Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2015

Advancements in mass spectrometry for biological samples: Protein chemical cross-linking and metabolite analysis of plant tissues

Adam T. Klein
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Analytical Chemistry Commons

Recommended Citation

Advancements in mass spectrometry for biological samples:

Protein chemical cross-linking and metabolite analysis of plant tissues

by

Adam Klein

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Program of Study Committee:
Young-Jin Lee, Major Professor
R. Sam Houk
Emily Smith
Edward Yu
Gustavo MacIntosh

Iowa State University

Ames, Iowa

2015

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Young Jin Lee, for the opportunity to work on challenging and exciting projects, as well as his encouragement to challenge myself in the process. I would also like to thank my committee members, Dr. Sam Houk, Dr. Emily Smith, Dr. Gustavo MacIntosh, and Dr. Ed Yu, for their guidance and support throughout the course of this research.

I would also like to thank everyone who has guided me down the road of chemistry, especially Dr. Cathy Haustein, who provided me with my first opportunity to do scientific research.

A special thanks is deserved for all of my family members. I couldn't have done it without their unwavering support, patience, and love.

ABSTRACT

This thesis presents work on advancements and applications of methodology for the analysis of biological samples using mass spectrometry. Included in this work are improvements to chemical cross-linking mass spectrometry (CXMS) for the study of protein structures and mass spectrometry imaging and quantitative analysis to study plant metabolites. Applications include using matrix-assisted laser desorption/ionization-mass spectrometry imaging (MALDI-MSI) to further explore metabolic heterogeneity in plant tissues and chemical interactions at the interface between plants and pests. Additional work was focused on developing liquid chromatography-mass spectrometry (LC-MS) methods to investigate metabolites associated with plant-pest interactions.

The first chapter includes an introduction into CXMS and MS-based metabolite analysis and the sixth and final chapter includes a brief summary of the work and future directions based on the work presented here.

The second chapter discusses the limitations of chemical cross-linking mass spectrometry, and steps taken to overcome these. This includes the incorporation of hydrogen-deuterium exchange mass spectrometry (HDX/MS) to identify the presence of structural distortion caused by the chemical cross-linking reaction. The use of extracted ion chromatograms (XIC) to increase cross-linking detection efficiency is also discussed.

The third chapter focuses on the application of MALDI-MSI to investigate seedling maize leaf tissues. Metabolites related to the photosynthetic process are compared at different locations along the leaves of two inbred genotypes. The localization and relative abundances of phosphatidylglycerols (PG) and sulfoquinovosyl diacylglycerols (SQDG) are described in detail.

The fourth chapter illustrates how imprinting MSI can be used to study the chemical interface between plants and pests. The interaction between soybean aphids and soybean leaves was studied using MALDI-MSI on imprints of infested leaves. These experiments revealed the spatial distribution of numerous metabolites, including compounds involved in the plant-pest interaction.

The fifth chapter discusses the development of a methodology to determine quaternary ammonium compounds (QACs) in plant samples. Hydrophilic interaction liquid chromatography (HILIC) mass spectrometry was used to quantitatively and qualitatively analyze common bean seed exudates. This work was focused on QACs that bacteria are known to uptake for osmoprotection and nutritional benefits.

CHAPTER 1

INTRODUCTION

This dissertation presents work in two distinct fields of mass spectrometry: chemical cross-linking for protein structure and small molecule analysis of plant tissues. Therefore, this introduction presents background information on both techniques separately.

Chemical Cross-Linking Mass Spectrometry

Protein structure

Proteins play an essential role in many biological functions including the transport of molecules within and between cells[1], replication of DNA[2], and catalysis of metabolic reactions[3]. The ability to carry out these functions depends largely on the structural conformation of the protein. Protein structure is often organized into three levels: primary structure, defined by the sequence of the linear amino acid chain; local secondary structures, such as alpha helices and beta sheets; and the 3-dimensional tertiary structure formed through non-covalent interactions such as hydrogen bonding, Van der Waals forces, hydrophobic packing, and ionic interactions. Investigating the 3-dimensional conformation of proteins can provide an insight into how they function.

Protein conformation studies are largely dominated by x-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy[4], which can provide structural

information with high resolution (1-2 Å)[5-6]. Over 95% of all protein structures listed in the Protein Data Bank were determined by one of these two techniques. Although x-ray crystallography and NMR spectroscopy dominate the field of protein structural analysis, both have inherent limitations: x-ray crystallography requires high-quality protein crystals, while NMR spectroscopy requires relatively large amounts of sample[7]. Crystallized proteins or proteins in high abundance are not native conditions and sample preparation can be prohibitively difficult using these methods. Moreover, a protein's structure in a crystal may not be reflective of its structure *in vivo*. The study of proteins under native conditions and at native concentrations could provide insight into the functional structure of a protein. As a result of these limitations in traditional analysis techniques, complementary mass spectrometric techniques have grown in popularity. Oxidative labeling[8], hydrogen exchange[9], and chemical cross-linking[10] mass spectrometry have all provided structural information for proteins under native conditions, albeit at much lower resolutions than available from crystallography or NMR spectroscopy. The advantages and disadvantages of these three techniques have been discussed in reviews by Konermann[11] and Sinz[12].

Cross-linking

Chemical cross-linking utilizes reagents consisting of two reactive groups separated by a spacer chain to form intramolecular connections in proteins. By analyzing the positions of the cross-links, it is possible to determine the maximum through-space distance between two amino acid residues. Cross-linking reagents are

available with different reactive moieties, spacer arm lengths, solubility, and possible enrichment properties, which allows cross-linking experiments tailored to many different experimental conditions and sample types. The reactive groups can be specific to amino acids, such as N-hydroxysuccinimide (NHS) esters that react with primary amines, or maleimide moieties that react with sulfhydryl groups, or non-specific reactions, such as aryl azide or diazirine photoreactive groups. Spacer arm lengths ranging from 0 Å, as with EDC, to 95 Å, as in $SM(PEG)_{24}$ are commercially available and provide a wide range of distance constraints. Water-soluble and membrane-permeable analogs are also commercially available and are suitable for cytosolic or membrane-bound proteins, respectively.

