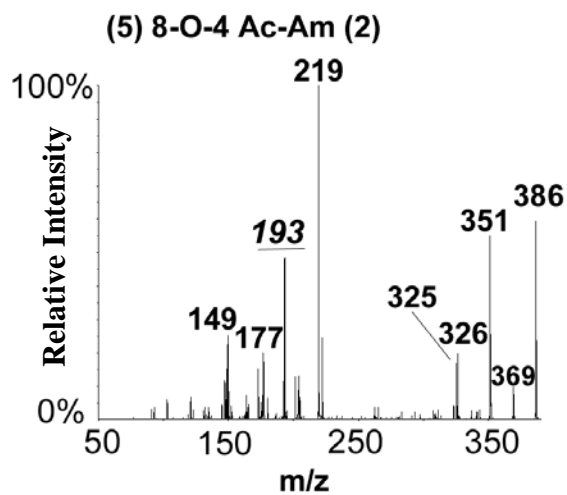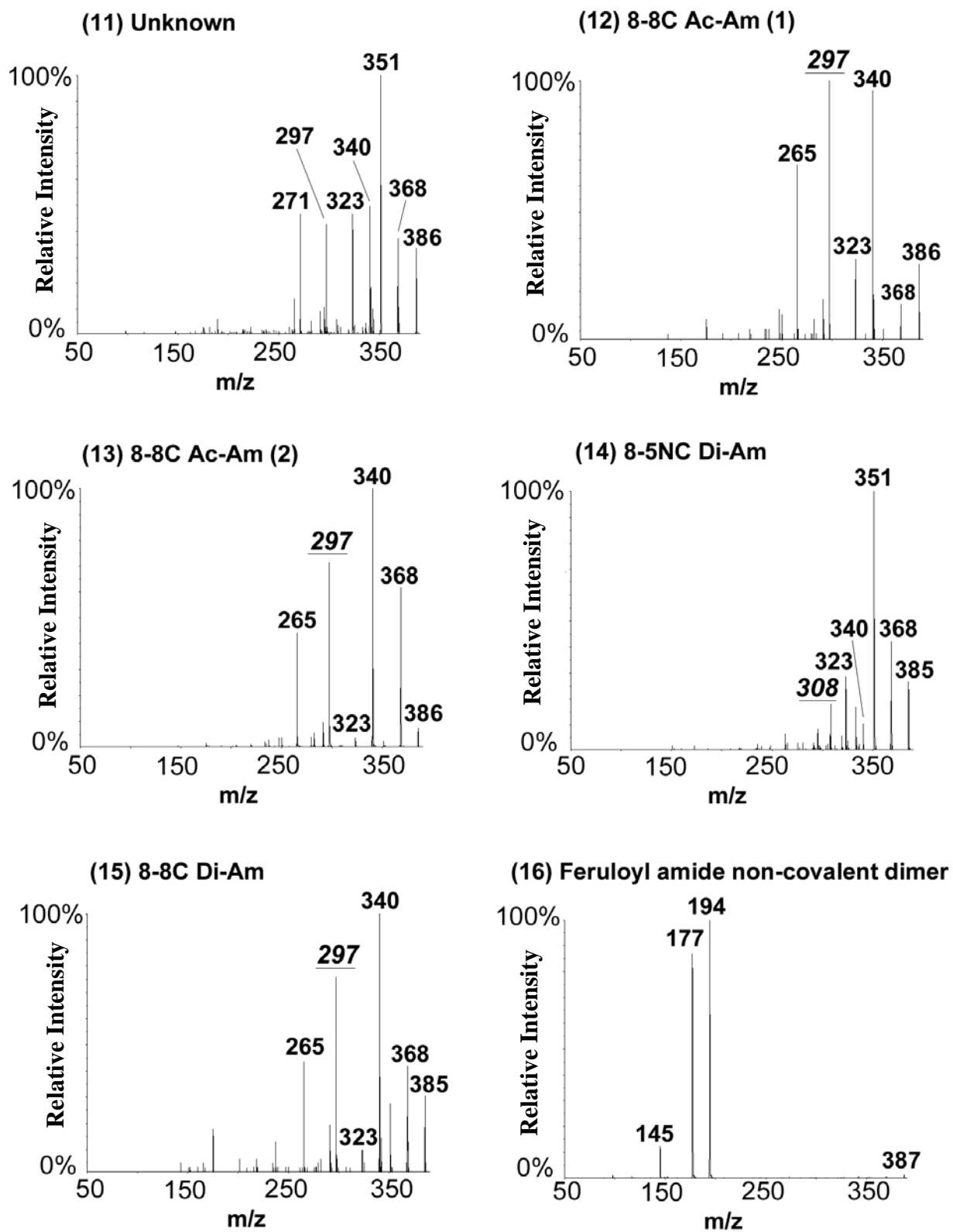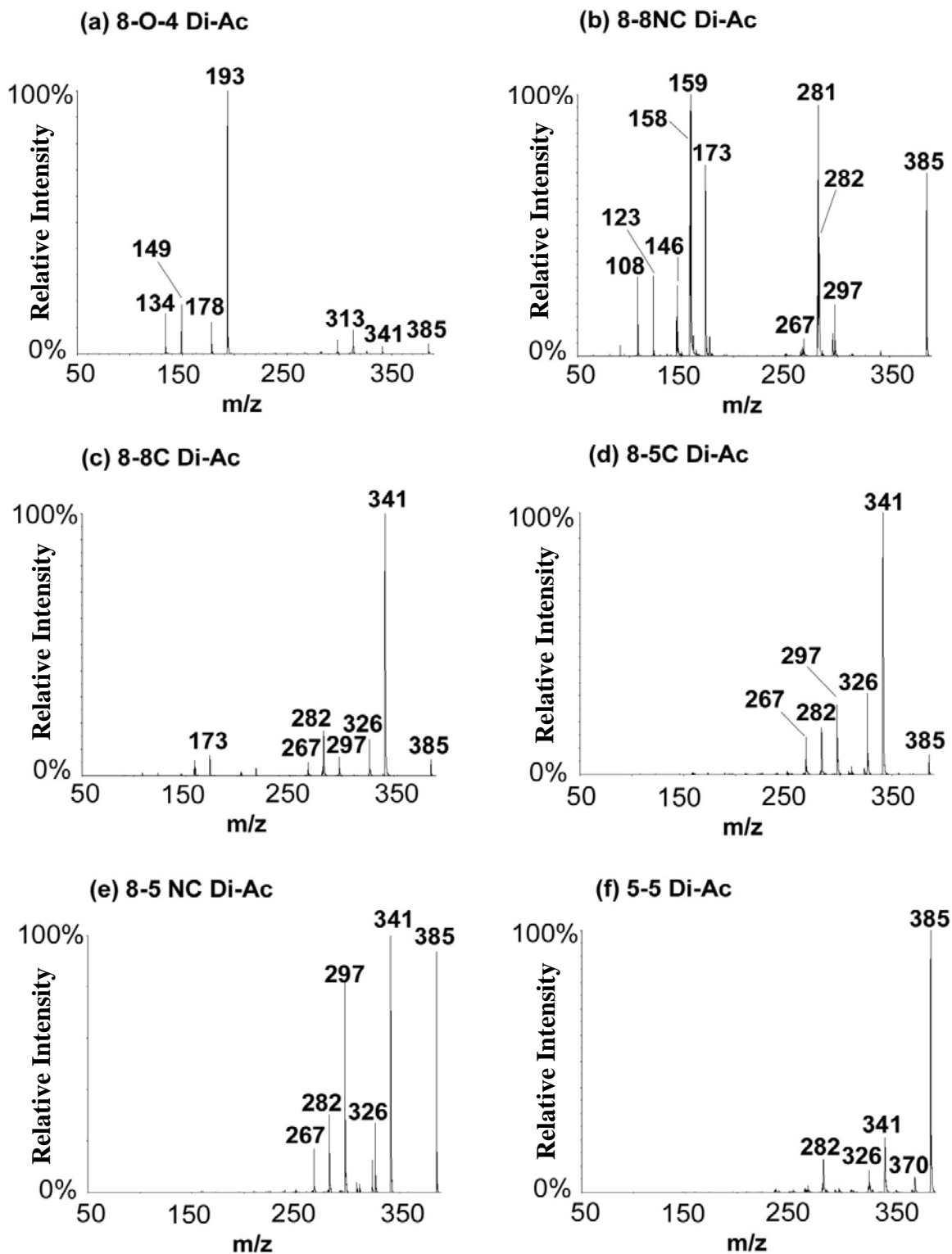
Figure 3-S6. CID MS/MS spectra from [M-H]$^{-}$ of diferulic acid authentic standards using 20 eV collision energy
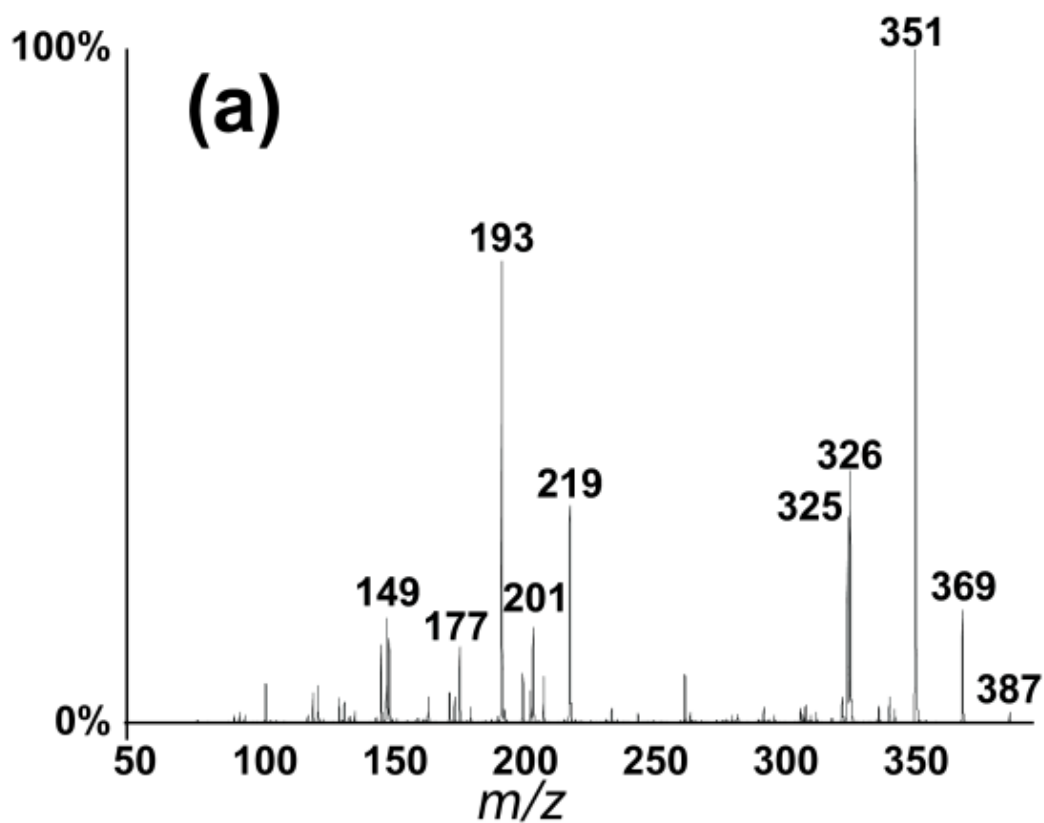
Figure 3-S7. CID spectra of diferulic acid authentic standards generated using (a) a hybrid linear ion trap (QTRAP 3200), (b) an orthogonal TOF instrument (LCT Premier) with nonselective CID, (c) a triple quadrupole (Quattro Premier XE) and (d) quadrupole/time-of-flight hybrid (QTOF Ultima API). MS/MS spectra are products of $[M+H]^{+}$.

Figure 3-S7. (cont'd)
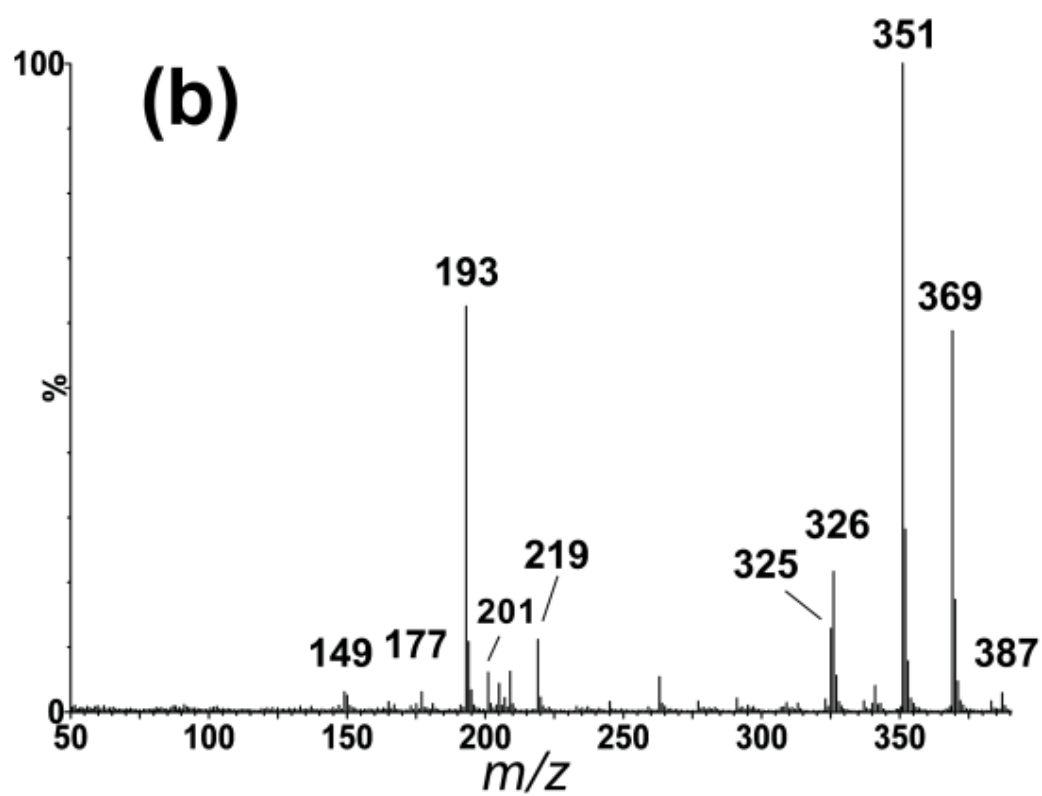
Figure 3-S7. (cont'd)
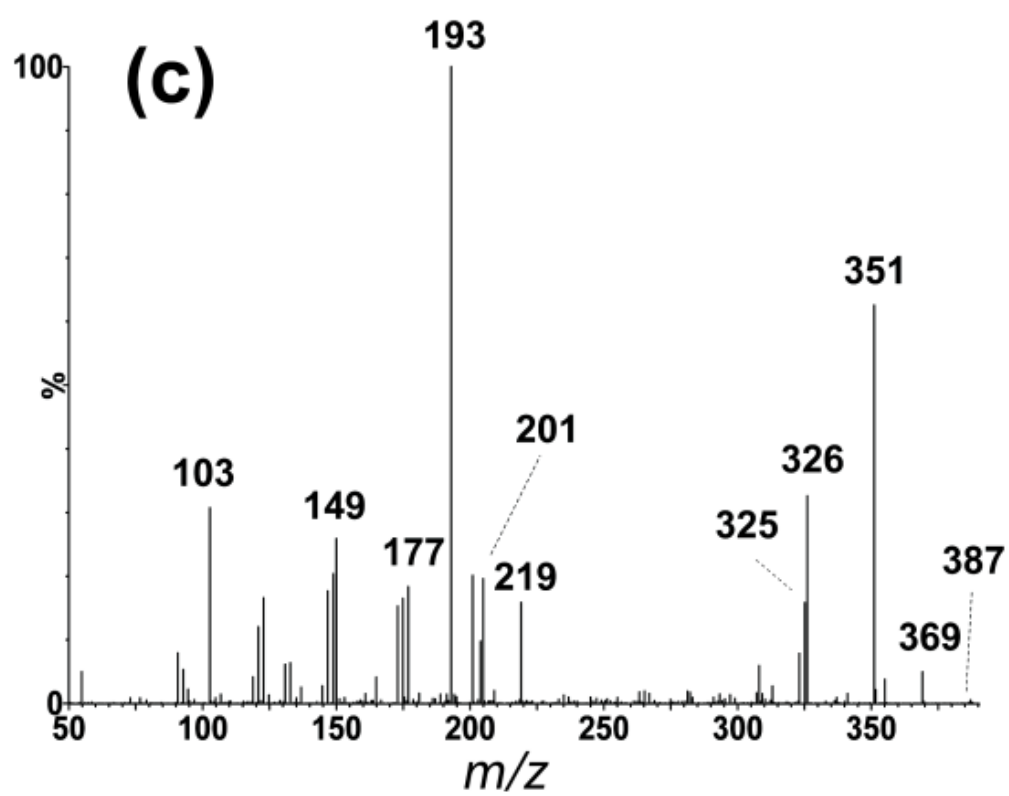
Figure 3-S7. (cont'd)
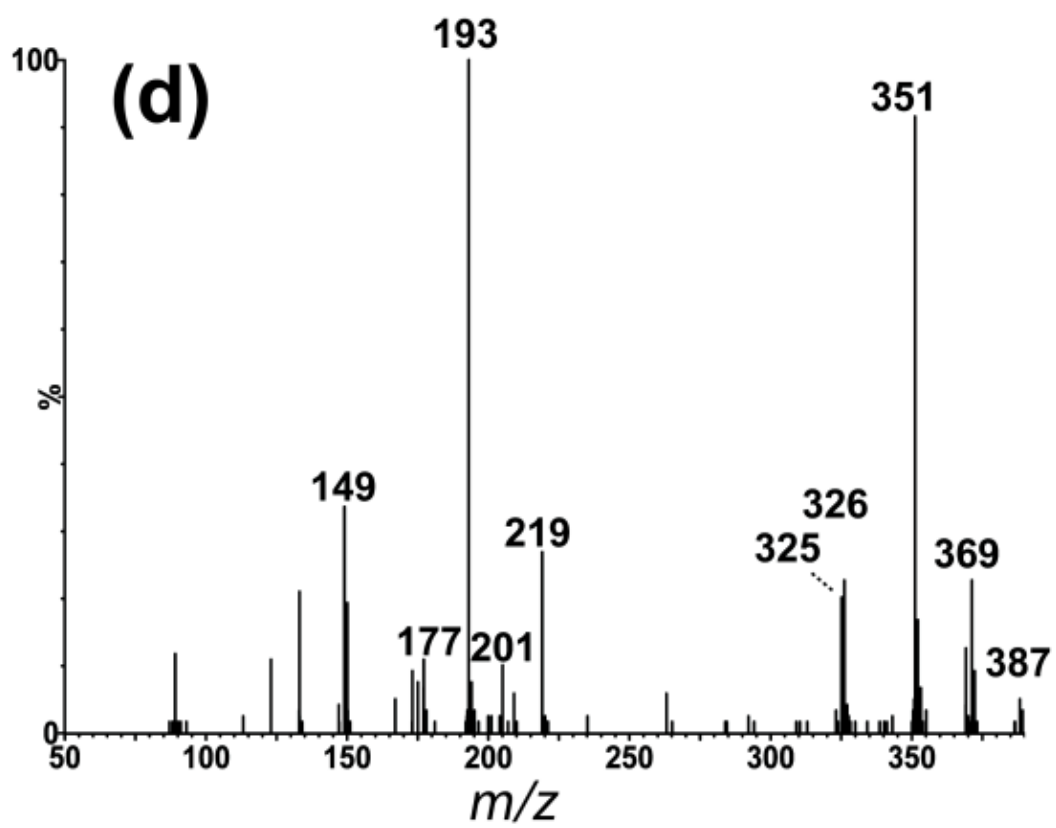
Fig. 3-S8. Extracted ion chromatograms for 5 MRM transitions from five authentic diferulic acid standards at equal concentrations (5 µM; 5 µL injection).
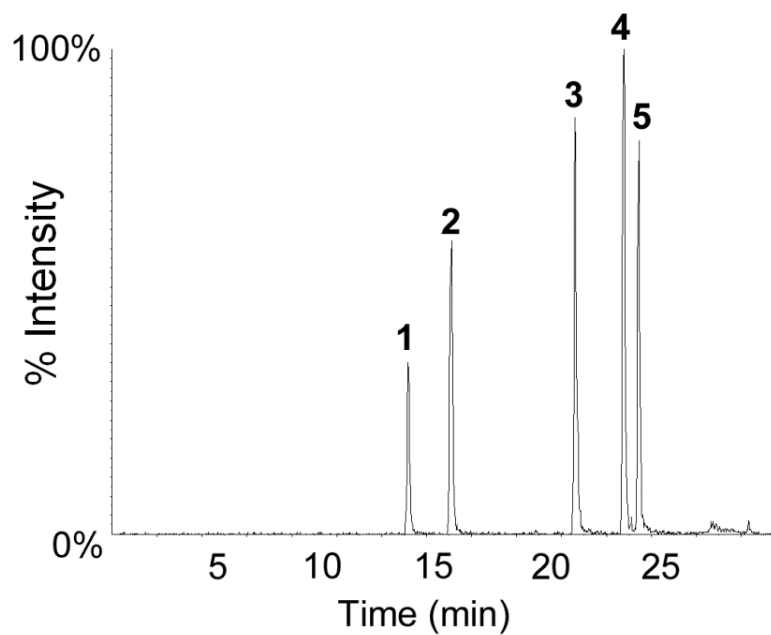
Peak #	Diferulic acid isomer	MRM transition in ESI positive mode	Relative response factor (relative to peak #1)
1	8–8C	387 > 297	1.00
2	8–8NC	387 > 245	1.76
3	5–5	387 > 319	2.37
4	8–O–4	387 > 193	2.47
5	8–5C	387 > 293	1.70

REFERENCES

REFERENCES

(1)   Last, R. L.; Jones, A. D.; Shachar-Hill, Y. *Nature Reviews Molecular Cell Biology* **2007**, *8*, 167-174.

(2)   Buanafina, M. M. D. *Molecular Plant* **2009**, *2*, 861-872.

(3)   Vogel, J. *Current Opinion in Plant Biology* **2008**, *11*, 301-307.

(4)   Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. In *Annual Review of Chemical and Biomolecular Engineering, Vol 2*; Prausnitz, J. M., Ed.; Annual Reviews: Palo Alto, 2011; Vol. 2, p 121-145.

(5)   Ralph, J.; Bunzel, M.; Marita, J. M.; Hatfield, R. D.; Lu, F.; KIm, H.; Schatz, P. F.; Grabber, J.; Steinhart, H. *Phytochemistry Reviews* **2004**, *3*, 79-96.

(6)   Ralph, J.; Quideau, S.; Grabber, J. H.; Hatfield, R. D. *Journal of the Chemical Society-Perkin Transactions 1* **1994**, 3485-3498.

(7)   Bunzel, M.; Heuermann, B.; Kim, H.; Ralph, J. *Journal of Agricultural and Food Chemistry* **2008**, *56*, 10368-10375.

(8)   Ralph, J. *Phytochemistry Reviews* **2010**, *9*, 65-83.

(9)   Jung, H. J. G.; Sahlu, T. *Journal of the Science of Food and Agriculture* **1986**, *37*, 659-665.

(10)   Grabber, J. H.; Hatfield, R. D.; Ralph, J. *Journal of the Science of Food and Agriculture* **1998**, *77*, 193-200.

(11)   Kabel, M. A.; Carvalheiro, F.; Garrote, G.; Avgerinos, E.; Koukios, E.; Parajo, J. C.; Girio, F. M.; Schols, H. A.; Voragen, A. G. J. *Carbohydrate Polymers* **2002**, *50*, 47-56.

(12)   Chundawat, S. P. S.; Vismeh, R.; Sharma, L. N.; Humpula, J. F.; Sousa, L. D.; Chambliss, C. K.; Jones, A. D.; Balan, V.; Dale, B. E. *Bioresource Technology* **2010**, *101*, 8429-8438.

(13)   Maslen, S. L.; Goubet, F.; Adam, A.; Dupree, P.; Stephens, E. *Carbohydrate Research* **2007**, *342*, 724-735.

(14)   Pauly, M.; Keegstra, K. *Plant Journal* **2008**, *54*, 559-568.

(15)   Izydorczyk, M. S.; Biliaderis, C. G. *Carbohydrate Polymers* **1995**, *28*, 33-48.

(16)   Ishii, T. *Plant Science* **1997**, *127*, 111-127.

(17)   Kato, Y.; Nevins, D. J. *Carbohydrate Research* **1985**, *137*, 139-150.

(18)   Grabber, J. H.; Hatfield, R. D.; Ralph, J.; Zon, J.; Amrhein, N. *Phytochemistry* **1995**, *40*, 1077-1082.

(19)   Muellerharvey, I.; Hartley, R. D.; Harris, P. J.; Curzon, E. H. *Carbohydrate Research* **1986**, *148*, 71-85.

(20)   Kato, A.; Azuma, J.; Koshijima, T. *Chemistry Letters* **1983**, 137-140.

(21)   Smith, M. M.; Hartley, R. D. *Carbohydrate Research* **1983**, *118*, 65-80.

(22)   Ishii, T.; Hiroi, T. *Carbohydrate Research* **1990**, *196*, 175-183.

(23)   Appeldoorn, M. M.; Kabel, M. A.; Van Eylen, D.; Gruppen, H.; Schols, H. A. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 11294-11301.

(24)   Hatfield, R.; Jones, B.; Grabber, J.; Ralph, J. *Plant Physiology* **1995**, *108*, 125-125.

(25)   Bunzel, M. *Phytochemistry Reviews* **2010**, *9*, 47-64.

(26)   Bunzel, M.; Allerdings, E.; Ralph, J.; Steinhart, H. *Journal of Cereal Science* **2008**, *47*, 29-40.

(27)   Allerdings, E.; Ralph, J.; Steinhart, H.; Bunzel, M. *Phytochemistry* **2006**, *67*, 1276-1286.

(28)   Allerdings, E.; Ralph, J.; Schatz, P. F.; Gniechwitz, D.; Steinhart, H.; Bunzel, M. *Phytochemistry* **2005**, *66*, 113-124.

(29)   Kim, H.; Ralph, J. *Organic & Biomolecular Chemistry* **2010**, *8*, 576-591.

(30)   Bunzel, M.; Allerdings, E.; Sinwell, V.; Ralph, J.; Steinhart, H. *European Food Research and Technology* **2002**, *214*, 482-488.

(31)   Barberousse, H.; Roiseux, O.; Robert, C.; Paquot, M.; Deroanne, C.; Blecker, C. *Journal of the Science of Food and Agriculture* **2008**, *88*, 1494-1511.

(32)   Lu, F.; We, L.; Azarpira, A.; Ralph, J. *J. Agric. Food. Chem.* **2012**, *60*, 8272-8277.

(33)   Dobberstein, D.; Bunzel, M. *Journal of Agricultural and Food Chemistry* **2010**, *58*, 8927-8935.

(34)   Bunzel, M.; Ralph, J.; Marita, J. M.; Hatfield, R. D.; Steinhart, H. *Journal of the Science of Food and Agriculture* **2001**, *81*, 653-660.

(35)   Morreel, K.; Kim, H.; Lu, F. C.; Dima, O.; Akiyama, T.; Vanholme, R.; Niculaes, C.;

Goeminne, G.; Inze, D.; Messens, E.; Ralph, J.; Boerjan, W. *Analytical Chemistry* **2010**, *82*, 8095-8105.

(36)   Morreel, K.; Dima, O.; Kim, H.; Lu, F. C.; Niculaes, C.; Vanholme, R.; Dauwe, R.; Goeminne, G.; Inze, D.; Messens, E.; Ralph, J.; Boerjan, W. *Plant Physiology* **2010**, *153*, 1464-1478.

(37)   Bily, A. C.; Burt, A. J.; Ramputh, A. L.; Livesey, J.; Regnault-Roger, C.; Philogene, B. R.; Arnason, J. T. *Phytochemical Analysis* **2004**, *15*, 9-15.

(38)   Chundawat, S. P. S.; Donohoe, B. S.; Sousa, L. D.; Elder, T.; Agarwal, U. P.; Lu, F. C.; Ralph, J.; Himmel, M. E.; Balan, V.; Dale, B. E. *Energy & Environmental Science* **2011**, *4*, 973-984.

(39)   Balan, V.; Bals, B.; Chundawat, S. P. S.; Marshall, D.; Dale, B. E. In *Biofuels: Methods and Protocols*; Mielenz, J. R., Ed.; Humana Press Inc: Totowa, 2009; Vol. 581, p 61-77.

(40)   Azarpira, A.; Lu, F.; Ralph, J. *Org Biomol Chem* **2011**, *9*, 6779-6787.

(41)   Grabber, J. H.; Ralph, J.; Hatfield, R. D. *Journal of Agricultural and Food Chemistry* **2000**, *48*, 6106-6113.

(42)   Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Jones, A. D.; Last, R. L. *Plant Journal* **2010**, *62*, 391-403.

(43)   Li, C.; Schmidt, A.; Shi, F.; Pichersky, E.; Jones, A. D. *Metabolomics* **2012**, *155*, 1999-2009.

(44)   Gu, L. P.; Jones, A. D.; Last, R. L. *Plant Journal* **2010**, *61*, 579-590.

(45)   Bunzel, M.; Ralph, J.; Marita, J.; Steinhart, H. *Journal of Agricultural and Food Chemistry* **2000**, *48*, 3166-3169.

(46)   Funk, C.; Ralph, J.; Steinhart, H.; Bunzel, M. *Phytochemistry* **2005**, *66*, 363-371.

(47)   Bunzel, M.; Ralph, J.; Funk, C.; Steinhart, H. *European Food Research and Technology* **2003**, *217*, 128-133.

(48)   Bunzel, M.; Ralph, J.; Funk, C.; Steinhart, H. *Tetrahedron Letters* **2005**, *46*, 5845-5850.

(49)   Rouau, X.; Cheynier, V.; Surget, A.; Gloux, D.; Barron, C.; Meudec, E.; Louis-Montero, J.; Criton, M. *Phytochemistry* **2003**, *63*, 899-903.

(50)   Bunzel, M.; Ralph, J.; Bruning, P.; Steinhart, H. *J. Agric. Food Chem.* **2006**, *54*, 6409.