Mass spectrometry

Cross-linked proteins can be analyzed by mass spectrometry using approaches similar to those established for traditional protein mass spectrometry, which can be performed using a top-down or bottom-up approach. In top-down analysis, intact proteins are directly fragmented and analyzed inside the mass spectrometer. Bottom-up analysis is more common and involves proteolysis prior to introducing the sample into the mass spectrometer. The resulting peptides are analyzed and their parent masses (MS) can be used as a mass fingerprint to identify proteins. Additionally, fragmentation (MS/MS) can be performed to obtain more detailed sequence information. Analysis of peptide fragments improves the confidence in identification by providing the peptides' amino acid sequences. Protein identification for the bottom-up

approach uses protein databases or predicted sequences, against which the fragment spectra are searched. A review comparing top-down and bottom-up approaches and their advantages and disadvantages is provided by Bogdanov and Smith[13].

Typical protein analysis by mass spectrometry uses either matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI) as the ionization technique. Both MALDI and ESI are soft ionization techniques and have minimal fragmentation of analytes during ionization, which makes them suitable for analysis of proteins/peptides. In MALDI, a light-absorbing matrix is applied to the sample and cocrystallizes with the analytes. The crystalline sample is then irradiated by a laser. The absorbing matrix is heated and desorbed, in turn leading to desorption of the sample material. Peptides are ionized either in the sample or in the gas phase[14] and are introduced into the mass spectrometer for analysis. ESI is performed by applying a voltage to the solvent containing the sample and the resulting aerosol is directed toward the mass spectrometer. The aerosol droplets from the spray are charged at the surface, and as the droplets get smaller through coulomb fission and desolvation, the charge is transferred to the analytes. ESI is advantageous over MALDI for peptide and protein analysis in that they may take on multiple charges, thus reducing the mass-to-charge ratio and extending the detection range of the mass spectrometer.

Protein identification by mass spectrometry is greatly enhanced by using tandem MS to fragment peptides and obtain more detailed sequence information. Peptide MS/MS is capable of producing predictable fragmentation patterns, and several fragmentation techniques are available. Electron capture dissociation (ECD) and

electron transfer dissociation (ETD), both of which reduce the charge state of multiply charged cations by introducing anions, tend to produce fragments along the peptide backbone between the α-carbon and amine (c and z-type ions)[15]. Collision induced dissociation (CID) is also a popular fragmentation method for peptide analysis, and works by accelerating and colliding analyte ions with an inert gas, typically N_2, He, or Ar. Peptides are mostly fragmented at the amide bonds in the peptide backbone, producing b and y-type ions[16-17]. CID works well for peptides with a low charge state and <20 amino acid residues, while ECD and ETD are superior for intact proteins and proteins with labile post-translational modifications[18].

CXMS detection

Identification of cross-linked peptides is done in a similar fashion to peptide identification; however, finding the low-abundance cross-linked peptides can be challenging when they are present in a sample containing a plethora of non-cross-linked peptides. Much work has been done to overcome the issue of low-abundance cross-links, including using the intrinsic properties of cross-linked proteins to enrich cross-linked peptides in a sample. Theoretically, peptides cross-linked together should be the summed mass and charge state of the non-cross-linked peptides. Because of this, strong cation exchange[19] and peptide size exclusion chromatography[20] have been used to separate cross-linked peptides in sample mixtures. Incorporating a third functional group into the cross-linking reagent can also be used to overcome sample complexity. Affinity tags, such as azide-labeled cross-linkers[21] and biotin-labeled

cross-linkers[22] have been applied to enrich cross-linked peptides. Detection of cross-linked peptides has also benefitted from using fluorogenic cross-linkers [23], isotope-labeled cross-linkers[24] and thiol-cleavable cross-linkers[25] in order to distinguish cross-linked peptides from non-cross-linked peptides.

Even with these sample enrichment strategies, data analysis to identify the cross-linked peptides can still be challenging. Manual interpretation of data would be time consuming and would present the additional challenge of identifying two peptides in one spectrum. Instead of seeing b and y fragments of one peptide, the spectra are complicated by having fragments from two peptides. Because of these challenges, software programs have been developed to improve the detection of cross-linked peptides. These programs such as Popitam[26], xQuest[27], and X!Link[28] have automated data analysis and improved cross-linking detection efficiency. A discussion of the advantages and limitations of these programs is available[29].

Cross-linking to study protein structure can be viewed as a post-translational modification (PTM), and natural PTMS have been shown to induce structural changes in proteins. For example, phosphorylation of protein kinases[30] or oxidation of immunoglobulin $\gamma1$[31] leads to a change in protein conformation. Even the presence of cations can induce proteins to change conformations, such as calmodulin in the presence of Ca^{2+} [32]. Since cross-linking experiments modify at least two sites on a protein, changing the protein conformation upon cross-linking is a possibility. Previous studies have overcome this by limiting cross-linking to one cross-link per protein, under the rationale that a single cross-link should have a minimal effect on the native

structure and any change in conformation would occur after the cross-linking reaction[10]. Therefore, the structure actually probed is that of the unmodified conformation.

In this work, we discuss advancements to overcome the two major shortcomings of chemical cross-linking: low detection efficiency and the possibility of structural distortion. We focus our efforts on improving cross-linking detection by increasing the number of cross-links attached to each protein, while ensuring that the native protein conformation is not disturbed.

Plant Metabolite Analysis by Mass Spectrometry

Metabolomics

Metabolomics is the study of small molecule metabolites in biological systems. Understanding the identity, location, and quantity of these small molecules inside the cells, tissues, or entire biological system can provide valuable information in understanding biological processes. Metabolites are classified as either primary or secondary metabolites. Primary metabolites are associated with growth, development, and reproduction, whereas secondary metabolites have roles in stress response and defense against other species. Additional roles for secondary metabolites include antioxidants, signaling compounds, and inhibitors to competitor species.

Thousands of metabolites from a range of chemical classes have been indentified in plants. For instance, terpene hydrocarbons, basic alkaloids, and phenolic isoflavones

have all shown to act as deterrents to pests. Small organic acids, quaternary ammonium compounds, glycosidic compounds, and phospholipids are also prevalent in plants.