Chapter Four: Profiling of Soluble Neutral Oligosaccharides from Treated Biomass using Solid Phase Extraction and Liquid Chromatography-Multiplexed Collision Induced Dissociation-Mass Spectrometry

4.1 Abstract

Thermochemical pretreatment of cellulosic biomass improves cell wall enzymatic digestibility, while simultaneously releasing substantial amounts of soluble oligosaccharides. Profiling of oligosaccharides released during pretreatment yields information essential for choosing glycosyl hydrolases necessary for cost-effective conversion of cellulosic biomass to desired biofuel/biochemical end-products. In this chapter, a methodology is presented for profiling of soluble neutral oligosaccharides released from ammonia fiber expansion (AFEX)-pretreated corn stover. This methodology employs solid phase extraction (SPE) enrichment of oligosaccharides using porous graphitized carbon (PGC), followed by high performance liquid chromatography (HPLC) separation using a polymeric amine based column (Prevail Carbohydrate ES) and electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS) in both positive and negative modes. For structural elucidation on the chromatographic time scale, nonselective multiplexed collision-induced dissociation was performed for quasi-simultaneous acquisition of accurate molecular and fragment masses of neutral oligosaccharide in a single analysis. These analyses revealed presence of glucans up to degree of polymerization (DP) 22 without side-chain modifications. Additionally, arabinoxylans up to DP=6 were detected in the pretreated biomass samples (post-enzymatic digestion). Cross-ring fragment ion abundances were consistent with assignment of linkages between sugar units in glucans and also xylose backbone in arabinoxylans as 1-4 linkages. Comprehensive profiling of soluble oligosaccharides also demonstrated that arabinoxylan acetylation was reduced more than 85% post-AFEX treatment.

4.2 Introduction

The biosynthesis of carbohydrate polymers represents one of the most prolific biochemical

141

transformations on Earth. It has been estimated that natural plant biosynthesis generates more than 10^{11} tons of biomass per year [1,2] with around half of this consisting of carbohydrate polymers cellulose and hemicelluloses. Carbohydrate polymers are responsible for plant cell wall structure and strength, storage of biochemical energy in the form of starch, and production of materials with far-reaching application including gelling and emulsifying agents and as drug delivery agents [3]. Humankind exploits only a small fraction of this biomass, and as a result, cell wall polysaccharides are attractive renewable resources. Abundance alone makes cell wall polysaccharides attractive renewable feedstocks for bioenergy, and specialty products.

Plant cell wall oligosaccharides are materials of daunting complexity, being composed of various (5 or 6 carbon) sugar monomers with different degrees and positions of branching, assorted chemical modifications including acetylation and feruloylation [4], and heterogeneity in molecular mass. These factors have strong influence on their solubility and digestibility [5]. Conversion of complex polysaccharides, particularly those from cellulosic biomass, to fermentable monosaccharides is often inefficient owing to chemical modifications including formation of diferulate crosslinks that take place *in vivo* during cell wall assembly. Yields of fermentable products are improved following application of hydrolytic and ammonolytic pretreatments including acid, alkali, and AFEX, which remove acetyl and phenolic acid esters from modified polysaccharides [6]. In order to better predict yields of conversion of cell wall biomass to fermentable sugars and to optimize design of efficient biorefineries, comprehensive profiling of mono- and poly-saccharides generated by pretreatment and during enzymatic processing is needed that goes beyond destructive conversion of glycopolymers to monomeric sugars.

Profiling of oligosaccharides derived from processing of cell walls starts best by defining

the molecular masses of the individual components in a mixture of oligosaccharides. Modern mass spectrometry readily provides molecular mass information through application of soft ionization methods including matrix-assisted laser desorption ionization (MALDI) [7] or electrospray ionization (ESI) [8]. When soft ionization is combined with CID to generate fragment ions, the resulting information allows for characterization of sequences of sugar monomers, linkages between sugars based on cross-ring fragment masses, and presence of branching. This approach has been exploited to investigate structures of oligosaccharides derived from plant tissues [9-13]. Most of these reports describe characterization of products of enzymatic digestions using soft ionization and MS/MS, frequently employing permethylation to improve information content in the mass spectra. One report demonstrated detection of fructans with DP>100 using HPLC based on PGC and electrospray ionization [14]. More recently, the combination of mass spectrometry with ion mobility (IM) separations of ions allowed for discrimination of oligosaccharides by shape, and not just mass [14]. Despite great advances in IM technology, this approach has yet to achieve resolution of the vast array of isomeric oligosaccharides without prior physical separation.

Another fast and powerful tool for structural identification/confirmation of analytes employs quasi-simultaneous acquisition of exact masses at high and low collision energies in a single analysis without mass filtering (MS^E) [15]. Fast data acquisition provided by time-of-flight mass spectrometry allowed for extension of this technique to use more than two collision conditions, termed multiplexed collision-induced dissociation. This approach yields CID mass spectra using multiple collision energies on the chromatographic time scale, and has driven discoveries of new plant metabolites and genes responsible for metabolite accumulation [16,17]. To our knowledge, multiplexed CID has not been reported for oligosaccharides.

Separation of oligosaccharides before mass spectrometry is essential for comprehensive oligosaccharide profiling owing to the complexity of plant oligosaccharide mixtures and the potential for numerous isomers. Separation of neutral oligosaccharides has long presented challenges because the polar functionality of these compounds affords limited opportunities for selective chromatographic retention, and for this reason, reversed phase LC has minimal utility. Stronger interactions between analyte and column must be exploited for analyte retention, and these require column materials with affinity for polar groups.

Some notable successes in separation of oligosaccharides were achieved through use of hydrophilic interaction chromatography (HILIC) [18,19,20.] High-pH anion exchange chromatography (HPAEC) has emerged as another common carbohydrate separation method that takes advantage of partial ionization of oligosaccharides at elevated pH [21-24], but the common nonvolatile mobile phase additives that achieve high pH are incompatible with electrospray ionization unless they are removed post-column. PGC has proved to be a suitable chromatographic stationary phase for retention of very polar compounds owing to both hydrophobic and electronic-type interactions between the analyte and the PGC surface [25,26]. PGC columns have been used to enrich or separate various sugars and sugar polymers including sugar phosphates from *Arabidopsis thaliana* [27], cell wall oligosaccharides [28] and human milk oligosaccharides [29,30]. PGC separations offer the advantage that mobile phases are often compatible with mass spectrometry analyses.

About a decade ago, the introduction of the Prevail Carbohydrate ES column (Alltech), a polymeric column with amine groups, yielded separations that resolved a variety of neutral mono- and oligosaccharides. This column yields separations similar to normal phase chromatography using water and acetonitrile as solvents, and these are compatible with ESI mass

144

spectrometry. Most of the applications of this column reported to date focused largely on analyses of mono- and disaccharides [31-35].

As mentioned above, pretreatment of biomass improves yields of conversion of carbohydrates to fermentable sugars, but the fundamental relationships between the severity of pretreatment and the digestibility of products remains uncertain. Pretreatment processes release complex mixtures of substances including phenolics, Maillard reaction products, and mono- and oligosaccharides [36]. Initial efforts that profiled soluble carbohydrates using a Bio-Rad Aminex 42-A column resolved carbohydrates with DP<5, but larger oligomers were not resolved. Furthermore, chromatographic peak areas of oligosaccharides detected using refractive index and mass spectrometric detection did not account for the total sugar monomers yielded by acid hydrolysis. This finding led us to suspect that losses of oligosaccharides were occurring during sample processing, and a strategy for oligosaccharide enrichment was pursued using SPE based on PGC as a prelude to LC/MS profiling.

In the current study, soluble neutral oligosaccharides including xylans and glucans were profiled in extracts of AFEX pretreated corn stover (AFEXTCS). Enrichment of larger oligomers using PGC-solid phase extraction (PGC-SPE) followed by analytical separation using a Prevail Carbohydrate ES column coupled to multiplexed collision-induced dissociation mass spectrometry provided a fast technique that yielded rich information for comprehensive profiling of neutral soluble oligosaccharides in extracts of untreated and pretreated plant material.

4.3. Methods

4.3.1 AFEX treatment

Corn stover (harvested in 2002 at the Kramer farm in Wray, CO, USA) was incubated in a reactor under pressure with liquid ammonia at 130 $^{\circ}$C with 60% moisture. The loading ratio was

1:1 (w/w) NH$_3$-to-biomass, and total residence time was 15 min. Detailed protocols can be found in our previous publications [36].

4.3.2 Enzyme hydrolysis

Oat spelt xylan (50 mg) mixture (Sigma-Aldrich, USA), 50 mg of birch wood xylan mixture (Sigma-Aldrich, USA), 50 mg of AFEXTCS and UTCS and 0.5 mL of AFEXTCS extract were all separately incubated with 25 µL of food grade xylanase, RE4 (40 mg/mL, from Genencor, USA) for 12 hours at 50 ^{o}C. Final volumes in all cases were 1 mL using sodium acetate buffer at pH=5. In each case, samples were injected directly for LC/ESI-TOF-MS analysis without further treatment.

4.3.3 Enrichment of oligosaccharides by PGC-SPE

PGC cartridges with bed weights of 1000 mg used for enrichment of oligosaccharides were HyperSep Hypercarb SPE columns from Thermo Fisher Scientific (USA). For evaluating retention capacities of cartridges, 7 mL of diluted (6 mg/mL total sugar) corn syrup (Kroger, USA) was loaded to one cartridge and non-retained materials after loading were analyzed using LC/ESI-TOF-MS for comparison with the diluted corn syrup sample. Methanol and acetonitrile were both tested for eluotropic strength, and methanol was a stronger eluent for oligosaccharides on the PGC cartridges. For enriching oligosaccharides in corn stover hydrolysate, 6 mL of hot water extract of AFEXTCS (20 mL water/g treated biomass) was loaded onto each cartridge (48 mL total of extract was loaded to 8 cartridges in parallel). 5 mL water was used to wash non- and poorly-retained compounds from the PGC cartridges. After washing, 30 mL of methanol was used to elute retained oligosaccharides. Eluants (total of 240 mL from 8 cartridges) were evaporated under vacuum and pooled to yield a final volume of 6 mL. This concentrated fraction

was injected directly for LC/ESI-TOF-MS analysis.

4.3.4 LC/ESI-TOF-MS separation and identification of oligosaccharides

Oligosaccharides from all extracts (enriched, enzyme digest or non-processed) were analyzed using a Prevail Carbohydrate ES column (150 × 2 mm, 5 µm; Alltech Associates, Deerfield, IL, USA) coupled to the mass spectrometer. The LC/MS system used in this work consisted of three Shimadzu HPLC pumps (LC-20AD) coupled to a Waters LCT Premier Time-of-Flight Mass Spectrometer (TOF-MS). Gradient elution was performed based on solvent A (0.15% aqueous formic acid) and B (acetonitrile) over 30 minutes. Total solvent flow was maintained at 250 µL/min, and gradient elution was performed using the following solvent compositions: initial, 95 % B, held for 1 min; linear gradient to 70 % B at 8 min and then to 50 % B at 18 min; hold at 50% B until 25 min, sudden increase to initial condition at 25.01 min and final hold at this composition until 30 min. Positive and negative mode electrospray ionization were performed along with multiplexed collision induced dissociation (CID) by switching among three different AP1 (Aperture 1 voltage in LCT Premier mass spectrometer) voltages (15, 40, and 65 V). Processing of LC/MS data was accomplished using MassLynx v. 4.1 software.

4.3.5 Acid hydrolysis and analysis of monosaccharides

Monomeric carbohydrate concentrations were measured before acid hydrolysis, while oligosaccharide concentrations were calculated by taking sugar concentrations measured after acid hydrolysis and subtracting the monomeric sugar concentrations. For acid hydrolysis, 500 µL of extract and 17.6 µL of 72% sulfuric acid were placed in tightly capped 10 mL culture tubes which were vortexed and placed in a preheated block heater (EL-02 Elite, Major Science, Saratoga, CA) at 121 ^{o}C for 1 hour, after which they were cooled on ice to room temperature.

For sugar analysis, the fractions were filtered using 0.22 μm PES syringe filters (Whatman, Piscataway, NJ), and were analyzed using a Prominence HPLC system (Shimadzu, Columbia, MD) using an Aminex HPX-87H HPLC carbohydrate analysis column, a Cation H guard column (Bio-Rad, Hercules, CA) and a RID-10A refractive index detector (Shimadzu, Kyoto, Japan). 5 mM aqueous H_2SO_4 was the mobile phase, delivered at a flow rate of 0.6 mL/min and 50 $^\circ$C. The injection volume was 10 μL.

4.4 Results and discussion

4.4.1 Optimization of conditions for PGC-SPE and enrichment of oligosaccharides

Initial LC/MS profiling of carbohydrates released from corn stover during AFEX processing detected low mass (DP<6) oligosaccharides [36], but mass balance calculations suggested that the analytical result failed to account for significant amounts of glucans. Working from the assumption that matrix constituents might suppress ionization, a fractionation scheme was developed, based on PGC-SPE (1000 mg cartridge), to remove matrix components and enhance the range of detected oligosaccharides. Based on the PGC-SPE manufacturer's guidelines, it was anticipated that the retention capacity of a PGC-SPE cartridge would be 10-50 mg of sugars. The performance of this fractionation was tested by loading 7 mL of a 6 g/L aqueous corn syrup solution (42 mg total sugars) onto the PGC cartridge. The original solution and the non-retained material were analyzed with LC/TOF-MS in negative ion mode. Figure 4.1-a shows the extracted ion LC/MS chromatogram of [M+Cl]$^-$ adduct ions with masses corresponding to corn syrup hexose oligomers [37], and Figure 4.1-b shows the corresponding chromatogram of the non-retained material. Comparing the total peak areas for [M+Cl]$^-$ ions from the identified hexose oligomers (DP=2-13) showed size-dependent retention of oligosaccharides, ranging from 67% for DP=2 to 100% for DP=10 and beyond, that is consistent