Plant metabolites are also highly dependent on factors such as the life stage of the plant, tissue type, and environmental conditions. Metabolites in seeds can differ greatly to those found in adult plants. Seeds are abundant in water-soluble nitrogenous metabolites that can easily be released during imbibition[33], while mature leaves have a protective waxy layer between the plant and the environment[34]. Roots contain metabolites associated with establishing a healthy rhizosphere[35-36], while the leaves can produce defense compounds specific to herbivores[37]. Even within plant tissues, metabolite distribution can vary. Flavonoids that serve as UV protectants are most abundant on the adaxial surface of the leaf[38]. Metabolite changes due to environmental factors, such as drought or temperature stress have been a focus of numerous studies[39-42]. Investigating the highly dynamic world of plant metabolites can provide an insight into the metabolic pathways associated with plant growth and response to stress conditions.

Mass spectrometry

Mass spectrometry is an excellent tool for determining metabolites in plants. Its versatility allows for analysis of a wide variety of metabolites, and experimental parameters and protocols can be modified to suit many different types of analytes. Metabolite identification or localization can be performed by qualitative MS analysis and quantitative analysis can determine metabolite abundance. The ongoing

improvement of ion transfer and detection efficiency in instrumentation has allowed for metabolite analysis to be expanded to low-abundance compounds. The development of multiple sample preparation and introduction techniques and several types of ionization modes allow for analysis of a vast range of tissue types and compound classes.

Instrument sensitivity is extremely important when analyzing plant metabolites. Many compounds are only present in small quantities in plant tissues. Mass spectrometers have routinely achieved femtomole-level analyses[43], and even yoctomole levels of detection have been reported[44]. This makes mass spectrometry a popular choice for analyzing low abundance compounds in plant tissues. The high sensitivity has also made it possible to progress from bulk tissue analysis down to the analysis of individual cells[45].

The most fundamental requirement for mass spectrometry is the need to ionize the analyte . Compounds containing functional groups are usually able to be protonated or deprotonated. Even hydrocarbons, which can be challenging to ionize because they lack functional groups, can be ionized under certain conditions[46-47]. Most mass spectrometers are able to switch polarity between positive and negative ion detection modes, either during analysis or between samples, which makes it possible to cover a wide range of metabolites.

Mass spectrometers are routinely coupled with liquid chromatography (LC) for analysis of plant extracts. This simplifies analysis by separating analytes from biological matrices and other metabolites. LC systems are compatible with atmospheric ionization

sources like electrospray ionization, as the samples are introduced in solution. Traditional LC analysis on reverse phase columns, suitable for analysis of slightly polar to nonpolar molecules, also uses solvents that are typically compatible with mass spectrometry analysis. Recent advancements with hydrophilic interaction liquid chromatography has also made the analysis of highly polar compounds suitable for LC-MS analysis, as the solvents used for HILIC separation are also compatible with MS. ESI sources have been shown to be useful for ionizing polar analytes in both positive and negative mode. Additional ionization sources, employing phenomena such as photoionization and chemical ionization, can also be incorporated in order to ionize highly nonpolar compounds.

Quantitative analysis is commonly performed by LC-MS. Detection modes such as selected ion monitoring (SIM) and multiple reaction monitoring (MRM) have been used to obtain quantitative information for selected metabolites[48].

Dried samples may also be analyzed by mass spectrometry by using matrix-assisted laser desorption/ionizataion (MALDI). Ionization in MALDI is facilitated by the matrix. Different organic matrices have been shown to be beneficial for certain classes of metabolites, such as α-cyano-4-hydroxycinnamic acid for peptides and proteins[49], 2,5-dihydroxybenzoic acid for small molecules and carbohydrates[50-51], 9-aminoacridine for nucleotides[52], and 1,5-diaminonaphthalene for lipids and small molecules[53-54]. Nanoparticle matrices have also been shown to be beneficial for harder to ionize analytes like terpenoids[55]. New matrices continue to be investigated and can provide further optimization for specific experiments.

11

MALDI-mass spectrometry imaging (MSI) has been used widely to show the distribution of metabolites in tissues[56-61]. These experiments are performed by scanning a region of interest on a surface, collecting spectra at a series of scan positions. The intensity for an ion in a spectrum can then be correlated to the x-y position of that spectrum. From this type of data, intensity maps can be generated that show the distribution of metabolites within tissue.

MALDI-MSI has been shown to be valuable in detecting and imaging metabolites in various plant tissues, such as roots [62], stems [63], leaves [64], and flowers [62]. Advancements in sample preparation techniques have helped expand the versatility of MALDI-MSI. Not all samples are suitable for direct MALDI-MSI analysis due to sample thickness or uneven morphology. Cryosectioning is commonly used to create thin tissue sections for the study of internal metabolites. Recent developments have used imprinting to extract leaf metabolites to a substrate prior to analysis[65]. These advancements have allowed for the analysis of all types of plant tissues. Thoughtful selection of sample preparation methodology, matrix, and ion polarity makes MALDI-MSI a suitable technique for a wide variety of metabolites in plants.

In this work, we expand on plant metabolite analysis by mass spectrometry. MALDI-MSI was used for qualitative analysis of photosynthesis-related metabolites in maize seedling leaves. Secondary metabolites associated with the defense response to aphids in soybean leaves were also studied. Finally, quantitative LC-MS analysis was used to study quaternary ammonium compounds released during germination of common bean seeds.

Dissertation Organization

This dissertation is divided into six chapters. The first chapter, seen above, serves as an introduction to chemical cross-linking mass spectrometry, as well as the use of mass spectrometry-based methods to investigate plant metabolites. Chapter two describes the limitations of chemical cross-linking mass spectrometry, and offers solutions to improve the detection efficiency and minimize distortions caused by cross-linking. The third chapter presents the application of MALDI-MSI to maize leaves to further understand the distribution of photosynthesis-related metabolites. Chapter four describes the application of imprinting MSI to study the plant-pest chemical interface on soybean leaves. The fifth chapter details the methodology used to study small polar metabolites in seeds that are associated with parasite osmoprotection and nutrition. The sixth and final chapter is a summary of this work and gives insight into future work based on the developed methodologies.