148

with a recent independent report [28]. To determine the effect of pH on retention of sugars, corn syrup was diluted with aqueous formic acid or ammonium hydroxide to give a pH of 2.3, 6.7 or 11.3. Increasing pH from 2.3 to 6.7 yielded 20-30% more retention as based on integrated peak areas. However, a further increase to pH 11.3 only yielded an additional 5% increase in retention, mainly for sugars of DP<5.
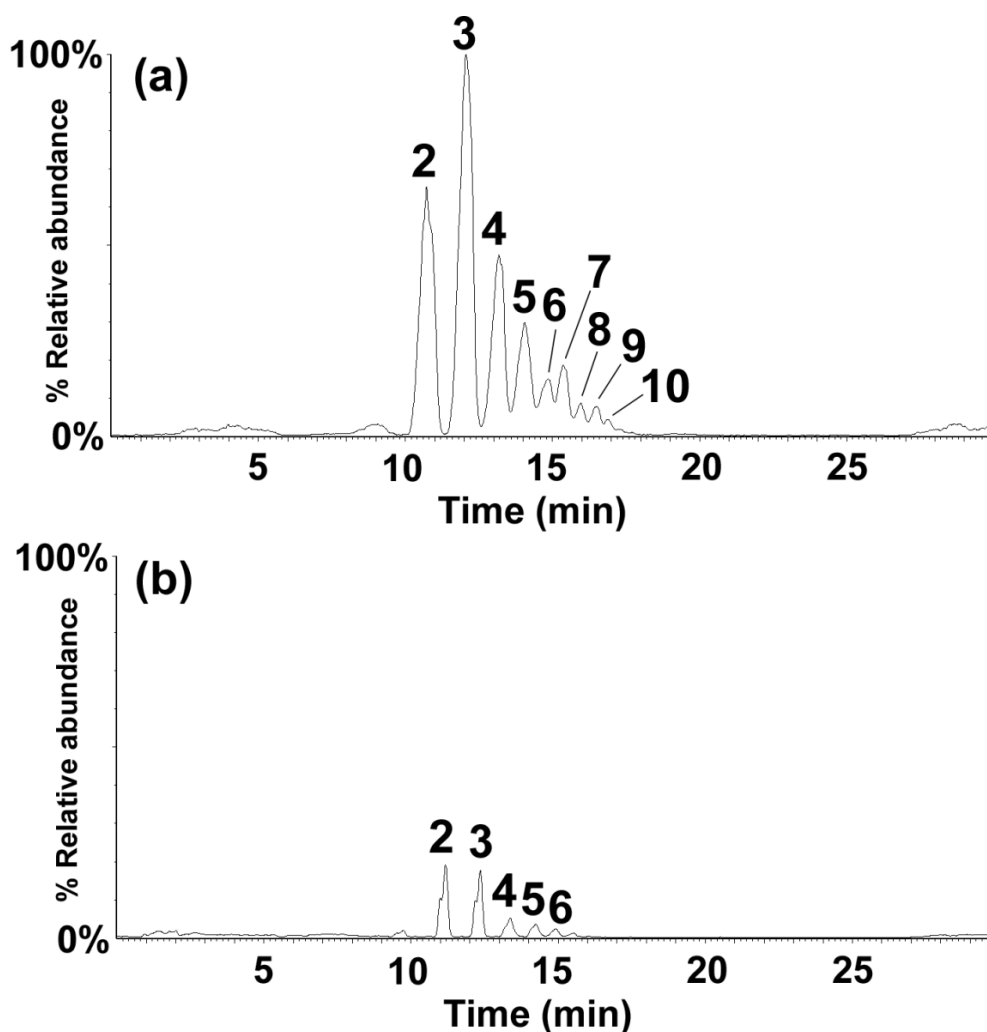
Figure 4.1. LC/ESI-TOF-MS extracted ion chromatograms of [M+Cl]⁻ ions from oligomers of hexose (DP=2-13) in corn syrup (a) and non-retained portion of corn syrup after passing through a PGC-SPE cartridge (b). Peaks are labeled with the DP value for each oligosaccharide. Scaling of the y-axis was performed to the same absolute signal for both chromatograms.

In order to select solvent for elution of oligosaccharides from PGC cartridges, corn syrup hexose oligomers were again used for method development and optimization. A range of solvents have been suggested for elution of retained solutes from PGC, and selection of solvents was guided by a desire to elute in solvents compatible with LC separation and mass spectrometric detection. Though acetonitrile has been used to elute oligosaccharides from PGC chromatography columns [28], our efforts compared oligomer recoveries using acetonitrile with methanol. After loading 7 ml of 6 g/L corn syrup to a PGC-SPE cartridge, elution with 20 mL methanol yielded >99% recovery of hexose oligomers from DP=6-13 as determined using LC/MS, whereas elution with the same amount of acetonitrile yielded only ~80% of the material loaded on column.

AFEXTCS includes a complex range of substances including phenolics and various polar constituents [36] that can compete for retention on SPE cartridges. Based on preliminary surveys, it was observed that loading more than 6 mL of AFEXTCS led to undesirable breakthrough of short-chain oligosaccharides (DP=2-4). Therefore, 6 mL of extract was loaded onto each PGC-SPE cartridge. After loading, 5.0 mL of water was used to elute salts and less retained substances, followed by elution of oligosaccharides using 30 mL methanol. A total of 48 mL AFEXTCS extract was loaded on 8 PGC-SPE cartridges in parallel, and methanol-eluted material (240 mL) was combined and reduced to 6 mL (an 8-fold concentration relative to the original extract) under vacuum for LC/ESI-TOF-MS analysis.

When unprocessed AFEXTCS extracts were concentrated by a factor of 8 by evaporation, LC/ESI-TOF-MS analyses of the solutions failed to detect oligohexoses larger than DP=5. During solvent evaporation of the crude extracts, substantial quantities of precipitate were formed, and this process was suspected to remove larger oligosaccharides by sedimentation as

150

the solvent evaporated. Filtered crude extracts (0.2 µm syringe filters) generated insoluble material upon standing at 25 °C for 24 hours. It is proposed that colloidal particles in the crude extracts adsorb oligosaccharides and coalesce to form insoluble precipitates. However, when freshly-prepared extract was loaded onto PGC cartridges, oligosaccharides were retained and eluted using methanol, with colloidal particles being retained in the column. Figure 4.2-a and 4.2b compare the total ion LC/MS chromatograms of AFEXTCS extract and the enriched oligosaccharides from the same extract. The results demonstrate that enriching larger oligosaccharides through removal of adsorptive constituents improves stability of oligosaccharide-containing solutions, yields dramatic increases in recoveries of DP>4, and allows for improved profiling of oligosaccharides released from biomass.
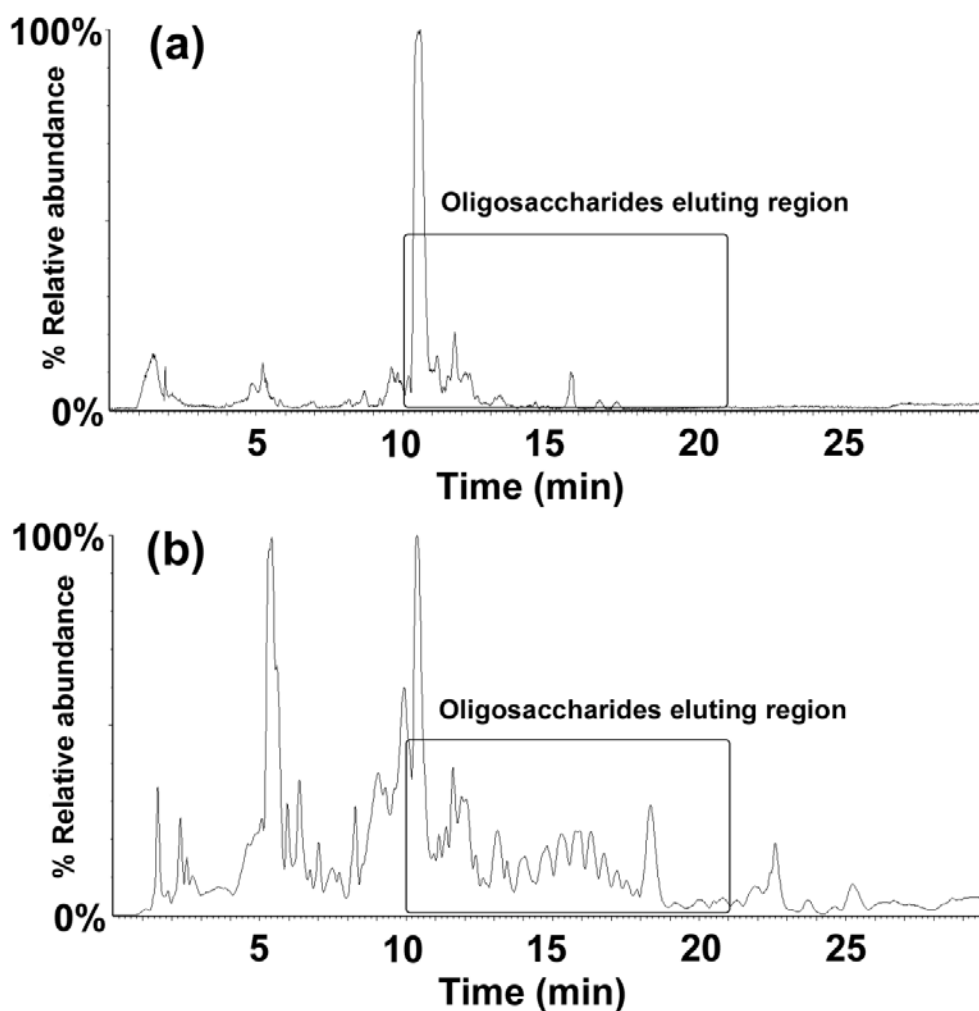
Figure 4.2. LC/ESI-TOF-MS total ion chromatograms of (a) unprocessed AFEXTCS extract and (b) PGC-SPE enriched oligosaccharides from the same extract. Scaling of the y-axis was performed to the same absolute signal for both chromatograms.

4.4.2 Hexose oligomers from the AFEXTCS extract

Neutral oligosaccharides eluted between 10-21 min using the Prevail Carbohydrate ES column, and ESI mass spectra were generated in both positive and negative ion modes. For representation of a more comprehensive assessment of oligosaccharide content, spectra were summed over portions of this retention time window (Figure 4.3). Combinations of accurate

152

mass measurements with observations of a series of ions differing in mass by 162 Da and generation of fragment ions using nonselective multiplexed CID [17] suggested presence of hexose oligomers of DP=2-22 in the enriched oligosaccharide fraction from AFEXTCS extract.
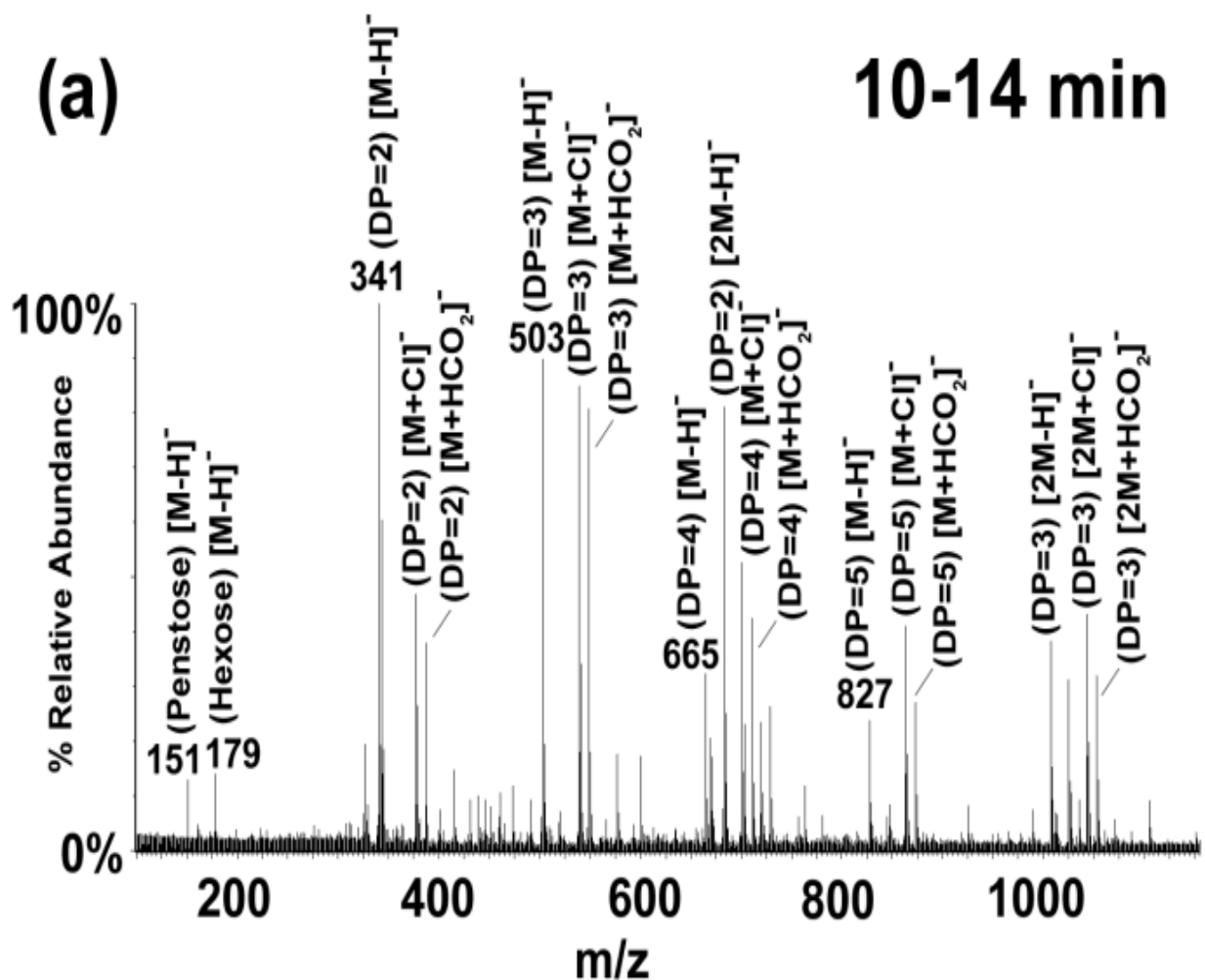
Figure 4.3. LC/ESI-TOF-MS spectra of PGC-SPE enriched oligosaccharides from AFEXTCS extract (Figure 4.2-b) averaged over retention time windows of (a) 10-14 min (b) 14-17 min (c) 17-19 min and (d) 19-21 min. These spectra were acquired using the lowest collision potential (15 V). Only [M-H]$^-$ ions are labeled with corresponding *m/z* values.

Figure 4.3 (cont'd)

(b) 14-17 min
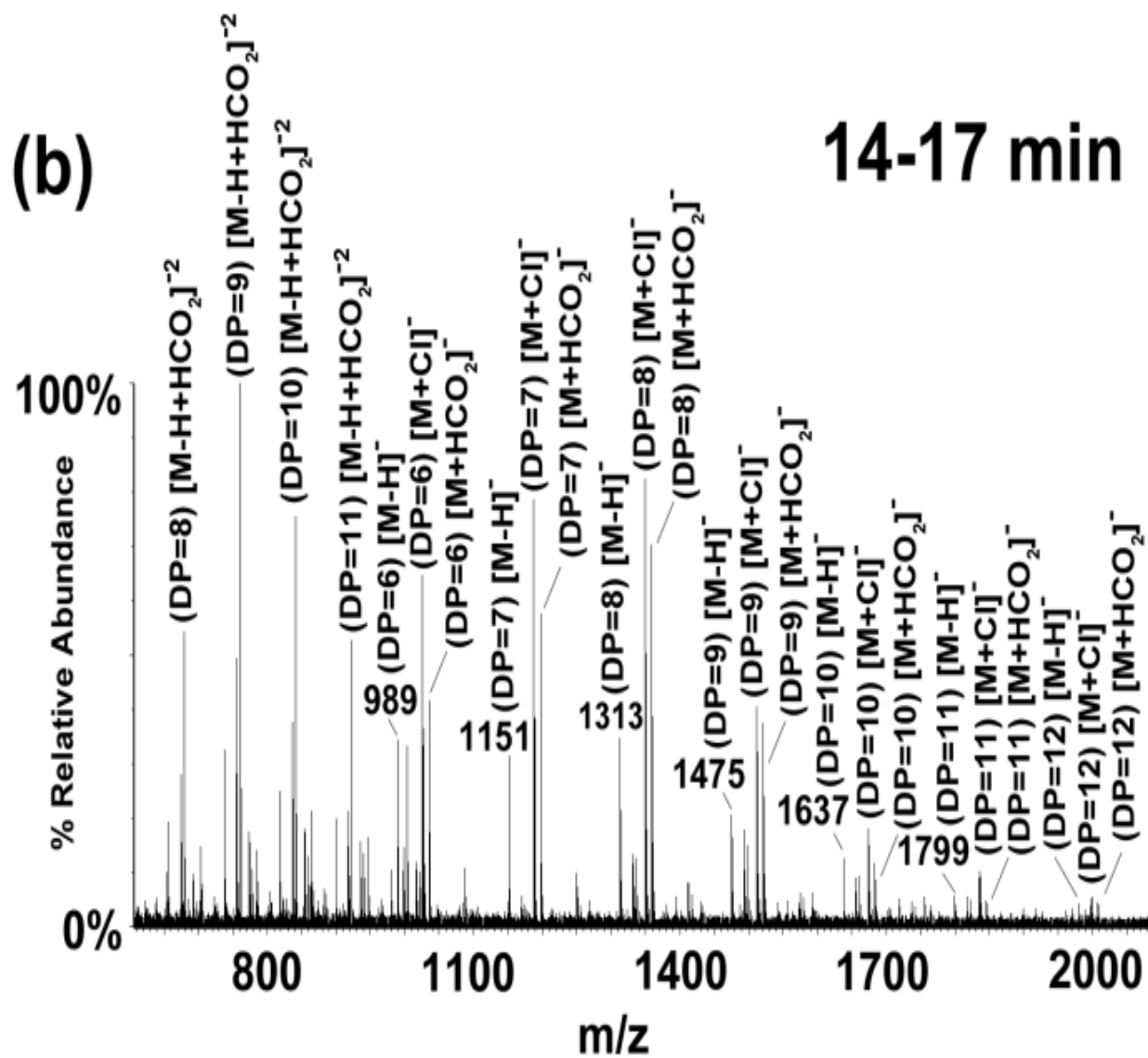
Figure 4.3 (cont'd)

Figure 4.3 (cont'd)

(d)

19-21 min
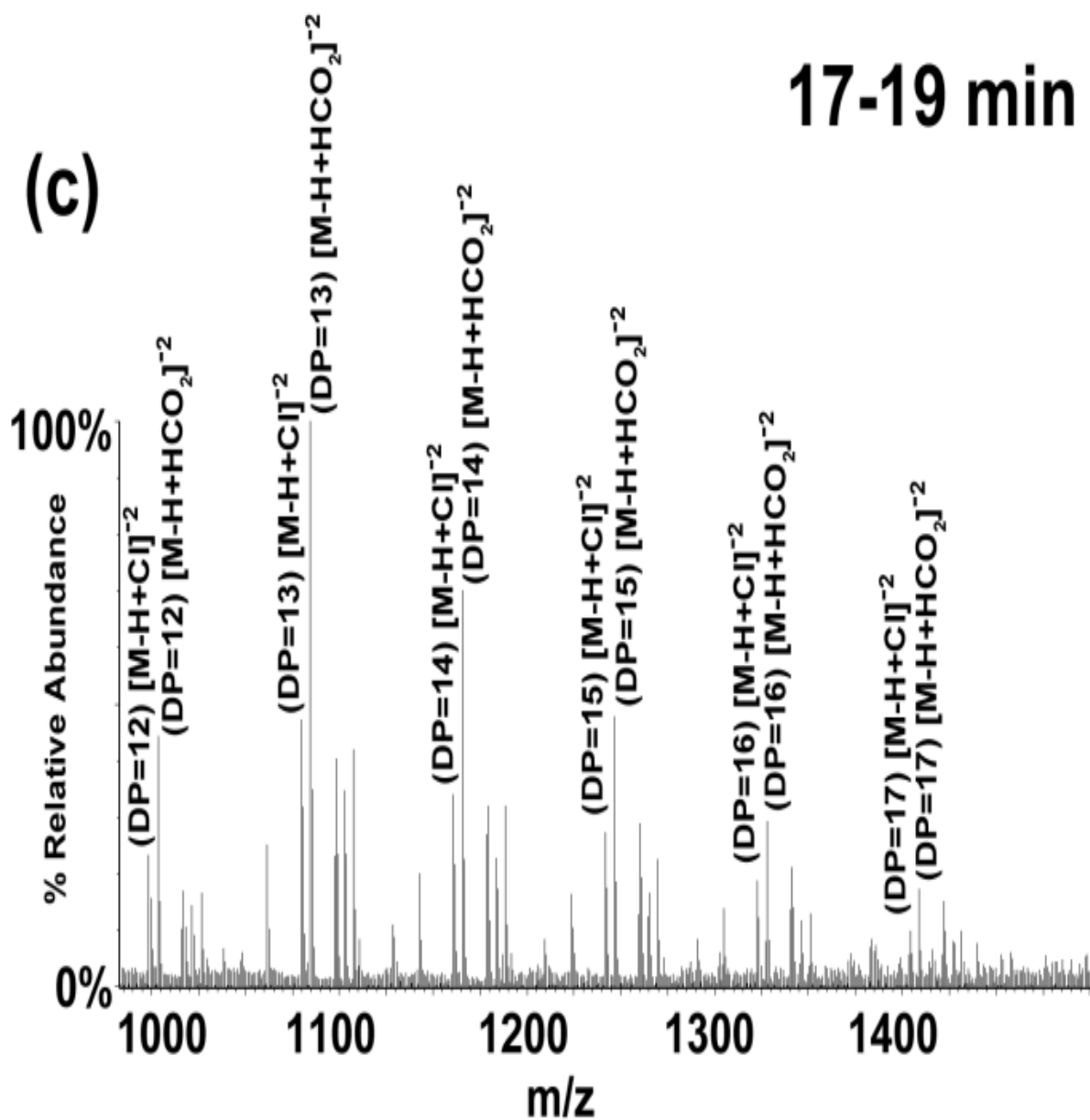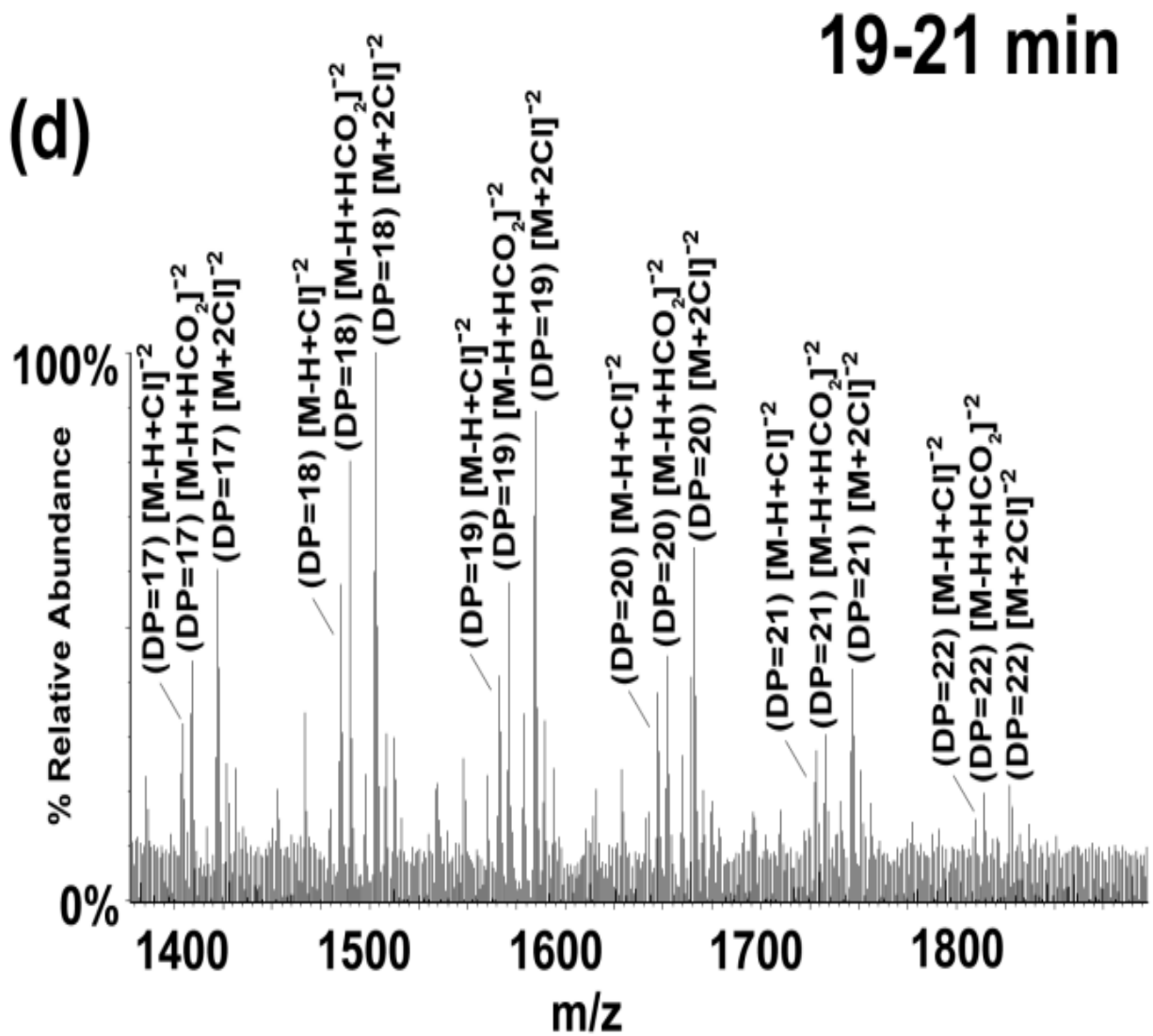
Oligosaccharides that dominated the mass spectra generated multiple ions indicative of individual molecular masses, including $[M-H]^-$, $[M+HCO_2]^-$ and $[M+Cl]^-$ ions for smaller hexose oligomers (DP<10) and doubly charged species such as $[M-H+Cl]^{2-}$, $[M-H+HCO_2]^{2-}$ and $[M+2Cl]^{2-}$ for larger oligomers (DP>10). Figure 4.4 demonstrates separation of oligomers in the form of extracted ion chromatograms for singly- and doubly-charged chloride adducts of oligohexoses up to DP=22 detected in AFEXTCS enriched fractions. Based on peak areas, abundances of hexose oligomers showed monotonic decrease with increasing DP.

While molecular masses of the dominant constituents corresponded to oligohexoses, the identities of monomeric subunits were not distinguished from molecular mass alone. To identify the subunit composition of the unprocessed AFEXTCS extract, acid hydrolysis was performed followed by analysis of monomeric sugars using HPLC and refractive index detection. The hydrolyzed material showed increases, relative to the nonhydrolyzed AFEXTCS extract, in concentrations of glucose, arabinose and xylose (Table 4.1). Since glucose was the only hexose detected in the hydrolysis products, it was concluded that the overwhelming majority of hexose oligomers in the AFEXTCS extract were glucans.

Table 4.1 Carbohydrate content in nonhydrolyzed (free sugar concentration) and acid hydrolyzed (total) AFEXTCS extract determined using HPLC and refractive index detection. The oligomeric sugar concentration is the difference between these values.

	Glucose (g/L)	Xylose (g/L)	Arabinose (g/L)
Free Sugar Concentration	0.12 ± 0.02	0.09 ± 0.01	0.07 ± 0.00
Total Sugar Concentration	0.54 ± 0.04	1.99 ± 0.09	0.48 ± 0.01
Oligomeric Sugar Concentration	0.42 ± 0.02	1.90 ± 0.08	0.41 ± 0.01
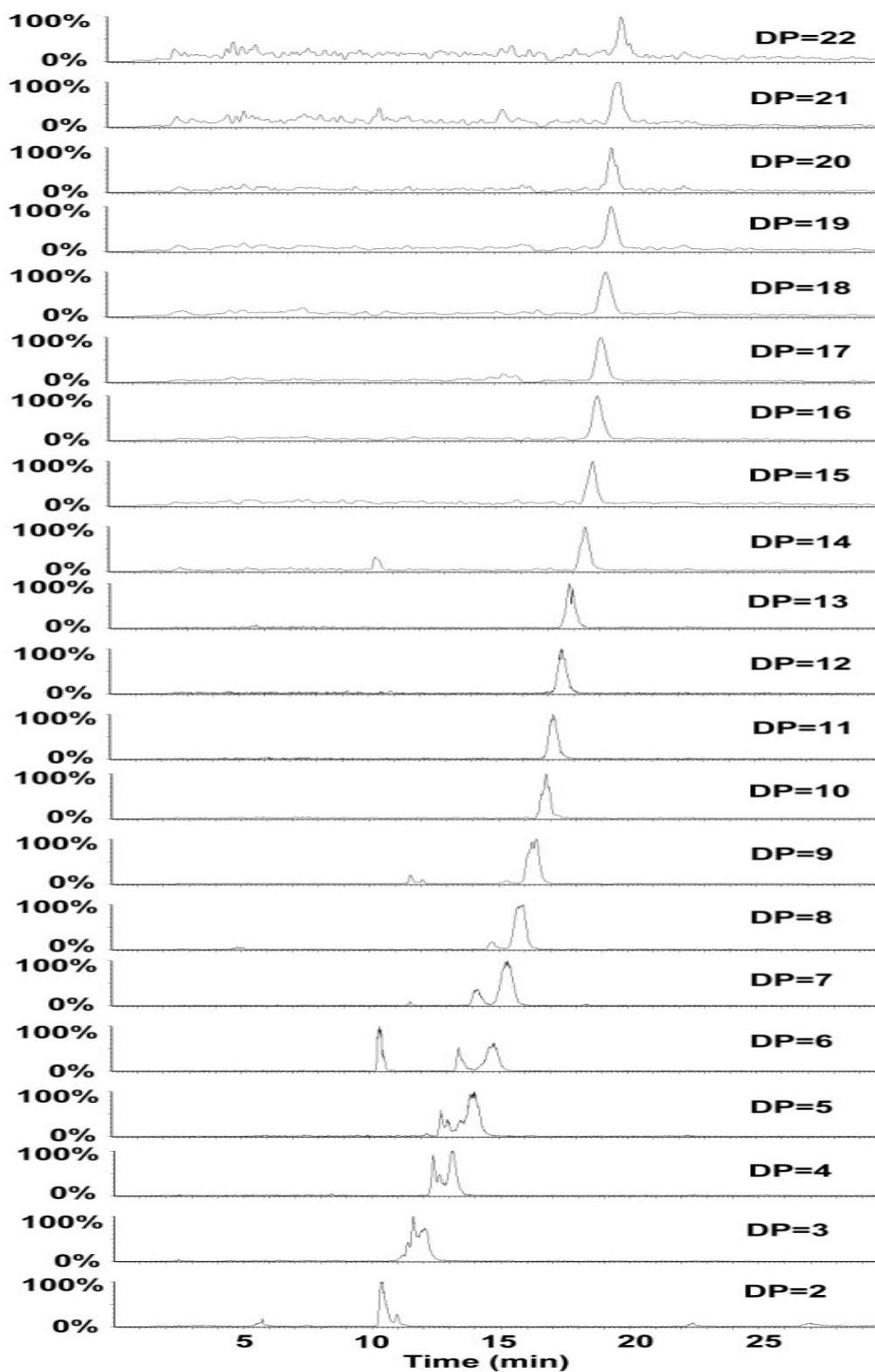
Figure 4.4. LC/ESI-TOF-MS extracted ion chromatograms ([M+Cl]$^-$ for DP=2-10 and [M-H+Cl]$^{-2}$ for DP=11-22) of hexose oligomers from the PGC-SPE enriched AFEXTCS extract.