CHAPTER 2

IMPROVING EFFICIENCY IN CHEMICAL CROSS-LINKING MASS SPECTROMETRY: MONITORING STRUCTURAL DISTORTION USING HYDROGEN-DEUTERIUM EXCHANGE AND CROSS-LINK DETECTION USING EXTRACTED ION CHROMATOGRAMS

Adam T. Klein and Young Jin Lee

Abstract

Chemical cross-linking mass spectrometry has become a valuable technique for studying structures of proteins and protein complexes. However, its utility has been scrutinized due to the possibility of structural distortion and the low abundance of cross-links. In this study we developed two methods that can dramatically increase the cross-link detection efficiency. In the first method, we propose to increase the average number of cross-links per protein while monitoring potential structural distortion by hydrogen-deuterium exchange mass spectrometry. In the case of cytochrome c, we demonstrated an average of two cross-links per protein retains virtually no structural distortion, in which the number of identified cross-links could be increased about twofold. In the second approach, we propose to perform chemical cross-linking reactions in two experimental conditions, a mild condition with almost no structural distortion but with few cross-links, and an aggressive condition with many more cross-links but also with a possibility that some of them might have come from structurally

distorted proteins. Cross-linked peptide sequences can be identified in MS/MS at the aggressive condition and their presence in the mild condition can be confirmed using extracted ion chromatograms (XIC).

Introduction

Three-dimensional protein structural analysis has been mostly performed by NMR or X-ray crystallography, which provides atomic level high-resolution structural information. However, these methods have some inherent limitations, such as the difficulty to study membrane proteins and macro-protein complexes, as well as proteins under native conditions[1]. Mass spectrometry has been applied to overcome some of these limitations and provide low-resolution structural information of proteins. Ion mobility-mass spectrometry directly accesses conformational space through collisional cross section measurements but the interpretation heavily relies on computational modeling[2].

Three other mass spectrometry techniques provide more detailed structural information: mass spectrometry combined with hydrogen-deuterium exchange (HDX), oxidative labeling, and chemical cross-linking. Hydrogen-deuterium exchange mass spectrometry (HDX/MS) has been the most successful with amino acid or peptide level structural information[3-6]. Specifically, with electron capture and electron transfer dissociation (ECD/ETD), it has become possible to directly obtain amino acid level structural information from intact proteins[7-9]. Mass spectrometric analysis of oxidized residues, after oxidative labeling, has the advantage of permanently modifying amino

acids, which then can be analyzed by a traditional proteomics approach by looking for oxidation as a post-translational modification. Oxidative labeling by synchrotron X-ray has the advantage of directly producing hydroxyl radicals from solvent radiolysis[10]; however, potential structural distortion during the minimum radiation time on a millisecond time scale is a major concern. Recently developed laser flash photolysis of hydrogen peroxide, called fast photochemical oxidation of proteins (FPOP), overcame this limitation by minimizing the reaction time to a microsecond time scale using a nano-second UV laser[11] and was applied to study protein folding dynamics[12] and epitope mapping[13].

Chemical cross-linking followed by mass spectrometric analysis (CXMS) is capable of directly providing distance information between amino acid residues[14-15], and is especially well suited for studying protein-protein interactions[16]. However, CXMS still faces challenges due to the possibility of protein structural distortion during cross-linking and the low abundance of cross-linked peptides. To avoid potential structural distortion, the average number of cross-links has been commonly limited to one cross-link per 10 kDa protein size [17]; i.e. even if there is structural distortion induced by chemical cross-linking, the second cross-link, which will probe such distortion, will be minimal. This strategy has been widely adopted; however, it has two critical limitations in that 1) such minimal cross-linking further limits the resulting cross-linked peptides that can be detected and 2) some proteins with two or more cross-links are still present in this condition (see Supplementary Figure1). Several approaches have been adopted to improve the detection of cross-linked peptides by

16

applying enrichment strategies such as affinity tags[18-19], strong cation exchange chromatography[20-22], or peptide size exclusion chromatography [23]. However, each enrichment strategy has its own limitation of losing some cross-links and it would be ideal to develop additional strategies independent of enrichment.

We have previously developed a shotgun approach to detect cross-links directly from MS/MS spectra, using a home-made program called X!Link[24]. This approach is much more efficient than comparing with a control in the precursor spectra, for the same reasons that shotgun proteomics is much more efficient than peptide finger printing[25]. We also developed a probability based scoring algorithm to improve confidence in finding these cross-links and minimize human efforts to confirm the cross-links[26]. Several other computational algorithms are now available that also allow for more efficient cross-link detection including xQuest[20], Popitam[27], and CrossWork[28]; yet, none of them increases the abundance of cross-linked peptides present in the given sample set.

One simple way to increase the abundance of cross-linked peptides is to induce more cross-linking reactions, i.e. more than one cross-link per protein, but to ensure there is no structural distortion in the given experimental condition. In this work, we pursue this possibility by adopting HDX/MS as a method to validate the absence of structural distortion while we induce more cross-links per protein. We are especially inspired by Back *et al.* who suggested the actual structural distortion in chemical cross-linking would be very minimal since cross-linking proteins to a resin surface is commonly adapted without losing enzymatic activity[29]. Our approach is especially

important as we directly probe structural distortion in CXMS, which has not been done as far as we are aware. To further improve the cross-link detection, we also developed a new method of comparing extracted ion chromatograms (XIC) of mild and aggressive cross-linking conditions. In this approach, we propose to find low abundance cross-links in a mild cross-linking condition without MS/MS spectra, by extrapolating XIC from an aggressive cross-linking condition where the MS/MS spectra are identified.

Experimental

Materials

Cytochrome c was purchased from Sigma (St. Louis, MO). Disuccinimidyl suberate (DSS) was purchased from Thermo Scientific Pierce (Rockford, IL). Trypsin endoproteinase was purchased from Promega (Madison, WI). All solvents and buffers were purchased from Sigma (St. Louis, MO).