4.4.3 Non-selective CID spectra of glucans

MS/MS spectra of oligosaccharides reveal valuable information about sequence, linkage and branching in the form of characteristic fragment ion masses. CID of oligosaccharide ions results in both cleavage of the glycosidic bonds (B, C, Y, and Z ions, Figure 4.5-a) and cross-ring fragmentation (A and X ions, Figure 4.5-a) [38]. Consecutive cleavages of the glycosidic bonds leads to sequence information but not branching or linkage type between the sugar units. However, type of linkage between sugar units can often be identified from characteristic cross-ring fragment ion masses. Figure 4.5 shows three CID spectra acquired at different collision energies for the oligosaccharide eluting at retention time 16.1 min (Figure 4.4, DP=8). At the lowest collision energy (15 V), the mass spectrum displays singly- and doubly-charged ions indicative of the oligosaccharide molecular mass. Mass spectra obtained under elevated CID voltages (aperture 1 (AP1); Figure 4.5-c and 4.5-d) of 40 and 65 V respectively were quasi-simultaneously acquired in a single analysis. Acquisition of mass spectra under multiple conditions from a single injection can be achieved because of the fast acquisition speed of the TOF analyzer (approx. 2×10^{4} spectrum transients/s). As expected, more fragment ions were observed for glucan with DP=8 at higher AP1 voltages. At an intermediate collision energy (40 V), the most prominent fragment ions correspond to C-type fragments, and their masses provide a useful indication of oligosaccharide sequences, and their monotonic increase in abundance with chain length suggest a linear oligosaccharide a suggested by Fernandez, Obel, Scheller & Roepstorff, 2004 [11]. At the highest collision energy (65 V), cross-ring (A-series) fragments were of sufficient abundance to suggest a linear oligosaccharide, consistent with 1,4-linkages based on the consistent presence of ^{0,2}A and ^{2,4}A fragment ions [39-43] which are 60 and 120 Da lighter than the C-series ions, and the absence of ^{0,3}A fragment ions that would have indicated 1,6-linkages [39,40,43]. Neutral losses of 78 Da corresponding to ^{0,2}A-H_2O were also

observed from [M-H]$^-$ precursors in β-1,4-linked oligosaccharides in previous studies [39,42,44]. Fragments from reducing and non-reducing end in the CID spectra in Figure 4.5-d could not be distinguished, as no derivatization was performed to identify the reducing end.

Although the multiplexed CID approach is non-selective (e.g. all ions are subjected to CID, and no specific *m/z* value was selected for fragmentation), the CID spectrum in Figure 4.5-d exhibits all of the abundant fragments generated by selective CID of *m/z* 1313 ([M-H]$^-$) using a linear ion trap mass analyzer (Q-Trap 3200). One of the major benefits of performing non-selective CID derives from collisional activation of all adduct ions and different charge states at the same time, which increases the yield of fragment ions. This is specifically helpful for larger oligosaccharides which ionize in a more diverse collection of singly- and doubly-charged pseudomolecular forms owing to the presence of various anions in the biomass extracts or mobile phase (Figure 4.3a-d).
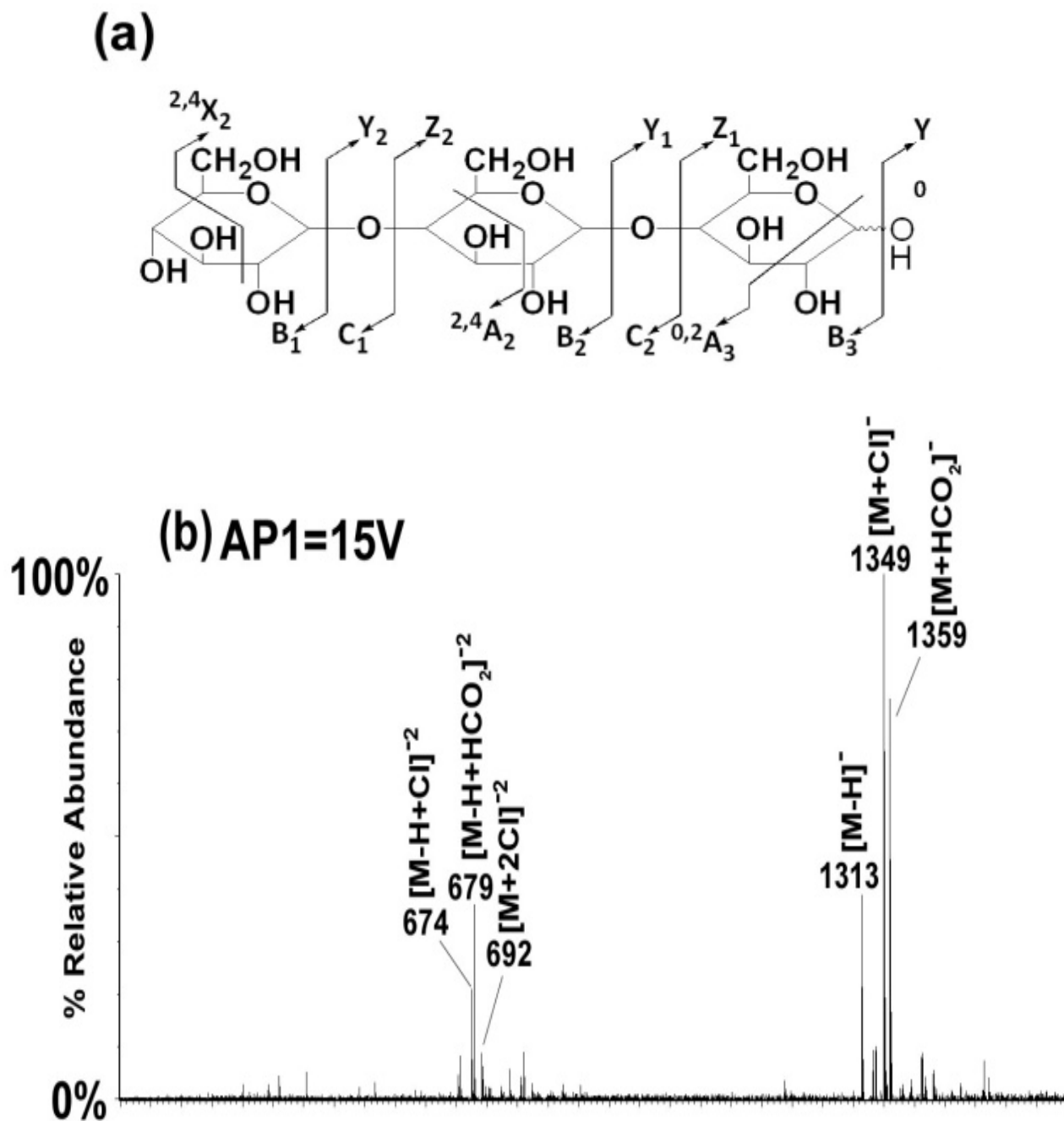
Figure 4.5. Structure of cellotriose, showing the Domon and Costello [38] nomenclature for fragment ions (a) LC/ESI-TOF-MS spectra of the glucan eluting at 16.1 min (DP=8) at three different AP1 (CID) voltages in negative ion mode: (b) 15 V, (c) 40 V, (d) 65 V. Peaks (d) labeled with black circles exhibit masses consistent with their annotation as B ions formed upon loss of water from corresponding C ions. ^{0,2}A and ^{2,4}A fragment ions have *m/z* values of corresponding C ion minus 60 or 120 units, respectively.

161

Figure 4.5 (cont'd)
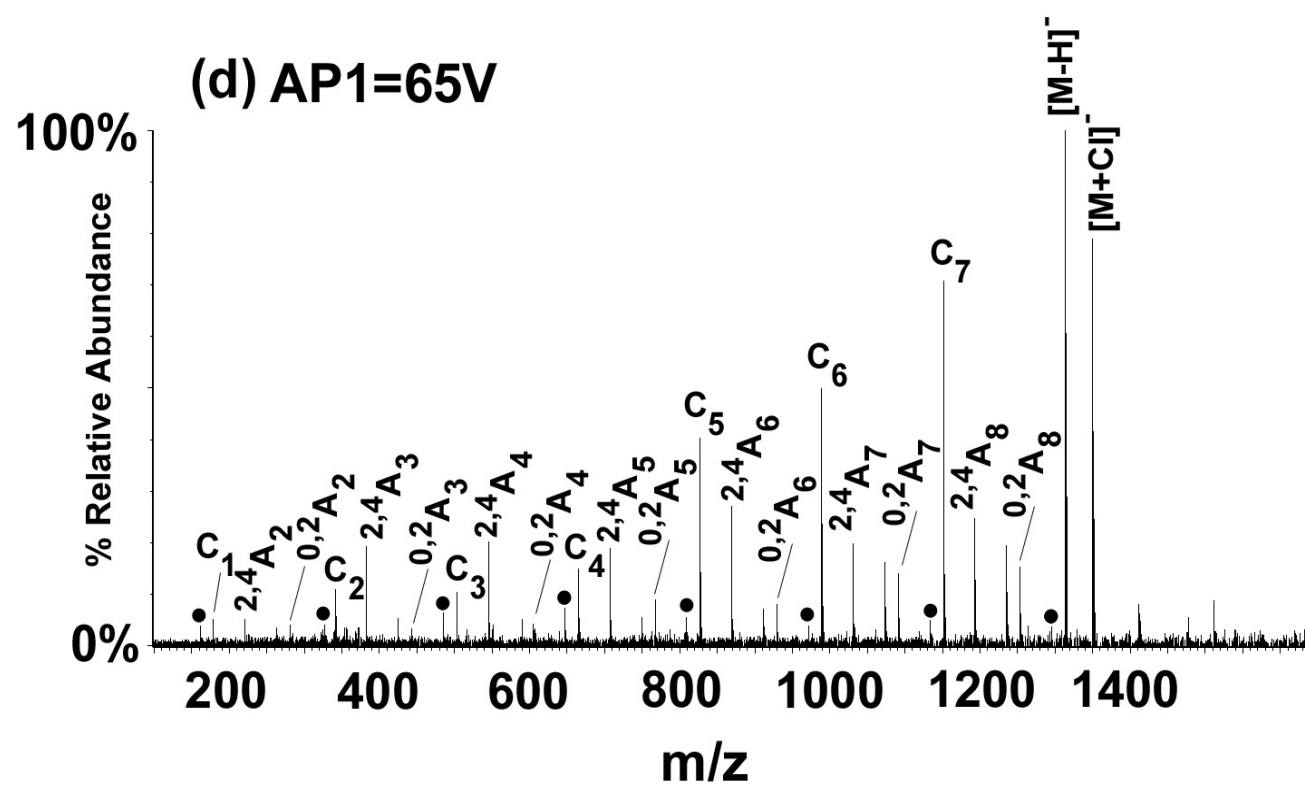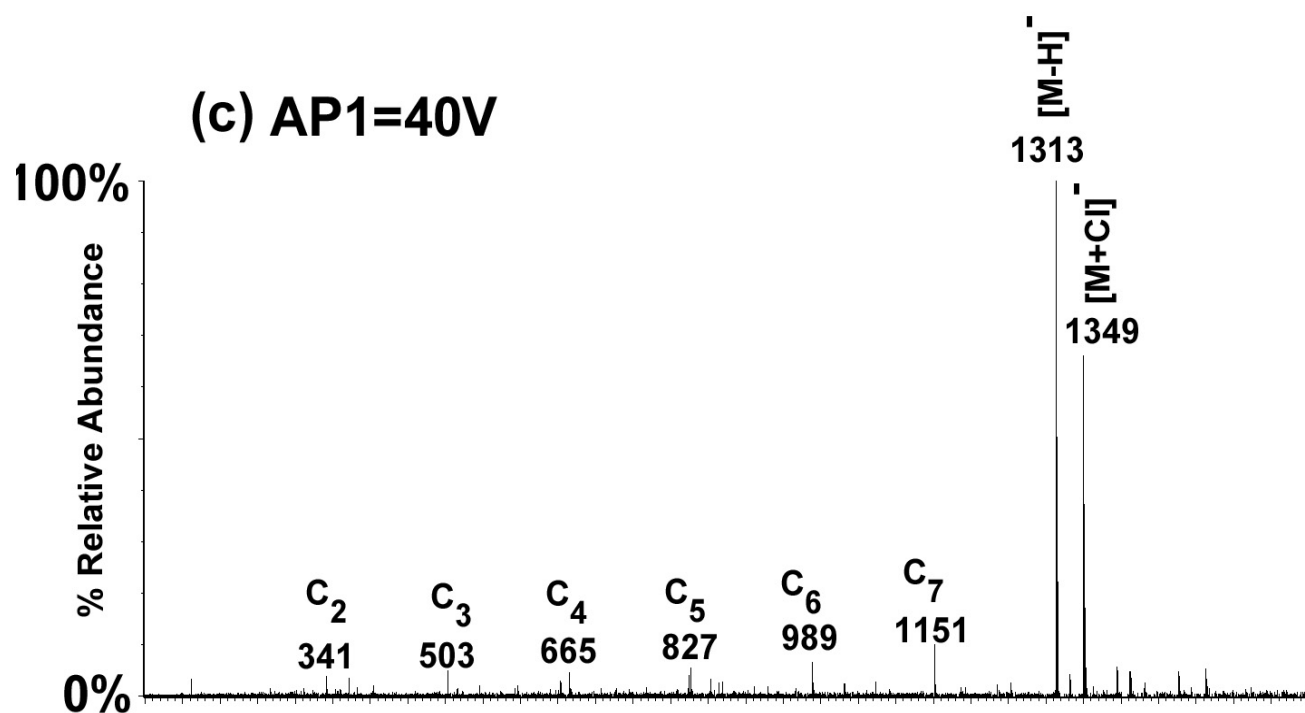
4.4.4 Pentose Oligomers from the AFEXTCS extract

Despite results from acid hydrolysis of AFEXTCS extract that determined 1.9 g/L of xylose and 0.4 g/L of arabinose in oligomeric forms (Table 4.1), LC/MS results yielded no clear evidence for pentose oligomers in either the original extract of the PGC-enriched fraction. Based on these findings, it was concluded that in contrast to glucans, soluble arabinoxylans released by AFEX treatment have high DPs (perhaps beyond the mass range acquired) or significant heterogeneity among numerous chemical forms, and need to be digested enzymatically to reduce molecular mass and complexity before detection using mass spectrometry. Characterization of more complex plant oligosaccharides has relied on enzymatic digestion [45-48] to reduce the complexity of the oligosaccharide forms to allow individual digestion products to be characterized. In keeping with this tradition, AFEXTCS extract was treated with endoxylanase in order to cleave high DP arabinoxylans. After enzymatic digestion, a flow injection analysis of AFEXTCS extract showed ions consistent with a range of arabinoxylans from DP=2 to DP=6. Figure 4.6-a displays a mass spectrum of endoxylanase digest of AFEXTCS extract using positive mode ESI. Identification was based on accurate mass measurements and comparison with commercial arabinoxylans from oat spelt and birch wood after the same enzymatic digestion process. Both ESI positive [11,49] and negative [44,49] modes were used for ionization of arabinoxylans in previous reports; however, positive mode ESI generated more abundant ions for digested arabinoxylans in AFEXTCS extract.

Support for structure annotation of the endoxylanase product was generated from multiplexed CID mass spectra. Figure 4.6-b shows a CID mass spectrum of an arabinoxylan (DP=6) identified from the AFEXTCS extract after enzymatic digestion at high CID voltage (65 V). Similar CID spectra for $[M+Na]^+$ of arabinoxylans with DP=6 were reported by Fernandez *et al.* [12]. Although formation of $^{0,2}A$ (loss of 60) and $^{0,3}A$ (loss of 90) fragment ions was

consistent with (1-4) linked pentose units [12,41,43,50,51] branching information could not be obtained as no derivatization (e.g. O-methylation) nor further selective MS^n were performed. However, based on the ratio of xylose to arabinose in corn stover cell walls (Table 1), it is of high probability that DP=6 contains at least one arabinose unit, and the prominence of the Y_4 fragment ion relative to other Y-ions is consistent with this conclusion. In summary, based on acid hydrolysis, accurate mass measurements, non-selective CID spectra and literature precedent, oligomers of pentose from the AFEXTCS extract were assigned as arabinoxylans with (1-4) linked backbone.
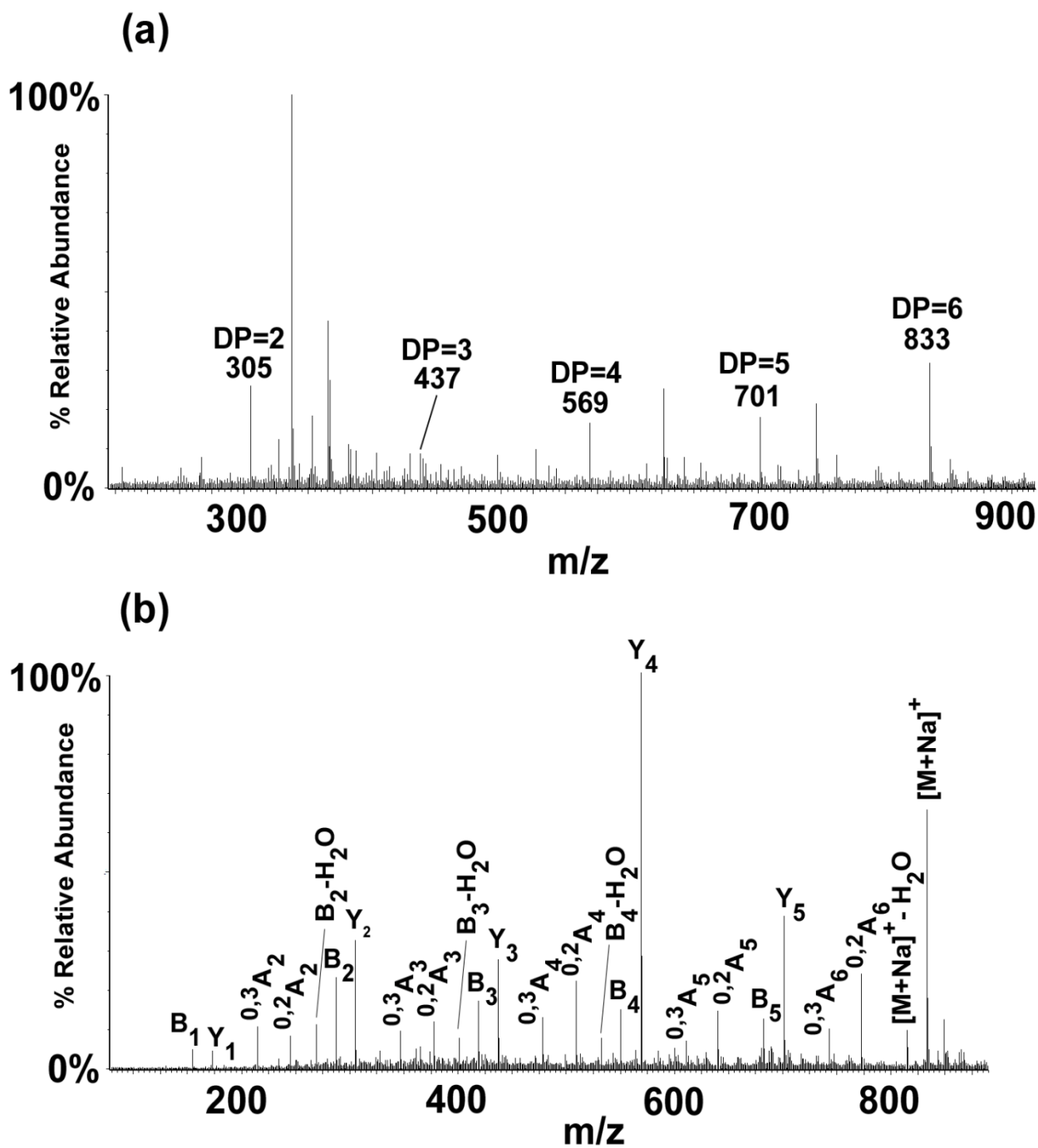
Figure 4.6. (a) ESI-TOF-MS spectrum of endoxylanase-digested AFEXTCS extract analyzed using flow injection in positive ion mode showing [M+Na]$^{+}$ ions from arabinoxylans ranging from DP=2 to DP=6. (b) LC/ESI-TOF-MS spectrum of arabinoxylan with DP=6 at elevated AP1 (CID) voltage (65 V) in positive ion mode.

It is known that *O*-acetyl units are attached to hemicellulose arabinoxylans in different hardwoods [45,52-54]. Cell wall acetylation has appreciable impact on enzymatic digestion as was shown in corn stover digestibility after different biomass treatment processes [52]. The acetyl groups in acetylated cell wall glycopolymers are estimated to account for 35 mg/g dry weight in corn stover [36]. AFEX treatment yields >85% removal of acetyl groups via formation of acetic acid and acetamide through hydrolysis and ammonolysis reactions respectively [36,55]. Solvolytic formation of acetic acid or acetamide depends on the severity of AFEX pretreatment conditions, which are otherwise mild compared to acid pretreatment methods [36]. In order to assess release of soluble arabinoxylans and removal of acetyl groups, UTCS and AFEXTCS were both digested with endoxylanase for 12 hours, and released arabinoxylans were profiled using LC/ESI-TOF-MS. A tentative comparison of peak areas of acetylated and non-acetylated arabinoxylans showed removal of >90% of the acetyl groups on detected arabinoxylans (DP=2-6) after AFEX treatment. Evidence of acetyl group removal is presented in Figure 4.7, which shows four extracted ion chromatograms for acetylated and unmodified arabinoxylans in AFEXTCS and UTCS after enzymatic digestion. Figure 4.7-a documents that arabinoxylans with DP=3 are more abundant in AFEXTCS compare to UTCS, which indicates that AFEX treatment released fermentable sugars from hemicellulose, as was demonstrated earlier using hydrolysis to monosaccharides [36,52]. On the other hand, comparison of levels of an acetylated arabinoxylan (DP=4) in UTCS and AFEXTCS digests documents that AFEX efficiently removed most of the acetyl groups under this specific condition, since the peak area in AFEXTCS (Figure 4.7-c) is less than 10% of the peak in UTCS (Figure 4.7-d). Similar reductions in amounts of poly-acetylated arabinoxylans were also observed.
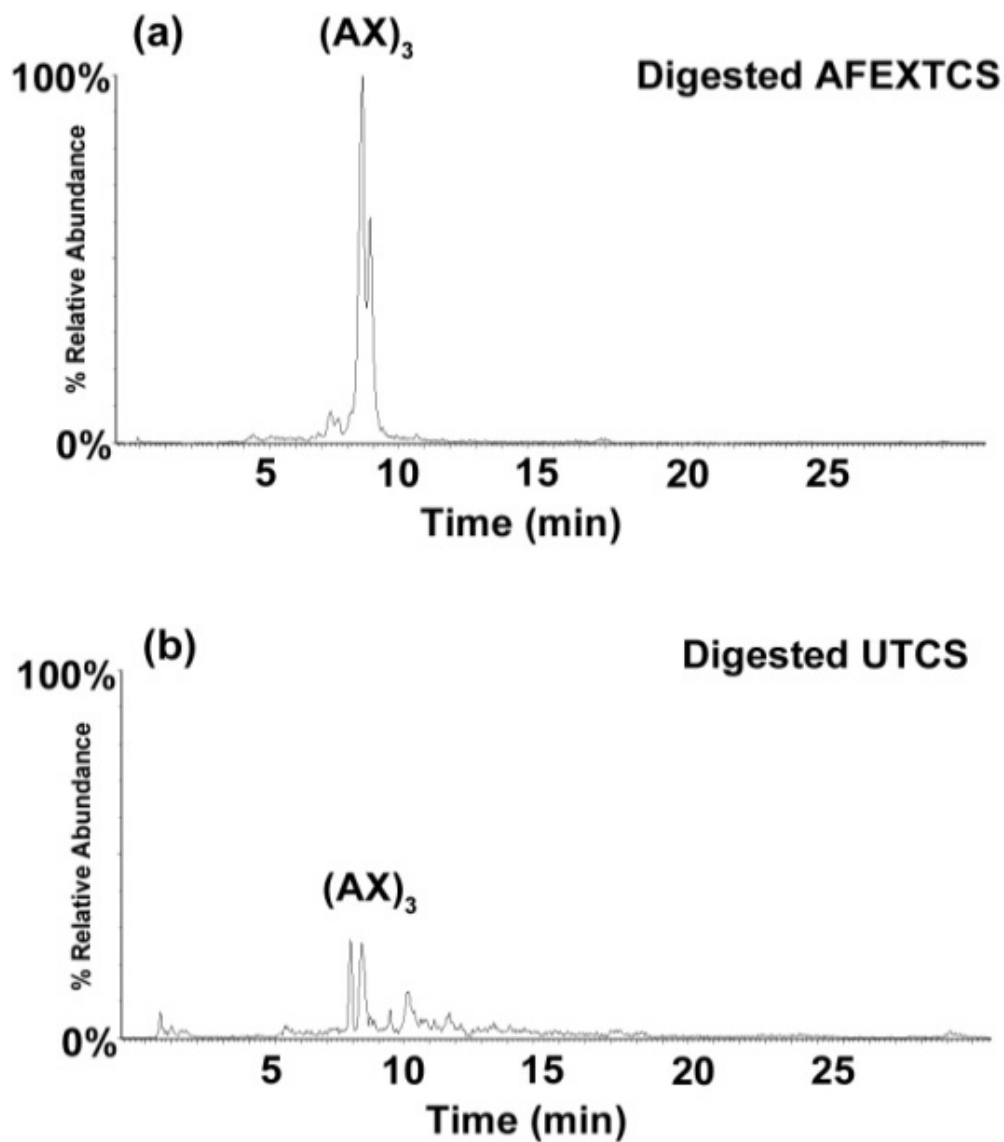
Figure 4.7. LC/ESI-TOF-MS extracted ion chromatograms for arabinoxylans (AX) with DP=3 from endoxylanase digested AFEXTCS (a) and digested UTCS (b), LC/ESI-TOF-MS extracted ion chromatograms for arabinoxylans with DP=4 and one acetyl group (Ac) in digested AFEXTCS (c) and digested UTCS (d). Scaling of the y-axis was performed to the same absolute signal for chromatograms (a) and (b) and also for (c) and (d).

167

Figure 4.7 (cont'd)
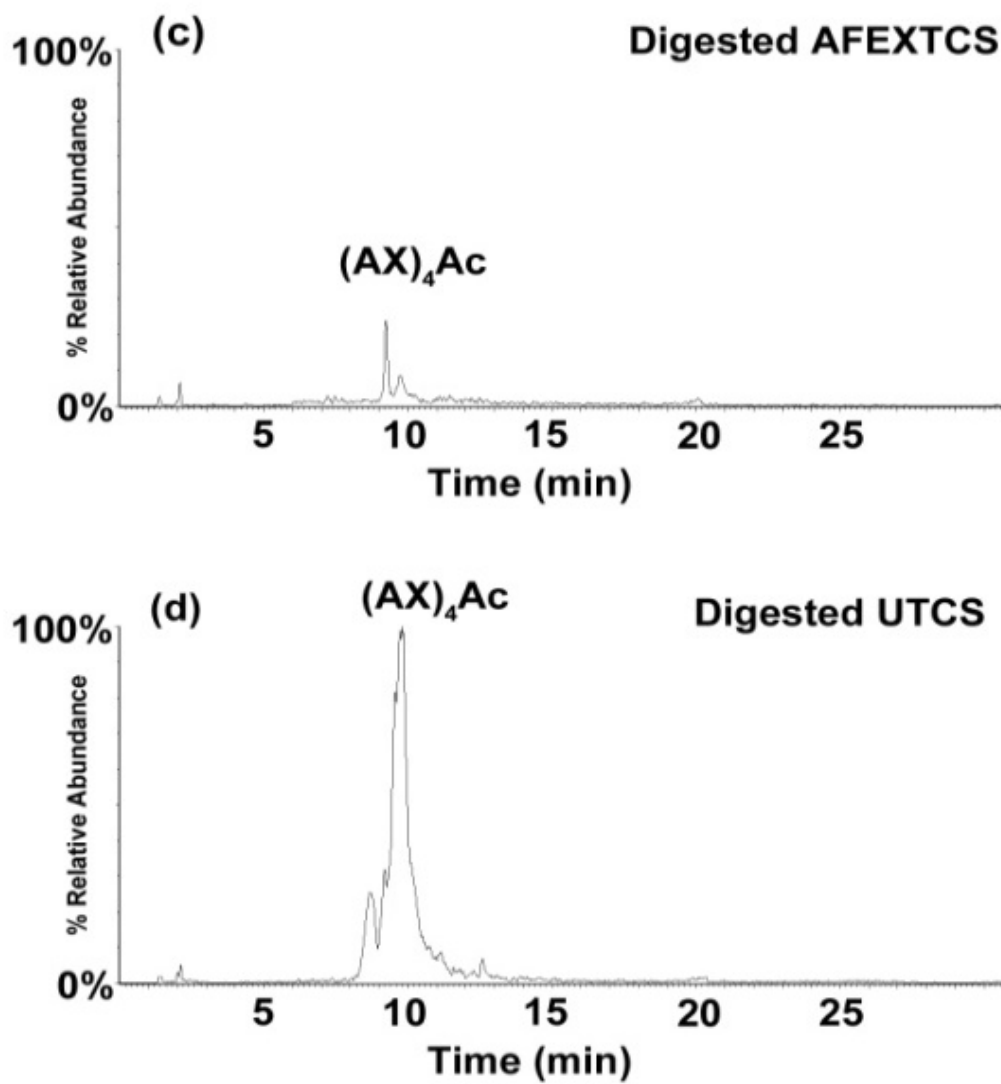
4.5 Conclusions

Identification of oligosaccharides released from pretreated biomass serves as an important step toward optimization of enzymatic digestion processes and reduction of the cost of biofuel production. The methodology here presents a simple, quick and powerful approach for enrichment, separation and identification of soluble neutral oligosaccharides. This approach was applied for profiling of glucans and arabinoxylans from the AFEXTCS hydrolysates, but the strategy can be applied for characterization of large neutral oligosaccharides from other sources as well. The Prevail Carbohydrate ES phase offers great utility for separation of large oligosaccharides including isomers, and the use of multiplexed CID provides acquisition of CID spectra for all separated oligosaccharides in a single LC/MS analysis, producing information-rich spectra that lead to rapid structure annotations. Although not reported here, this technique has potential use in combination with oligosaccharide derivatization for fast and reliable characterization of complex oligosaccharides, and efforts are already underway to profile arabinoxylans released from cellulosic biomass by AFEX and related processes.