Time dependent cross-linking for HDX

Three samples of cytochrome c (5 µM in 20 mM HEPES buffer) were cross-linked with DSS for the reaction times of 10 min, 30 min, and 60 min, respectively, at 25°C. DSS was initially dissolved in dimethyl formamide (DMF) at 25 mM concentration and added to the samples for a final concentration of 500 µM, which corresponds to 100:1 cross-linker to protein molar ratio. The cross-linking reaction was quenched by adding ammonium bicarbonate to a final concentration of 50 mM. All samples were then spin filtered (Amicon Ultra, molecular weight cutoff of 10 kDa; EMD Millipore, Billerica, MA)

with 10 mM ammonium acetate three times, to remove excess cross-linker, exchange the buffer, and concentrate the sample. The samples were reconstituted in 10 mM ammonium acetate to a final cytochrome c concentration of ~25 μM.

The HDX/MS experiment was performed using a three syringe system [9]. Syringe 1 contained the 25 μM cross-linked cytochrome c solution and had a flow of 4 μL/min. Syringe 2 contained a 10 mM ammonium acetate solution in D_2O and had a flow of 16 μl/min. Syringe 1 and 2 were mixed with a zero-volume tee that allowed the hydrogen-deuterium exchange reaction to occur for 5 seconds before being quenched by the solution from the third syringe. Syringe 3 contained a quenching solution, a mixture of acetonitrile, H_2O, D_2O, and formic acid at the ratio of 10:18:72:0.4 with pH 2.5, and was at a flow of 20 μL/min prior to being mixed with the mixture from Syringes 1 and 2, and introduced into the mass spectrometer. MS data was collected on an Agilent QTOF 6540 using a capillary voltage of 3500 V, and a fragmentor voltage of 175 V, skimmer voltage of 65V, and scanning m/z range of 500–3000.

Mild vs. aggressive cross-linking

Cytochrome c (5 μM) was cross-linked with DSS in 20 mM HEPES buffer at reaction temperatures of 25°C for the mild condition and 45°C for the aggressive condition. The DSS concentrations for cross-linking reaction were 100 μM for the 25°C reaction temperature, and 500 μM, 2500 μM, and 12.5 mM for the 45°C reaction temperature. These concentrations correspond to cross-linker to protein molar ratios of 20:1, 100:1, 500:1, and 2500:1, respectively. Reaction times were varied with

19

reaction temperature in an attempt to keep the amount of cross-linking consistent. The final reaction times used were 1 hour at 25°C and 30 minutes at 45°C. The cross-linking reaction was quenched by adding ammonium bicarbonate to a final concentration of 50 mM. Buffer was changed to 50 mM ammonium bicarbonate by spin filtering three times, removing over 99% of HEPES and excess DSS. Samples were digested overnight by adding acetonitrile to 10% and trypsin at a substrate-to-enzyme ratio of 20:1 (w/w). After the digestion, the samples were quenched by adding formic acid to 1% and kept at -80°C before LC-MS/MS analysis.

LC-MS/MS of cross-linked peptides

Digested samples were analyzed using a hybrid linear ion trap Orbitrap high-resolution mass spectrometer (LTQ-Orbitrap; Thermo, San Jose, CA) coupled with liquid chromatography (Paradigm MS4; Michrom Bioresources, Auburn, CA). A home-built capillary column (75 μm id x 15 cm) packed with reverse phase packing material (C18AQ, 5 μm, 100 Å; Michrom) was used with a 2 hour gradient at 300 nl/min. The gradient included a 5 minute loading step at 5% mobile phase B, followed by a linear gradient from 5% to 34% B for 70 minutes, a 40 minute gradient from 34% to 80% B, a five minute hold at 80% B, and a five minute re-equilibration at 5% B. Mobile phase A consisted of 2% acetonitrile and 0.1% formic acid and mobile phase B consisted of 60% acetonitrile and 0.1% formic acid. Spray voltage was optimized for each experiment.

Precursor scans were acquired with the Orbitrap high-resolution mass spectrometer, followed by four data-dependent MS/MS scans with the ion trap mass

spectrometer. Dynamic exclusion was used with a repeat count of 1, a repeat duration of 30 seconds, an exclusion list of 50, and an exclusion duration of 180 seconds. MS/MS of singly and doubly charged ions were excluded, determined from the pre-scan.

Data analysis

Deisotoping was done with Mascot Distiller (v2.4.2.0; Matrix Science, London, UK) with parameters specified previously [24]. Cross-links were identified using X!Link, which reports three E-values: one for the cross-linked peptide, and two for each individual peptide. Positive identification was based on the maximum E-values of 0.03 for the cross-linked peptide and 0.3 for each individual peptide. The X!Link program is freely available by contacting the corresponding author.

Results & Discussions

Probing structural distortions using HDX/MS

Here, we propose hydrogen-deuterium exchange mass spectrometry (HDX/MS) as a tool to investigate structural distortions induced by chemical cross-linking under the given cross-linking conditions. HDX/MS is a very attractive structural tool due its high sensitivity, ability to probe heterogeneous distributions, and accessibility to surface exposed amino acids. As a first step to test the plausibility and develop a methodology easily adaptable to complex cross-linked proteins, we studied the structural distortion of cytochrome c with the addition of 0%, 5%, 10%, and 20% acetonitrile using HDX/MS. Figure 1 shows the 15+ ion of cytochrome c directly

eletrosprayed after incubation with deuterium for 5 seconds through a three syringe pump system (see experimental section for the details). The most abundant peak positions are at m/z 823.52, 823.87, 823.94, and 824.34 for Figures 1A-1D. Considering the peak position without HDX is m/z 816.29 (not shown), they correspond to a hydrogen-to-deuterium exchange of 108.5, 113.7, 114.8, and 120.8, respectively. This suggests that even only 5% organic solvent begins unfolding the cytochrome c from its native state, although very slightly (five HDX). Another notable feature in these HDX/MS measurements is a wider isotope envelope as the protein becomes denatured, as seen in Figures 1C and 1D. This is due to a subpopulation of the sample being unfolded to a greater extent. This broadening could be a better barometer to indicate the structural distortion compared to the simple peak shift, because it gives information about heterogeneous structural distortion in a subpopulation.