REFERENCES

REFERENCES

(1)   Duchesne, L. C.; Larson, D. W. *Bioscience* **1989**, *39*, 238-241.

(2)   Pauly, M.; Keegstra, K. *Plant Journal* **2008**, *54*, 559-568.

(3)   Sinha, V. R.; Kumria, R. *International Journal of Pharmaceutics* **2001**, *224*, 19-38.

(4)   Ishii, T. *Plant Science* **1997**, *127*, 111-127.

(5)   Scheller, H. V.; Ulvskov, P. In *Annual Review of Plant Biology, Vol 61*; Merchant, S., Briggs, W. R., Ort, D., Eds.; Annual Reviews: Palo Alto, 2010; Vol. 61, p 263-289.

(6)   Sousa, L. D.; Chundawat, S. P. S.; Balan, V.; Dale, B. E. *Current Opinion in Biotechnology* **2009**, *20*, 339-347.

(7)   Harvey, D. J. *Mass Spectrometry Reviews* **1999**, *18*, 349-450.

(8)   Reinhold, V. N.; Reinhold, B. B.; Costello, C. E. *Analytical Chemistry* **1995**, *67*, 1772-1784.

(9)   Maslen, S. L.; Goubet, F.; Adam, A.; Dupree, P.; Stephens, E. *Carbohydrate Research* **2007**, *342*, 724-735.

(10)   Van Dongen, F. E. M.; Van Eylen, D.; Kabel, M. A. *Carbohydrate Polymers* **2011**, *86*, 722-731.

(11)   Fernandez, L. E. M.; Obel, N.; Scheller, H. V.; Roepstorff, P. *Carbohydrate Research* **2004**, *339*, 655-664.

(12)   Fernandez, L. E. M.; Obel, N.; Scheller, H. V.; Roepstorff, P. *Journal of Mass Spectrometry* **2003**, *38*, 427-437.

(13)   Harrison, S.; Fraser, K.; Lane, G.; Hughes, D.; Villas-Boas, S.; Rasmussen, S. *Analytical and Bioanalytical Chemistry* **2011**, *401*, 2955-2963.

(14)   Munisamy, S. M.; Chambliss, C. K.; Becker, C. *Journal of the American Society for Mass Spectrometry* **2012**, *23*, 1250-1259.

(15)   Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K. *Rapid Communications in Mass Spectrometry* **2006**, *20*, 1989-1994.

(16)   Gu, L. P.; Jones, A. D.; Last, R. L. *Plant Journal* **2010**, *61*, 579-590.

(17)   Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Jones, A. D.; Last, R.

L. *Plant Journal* **2010**, *62*, 391-403.

(18)   Leijdekkers, A. G. M.; Sanders, M. G.; Schols, H. A.; Gruppen, H. *Journal of Chromatography A* **2011**, *1218*, 9227-9235.

(19)   Churms, S. C. *Journal of Chromatography A* **1996**, *720*, 75-91.

(20)   Karlsson, G.; Swerup, E.; Sandberg, H. *Journal of Chromatographic Science* **2008**, *46*, 68-73.

(21)   Guignard, C.; Jouve, L.; Bogeat-Triboulot, M. B.; Dreyer, E.; Hausman, J. F.; Hoffmann, L. *Journal of Chromatography A* **2005**, *1085*, 137-142.

(22)   Cataldi, T. R. I.; Campa, C.; De Benedetto, G. E. *Fresenius Journal of Analytical Chemistry* **2000**, *368*, 739-758.

(23)   Lee, Y. C. *Journal of Chromatography A* **1996**, *720*, 137-149.

(24)   van der Hoeven, R. A. M.; Hofte, A. J. P.; Tjaden, U. R.; van der Greef, J.; Torto, N.; Gorton, L.; Marko-Varga, G.; Bruggink, C. *Rapid Communications in Mass Spectrometry* **1998**, *12*, 69-74.

(25)   Hennion, M. C. *Journal of Chromatography A* **2000**, *885*, 73-95.

(26)   West, C.; Elfakir, C.; Lafosse, M. *Journal of Chromatography A* **2010**, *1217*, 3201-3216.

(27)   Antonio, C.; Larson, T.; Gilday, A.; Graham, I.; Bergstroem, E.; Thomas-Oates, J. *Journal of Chromatography A* **2007**, *1172*, 170-178.

(28)   Westphal, Y.; Schols, H. A.; Voragen, A. G. J.; Gruppen, H. *Journal of Chromatography A* **2010**, *1217*, 689-695.

(29)   Zhang, J. H.; Xie, Y. M.; Hedrick, J. L.; Lebrilla, C. B. *Analytical Biochemistry* **2004**, *334*, 20-35.

(30)   Strum, J. S.; Aldredge, D.; Barile, D.; Lebrilla, C. B. *Analytical Biochemistry* **2012**, *424*, 87-96.

(31)   Slimestad, R.; Vagen, I. M. *Journal of Chromatography A* **2006**, *1118*, 281-284.

(32)   Wan, E. C. H.; Yu, J. Z. *Environmental Science & Technology* **2007**, *41*, 2459-2466.

(33)   Agblevor, F. A.; Murden, A.; Hames, B. R. *Biotechnology Letters* **2004**, *26*, 1207-1210.

(34)   Kalay, H.; Ambrosini, M.; van Berkel, P. H. C.; Parren, P.; van Kooyk, Y.; Vallejo, J. J. G. *Analytical Biochemistry* **2012**, *423*, 153-162.

(35)   Vinjamoori, D. V.; Byrum, J. R.; Hayes, T.; Das, P. K. *Journal of Animal Science* **2004**, *82*, 319-328.

(36)   Chundawat, S. P. S.; Vismeh, R.; Sharma, L. N.; Humpula, J. F.; Sousa, L. D.; Chambliss, C. K.; Jones, A. D.; Balan, V.; Dale, B. E. *Bioresource Technology* **2010**, *101*, 8429-8438.

(37)   Megherbi, M.; Herbreteau, B.; Faure, R.; Salvador, A. *Journal of Agricultural and Food Chemistry* **2009**, *57*, 2105-2111.

(38)   Domon, B.; Costello, C. E. *Glycoconjugate Journal* **1988**, *5*, 397-409.

(39)   Garozzo, D.; Giuffrida, M.; Impallomeni, G.; Ballistreri, A.; Montaudo, G. *Analytical Chemistry* **1990**, *62*, 279-286.

(40)   Spengler, B.; Dolce, J. W.; Cotter, R. J. *Analytical Chemistry* **1990**, *62*, 1731-1737.

(41)   Hofmeister, G. E.; Zhou, Z.; Leary, J. A. *Journal of the American Chemical Society* **1991**, *113*, 5964-5970.

(42)   Pasanen, S.; Janis, J.; Vainiotalo, P. *International Journal of Mass Spectrometry* **2007**, *263*, 22-29.

(43)   Zhou, Z. R.; Ogden, S.; Leary, J. A. *Journal of Organic Chemistry* **1990**, *55*, 5444-5446.

(44)   Quemener, B.; Ordaz-Ortiz, J. J.; Saulnier, L. *Carbohydrate Research* **2006**, *341*, 1834-1847.

(45)   Kabel, M. A.; Carvalheiro, F.; Garrote, G.; Avgerinos, E.; Koukios, E.; Parajo, J. C.; Girio, F. M.; Schols, H. A.; Voragen, A. G. J. *Carbohydrate Polymers* **2002**, *50*, 47-56.

(46)   Teleman, A.; Nordstrom, M.; Tenkanen, M.; Jacobs, A.; Dahlman, O. *Carbohydrate Research* **2003**, *338*, 525-534.

(47)   Izydorczyk, M. S.; Biliaderis, C. G. *Carbohydrate Polymers* **1995**, *28*, 33-48.

(48)   Hoffman, M.; Jia, Z. H.; Pena, M. J.; Cash, M.; Harper, A.; Blackburn, A. R.; Darvill, A.; York, W. S. *Carbohydrate Research* **2005**, *340*, 1826-1840.

(49)   Wang, J.; Yuan, X. P.; Sun, B. G.; Cao, Y. P.; Tian, Y.; Wang, C. T. *Food Chemistry* **2009**, *115*, 1529-1541.

(50)   Lemoine, J.; Fournet, B.; Despeyroux, D.; Jennings, K. R.; Rosenberg, R.; Dehoffmann, E. *Journal of the American Society for Mass Spectrometry* **1993**, *4*, 197-203.

(51)   Lemoine, J.; Strecker, G.; Leroy, Y.; Fournet, B.; Ricart, G. *Carbohydrate Research* **1991**,

221, 209-217.

(52)   Selig, M. J.; Adney, W. S.; Himmel, M. E.; Decker, S. R. *Cellulose* **2009**, *16*, 711-722.

(53)   Kabel, M. A.; Schols, H. A.; Voragen, A. G. J. *Carbohydrate Polymers* **2002**, *50*, 191-200.

(54)   Ishii, T. *Phytochemistry* **1991**, *30*, 2317-2320.

(55)   Humpula, J. F.; Chundawat, S. P. S.; Vismeh, R.; Jones, A. D.; Balan, V.; Dale, B. E. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* **2011**, *879*, 1018-1022.

Chapter Five: Concluding Remarks

Climate change and depletion of fossil fuels mandates a need to implement an energy transition from non-renewable to renewable energy resources that will last for centuries to come. Lignocellulosic biomass as a resource for production of renewable energy in form of liquid transportation biofuel represents an alternative to food sources with great opportunities and economic challenges. Efficient utilization of lignocellulosic biomass will be essential before this resource becomes economically viable for large scale production.

Thermochemical pretreatment of lignocellulosic biomass is necessary to ensure adequate yields of monosaccharides needed for liquid biofuel production, but also represents the most costly stage in a lignocellulosic biorefinery plant. Biomass pretreatment aims to alter or remove structural and compositional impediments to hydrolysis in order to improve rates of enzyme hydrolysis and increase yields of fermentable sugars. Understanding the chemistry of pretreatment and quantitative determination of products of chemical reactions that occur during pretreatment is required to support development of effective models that can guide rational design of pretreatment process conditions, and is crucial for process optimization. Most current knowledge regarding products formed during biomass processing has relied on GC/MS targeted analyses of a limited number of lignin and sugar derived components. Although these efforts generated a substantial body of knowledge about pretreatment processing but this information has not yet yielded the science base for successful predictions regarding how best to optimize the pretreatment processes.

In light of our limited understanding of biomass treatment fundamentals, multifaceted analytical approaches are required to analyze treated lignocellulosic biomass extracts owing to the wide range of molecular masses and physical properties of cell wall degradation products. Several important categories of these products that should be monitored qualitatively and

176

quantitatively in the process include: small phenolics and carboxylic acids (molecular weight 50-300 Da), volatile heterocyclic compounds (molecular weight 50-200 Da), fermentable monosaccharides (molecular weight 100-200 Da), and oligosaccharides and oligolignols (molecular weight 300 and above). Characterization and quantification of these products is a demanding task that requires fast and comprehensive analysis methods suitable for large-scale analyses generated during process optimization.

Developments in analytical technologies have moved toward establishing high-throughput, cost-effective and comprehensive methodologies that can identify and quantify numerous diverse analytes all at once, with minimal sample preparation. Replacing GC/MS methodologies that require long and labor-intensive sample preparation and derivatization with short LC/MS methods provides faster and more cost-effective analyses of biorefinery process streams. To move application of these technologies forward and push the limits of the scope of identified compounds formed during biomass processing, the research described in this dissertation has presented development of fast mass spectrometry-based analytical strategies to detect, characterize and quantify biomass degradation products in including oligosaccharides released during the process.

To this end, Chapter Two of this dissertation focused on profiling of nitrogenous compounds present in AFEXTCS extracts but absent in extracts of untreated corn stover. It was revealed that due to ammonolysis reactions, in addition to feruloyl and coumaroyl amides, diferuloyl amides, nitrogenous heterocyclics, and several other abundant nitrogenous compounds were present only in the AFEXTCS extract, of which the majority are suggested to be phenolic amides based on accurate mass measurements and RMD values determined using TOF mass spectrometry. The developed LC/MS and GC/MS strategies were implemented to quantify

177

important degradation products in extracts of corn stover which had been treated with varying severities of AFEX. Quantification of all detected nitrogenous compounds accounted for about 45-50% of ammonia loss that occurs during AFEX. Continuing efforts in this area are encouraged, since the remaining losses of ammonia have yet to be accounted for. Methods described in Chapter Two are appropriate for fast profiling of key reaction products in extracts of treated biomass that guides process optimization. Although the methodology explained in this chapter offered prospects for improved recognition and classification of chemical classes present in biomass extracts, the combination of accurate mass measurement and RMD reported in this dissertation has further applications extending beyond biomass research, and are appropriate for classification of metabolites in metabolomics and identification of other natural product-derived materials.

Chapter Three explained the chemical diversity of diferulate cross-linkers in plant cell walls and methodologies for comprehensive determination of the isomeric forms based on MS/MS product ions mass spectra combined with liquid chromatography retention times. Unlike conventional GC/MS methods for identification of diferulate isomeric forms, this approach avoids derivatization and can distinguish each diferulate isomer based on specific distinguishing fragment ions formed upon CID of diferulate $[M+H]^+$ ions. The LC/MS method was applied to demonstrate the diversity of diferulates in corn stover cell walls upon NaOH and AFEX treatments. This diferulate analysis method offers the potential to reveal treatment efficiency for a range of biomass resources, and also for guiding the breeding of grasses for improved digestibility and resistance to pests. It is hoped that the fragmentation chemistry demonstrated in this work will guide future efforts to characterize higher ferulate oligomers and establish the levels and roles of oligoferulates in cell wall structures of wild and engineered plant materials.

Profiling of oligosaccharides released during pretreatment yields information essential for choosing glycosyl hydrolases necessary for cost-effective conversion of cellulosic biomass to desired biofuel/biochemical end-products. In Chapter Four, LC/MS strategies were presented for characterization of oligosaccharides in the AFEXTCS extract.. LC/MS methods were developed to separate glucans with varying DP values from 2 to 22 and arabinoxylans. Physical separation using an amine-based chromatography column combined with mass spectrometry detection using a non-selective multiplexed-CID technique performed in a TOF mass analyzer revealed structural information of identified arabinoxylans and glucans in the AFEXTCS extract. Profiling of soluble arabinoxylans also demonstrated that arabinoxylan acetylation was reduced more than 85% during AFEX treatment.

The methodology explained in Chapter Four enables profiling of oligosaccharides using a consistent protocol for enrichment, which is often essential for oligosaccharide recovery, and characterization yielding accurate measurements of molecular and fragment masses in a collision energy resolved fashion. Distinguishing isomeric substances based on the energy-dependence of fragment ion abundances is an area of research still in its infancy, and is enabled by improvements in instrumentation and computational methods. These improvements in analytical throughput allow faster oligosaccharide profiling for characterization of oligosaccharides important in bioenergy production but also in production of foods, biomedical materials, and other consumer products that are based upon oligosaccharides.