For the convenience of the data analysis, an exponentially modified Gaussian function (EMG), h(x), is adapted to fit the isotope envelope in the spectra[30].

$$h(x) = \int_0^x G(y)H(x-y)dy$$

Where G(y) is a Gaussian function and H(x) is an exponential decay function given by,

$$H(x) = \frac{1}{\tau}\exp(-\frac{x}{\tau})$$

The amount of tailing in the spectra is mostly determined by the variable τ, which is affected by natural isotopic distribution and exchanged deuterium distribution. With all other parameters kept constant, the best fitting tau value for the organically denatured cytochrome c increases with the amount of denaturant; 14, 21, and 52, respectively, from 12 for the undenatured protein. EMG fitting is relatively insensitive to slight

structural changes (e.g., Figure 1A vs 1B), whereas very sensitive to subpopulation changes (e.g., Figure 1B vs. 1C). This is in contrast to the simple peak shift comparison; uptake of 5 additional deuteriums from Figure 1A to 1B, but only one more deuterium uptake from Figure 1B to 1C. According to our simulation, EMG fitting can clearly differentiate 10% change in FWHM (Supplementary Figure 2).

To investigate the utility of HDX/MS in probing structural distortion in cross-linking, cytochrome c is cross-linked with DSS for various durations to allow for different amounts of cross-links to be attached to the protein. DSS is a bifunctional cross-linker with an NHS-ester reactive group at the end of each arm. It reacts with the primary amines on the lysine side chain or an unmodified protein N-terminus and forms amide bonds. DSS has a spacer chain length of 11.4 Å and can cross-link two lysines with inter-alpha carbon (Cα-Cα) distance up to 24 Å apart. Considering uncertainty in X-ray or NMR measurement, the maximum inter-alpha carbon distance between the two cross-linkable lysines could be up to about 27 Å when compared with 3D protein structures available in a protein database. Rappsilber et al. also suggested the maximum distance of 27.4 Å, which is confirmed in their large-scale analysis of RNA polymerase II-TFIIF complex[21].

Figure 2 compares the three spectra of cytochrome c cross-linked with DSS at 25°C for 10, 30, and 60 minutes, and electrosprayed with on-line deuterium labeling using the three syringe pump system[9]. These spectra correspond to an average number of cross-linking of 2.1, 3.4, and 3.6, respectively. It should be noted that 12+ ion series of cytochrome c were most abundant after cross-linking, due to the modification

of some lysine residues, and used in this analysis, instead of 15+ ions in Figure 1. For the cytochrome c reacted for 10 minutes (Figure 2A), the non-cross-linked cytochrome c was observed, along with up to four cross-links per protein. Each cross-link series contain multiple peaks due to the combination of intra-protein cross-links and dead-end modifications (one end is attached and the other end is hydrolyzed), resulting in the protein mass increase of +138Da and +156Da, respectively. Hence, one cross-linked protein has two peaks corresponding to 1 cross-link and 1 dead-end, two cross-liked protein has three peaks corresponding to 2/0, 1/1, and 0/2 modifications of cross-link/dead-end, and three cross-linked protein has four peaks corresponding to 3/0, 2/1, 1/2, and 0/3 modifications of cross-link/dead-end. For longer cross-linking reaction times (Figures 2B and 2C), non-cross-linked proteins were not observed and the number of cross-links was up to five cross-links per protein or higher.

To get insight about the possible structural distortions induced by chemical cross-linking, the spectra in Figure 2 are fitted with the exponential modified Gaussian equation. In Figure 2A, the tau value of 36 could best fit the peak profile for non-cross-linked cytochrome c, corresponding to the natural broadening of cytochrome c without cross-linking. It should be noted that the normal tau value is larger than that of the 15+ ions in Figure 1A because the 12+ ion is broader in m/z space. For up to 3 cross-links, the same tau value of 36 could be used to fit the profiles, suggesting there is no significant structural distortion up to 3 cross-links attached to the protein. However, a higher tau value of 65 gives the best fitting for the species containing 4 cross-links. It was the same for Figures 2B and 2C; i.e., a tau value of 36 could fit the peak profile up to

3 cross-links, whereas the tau value of 65 was needed for 4 and 5 cross-links. In Figure 2, we did not fit all the tails for simplicity (such as sodium ion adducts). Even when we took all these into account, we still could not fit the spectrum with tau value of 36 (Supplementary Figure 3), indicating 4 or higher cross-links significantly affect the protein structure of cytochrome c.

To conclude, in the case of cytochrome c with the given experimental condition, HDX/MS revealed 4 or more cross-links would structurally distort the protein structures. One thing to note is there is no apparent structural distortion for up to three cross-links in cytochrome c (Figure 2A). Considering significant structural distortion is induced by the fourth cross-linking reaction, the proteins with four cross-links could still be used for cross-link analysis. Namely, the fourth cross-linking reaction might have been probing three cross-link attached proteins, which did not have significant structural distortion at the time. Therefore, the cross-linking condition for Figure 2A, with the average of two cross-links per protein and up to four cross-links attached, is still a reliable experimental condition for cross-link analysis with almost no structural distortion. However, the fourth cross-link might have induced significant changes in the protein structure afterwards, and five or more cross-link containing proteins are not reliable for structural analysis. It should be noted that less than 1.6% of five cross-linked proteins were present in Figure 2A, which is usually not expected to be a problem or could be safely removed from XIC comparison as will be discussed in the next section.

This result suggests that the commonly adopted guideline of one cross-link per 10 kDa protein size is unnecessarily tight, and up to two cross-links per protein might be acceptable for some proteins. This work also confirms cross-linking does distort the protein structures in aggressive reaction conditions, significantly enough that there is a noticeable difference in the number of exchangeable hydrogens after cross-linking. Our conclusion obtained for cytochrome c should be carefully generalized to other proteins, because each protein might be different in terms of their structural sensitivity to chemical cross-linking. However, considering cytochrome c is a relatively small protein with MW of only 12 kDa and many surface exposed lysines, we suspect it is probably one of the more structurally sensitive proteins to cross-linking. Other proteins, especially much bigger proteins with larger surface areas, are less likely to be affected by chemical cross-linking. Hence, we propose the conclusion obtained in cytochrome c, i.e. there is little or no structural distortion up to two cross-links per protein, is a relatively safe guideline in most CXMS experiments.

Use of extracted ion chromatograms to improve cross-link detection

By using the average of two cross-links per protein as suggested above instead of one, the cross-link detection efficiency can be improved about twofold. However, it is necessary to improve the efficiency even further, considering a significant amount of these cross-links are either dead-end modifications or structurally uninformative "close-by" cross-links (e.g., intra-peptide cross-links). To further improve cross-link detection efficiency, we propose a simple method based on extracted ion

26

chromatograms. The basic idea came from the fact that many of the low abundance cross-links may not be selected for MS/MS data acquisition, in spite of their presence in precursor mass spectra. In this approach, the cross-linking reaction would be performed in two different reaction conditions; a mild experimental condition and an aggressive reaction condition. The former (mild condition) would produce no structural distortion but have far less cross-links to detect and the latter (aggressive condition) would identify many more cross-links but have some cross-links from structurally distorted proteins. For those cross-links identified with MS/MS in the aggressive reaction condition but not in the mild condition, the extracted ion chromatograms (XICs) would be produced for both the data sets with very narrow accurate mass windows (5 ppm). Then, the cross-link is considered to be present in the mild condition if the XIC is clearly observed at the same retention time as the aggressive reaction condition. This approach would identify low abundance cross-links that are present in the precursor MS spectra, but their MS/MS were not obtained or obtained with very low quality.

This methodology is tested with cytochrome c as a model system and DSS as a cross-linker. We used a probability based cross-link identification program, X!Link, we have previously developed[26]. This algorithm uses three E-values to distinguish true cross-links from false positives; one for the entire cross-linked peptide, and the other two to confirm each individual peptide. E-values of 0.03 and 0.3 for the cross-link and each peptide, respectively, were suggested as effective cutoff values that would remove most random matching even in a search against large protein sequence databases. In

this proof of concept experiment, we used the mild experimental condition at 25°C for 60 min with a molar ratio of 20:1, a comparable condition for Figure 2A, and the aggressive reaction conditions was at 45°C for 30 min with molar ratios of 100:1, 500:1, and 2500:1, similar or more aggressive than the conditions for Figure 2B or 2C.

Table 1 summarizes the number of identified cross-links in cytochrome c at the various cross-linking reaction conditions. Fourteen unique cross-links were identified in the mild condition and 18, 24, and 33 unique cross-links were identified at the aggressive reaction conditions with molar ratios of 100:1, 500:1, and 2500:1, respectively. All the detected cross-links have an inter C_α distance of 23 Å or below between the two cross-linked lysines in the mild condition (maximum distance is 18.7 Å), but some of them are stretched to 23-27 Å in the aggressive condition. It should be noted that the inter C_α distance of 23-27 Å is possible but only when the two lysines are arranged to stretch out toward each other without significant steric hindrance. Details of identified cross-linked peptides are presented in Supplementary Tables 1A-D for each experimental condition.

Figure 3 demonstrates how we detect low abundance cross-links using XIC. In this example, XIC of m/z 881.4535 corresponding to the cross-link of ^{40}TGQAPGFSYTDAN**K**NK^{55}-^{73}**K**YIPGTK^{79} (3+ charge state) were compared at the mild and aggressive conditions. MS/MS spectra corresponding to TGQAPGFSYTDAN**K**NK-**K**YIPGTK were collected under the aggressive condition of 45°C and a molar ratio of 100:1, but no MS/MS spectra were collected at the mild condition. Although there is no MS/MS, the existence of this cross-link at the mild condition is evident from the exact

match of the XIC with the harsh condition. The inset precursor mass spectra are almost identical, in the mass values and isotope envelopes, further supporting our claim. This method is so sensitive that it may detect some very low abundance cross-links from structurally distorted proteins present in extremely low amounts, e.g., five cross-linked proteins in Figure 2A; however, this possibility could be also carefully considered and removed by quantitatively comparing XIC. Using the above XIC approach, a total of eight additional cross-links were initially considered to be present in the sample cross-linked at the mild condition. Table 2 quantitatively compares XIC peak areas of the eight cross-links at the mild and the aggressive conditions. Of the eight potential cross-links, three of them (K39-K100, K25-K39, and K39-K86) have extremely small peak areas compared to the most aggressive condition (less than 0.5%); these three are ignored because of the possibility they may have come from very low abundance structurally distorted proteins. Especially, their dramatic increase of ion abundance at 2500:1 compared to 500:1 (>20 times signal increase for 5 times more cross-linking molar ratio) indicates that they are sensitive to structural distortion. Additionally, the absolute ion abundances for the three cross-links are extremely low in the precursor mass spectra in the mild condition, less than 0.01% of the uncross-linked peptide used for the normalization purpose. Hence, the removal of these three cross-links can be justified. It is interesting to find all three cross-links contain K39, and the 2500:1 condition has many new K39 containing cross-links (Supplementary Table 1D) whereas there is no K39 containing cross-links in the mild condition (Supplementary Table 1A). K39 forms a beta bridge with T58 through hydrogen bonding (PDB:2B4Z)

and the denaturation of the protein seems to break the hydrogen bonding to make K39 available for cross-linking with other lysines.

The other five cross-links are present at a much higher abundance without abnormalities; i.e., gradual increase or even decrease of XIC peak area as cross-linker-to-protein molar ratio increases. Gradual increase of XIC is due to more cross-linking reactions and the decrease is due to competition with structurally distorted proteins that have become available. Table 3 summarizes the detailed information about the five additionally identified cross-links. Three of them had MS/MS spectra acquired at the mild condition but were rejected due to high E-values. For example, MS/MS spectra corresponding to ^{8}**K**IFVQK^{13}-^{100}**K**ATNE^{104} (m/z 487.9486 for 3+ charge state) were detected in both mild and aggressive reaction conditions. However, E-values in the mild condition (E=0.032, E-alpha=0.027, E-beta=0.60) are slightly higher than the cutoff values used (E=0.03, E-alpha=0.3, E-beta=0.3) and rejected; its high E-values are attributed to its low ion signals and corresponding low quality MS/MS spectra. In contrast, the cross-link is validated in the aggressive condition for its low E-values (0.0001, 0.0001, and 0.23, respectively). Although its MS/MS spectrum from the mild condition cannot validate the presence of **K**IFVQK-**K**ATNE, the XIC is exactly matching (Supplementary Figure 4) with the aggressive reaction condition where the MS/MS is validated, confirming its presence.

Of the five additional cross-links identified, four cross-links had Cα-Cα distances between the Lys pairs below 17 Å, which are well within cross-linkable distance (<27 Å). K79-K88 has a slightly tight distance of 23.4 Å but there is no steric hindrance in the

direct path. In addition, XIC peak area for K79-K88 increases only mildly as the cross-linker-to-protein molar ratio increases, suggesting this is not from structurally distorted proteins. The additional identification of five cross-links corresponds to a 36% increase from the previous 14 cross-links. It should be noted that a total of 32 cross-links were identified exclusively in the aggressive conditions but only five of them were confirmed to be present in the mild condition using the XIC method. The rest of the 24 cross-links are considered as either structurally prohibited or simply were not cross-linked in sufficient amount in the given mild condition.

Conclusions

In this paper, we developed two methods that can dramatically improve cross-link detection in CXMS without any special labeling agents or enrichment strategies, and demonstrated the methods using cytochrome c as an example. In the first method, we proposed to use a CXMS experimental condition with more than one cross-link per protein but still without structural distortion, confirmed by HDX/MS. In the case of cytochrome c, we demonstrated there is almost no structural distortion up to two cross-links per protein at 25°C. In the second method, we proposed to perform cross-linking reactions in mild and aggressive experimental conditions, then identify the cross-links with MS/MS at aggressive conditions and confirm its presence at the mild condition by comparing XIC. The second method is so sensitive that there is a risk it may detect very low abundance cross-links from structurally distorted proteins, which, however, can be distinguished from the quantitative evaluations of XIC peak areas at

31

various mild and aggressive reaction conditions. By combining the two methods, we expect cross-link detection efficiency can be dramatically increased making CXMS much more useful. These methods are independent of chemical reagents or enrichment protocols and could be utilized in combination with other protocols, e.g. SCX enrichment, if a further increase in cross-linking detection efficiency is desired.

Furthermore, this is the first direct and rigorous demonstration for the presence of structural distortion in aggressive cross-linking conditions and absence of structural distortion in a mild condition. Other methods such as circular dichroism or NMR could be also used to detect structural distortions, but these techniques may not be sensitive to probe minor structural changes occurring in a limited subpopulation. In contrast, HDX/MS could easily characterize low abundance structurally distorted proteins, because of the inherent advantage of mass spectrometry in probing heterogeneous molecular populations.
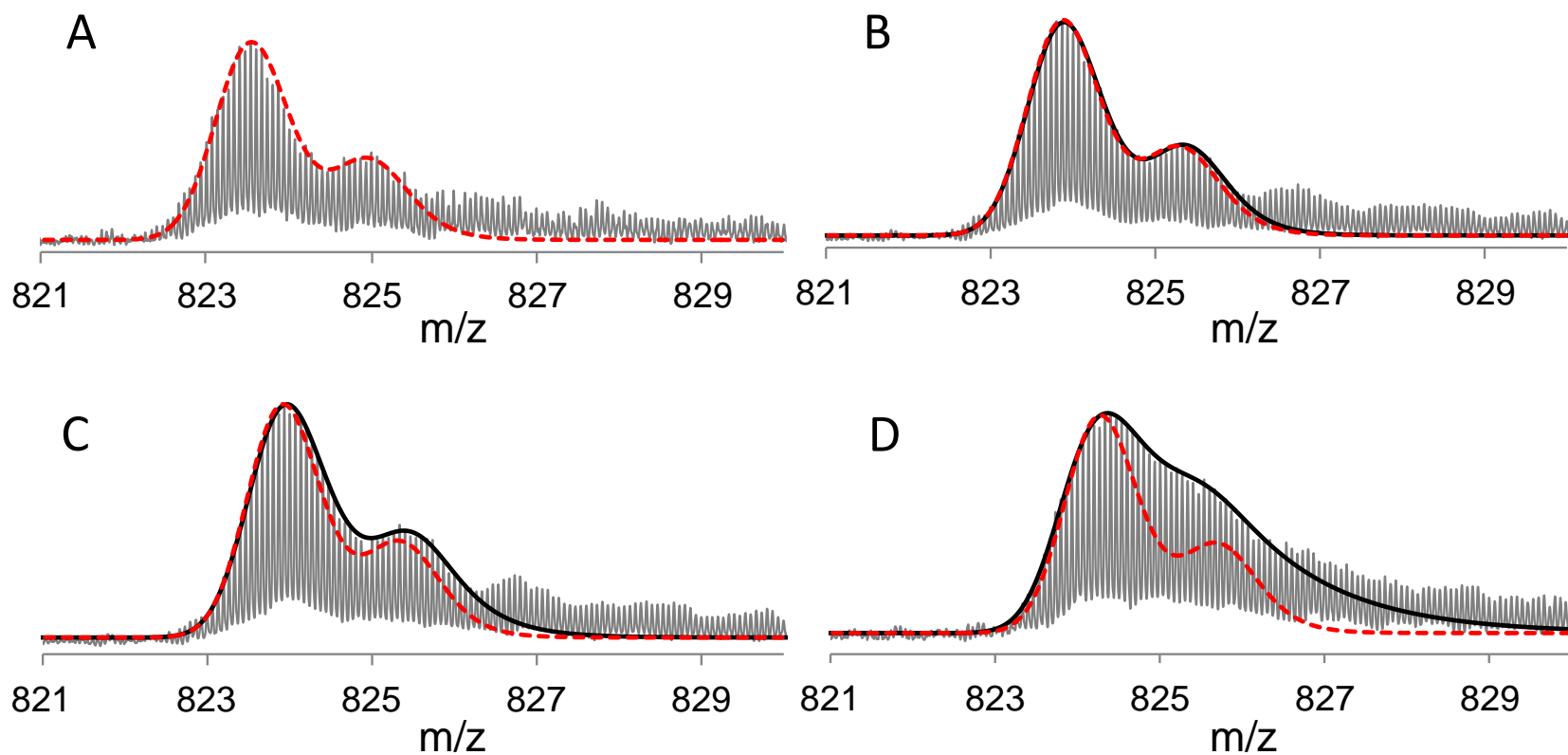
Figure 1. HDX/MS of denatured cytochrome c. HDX/MS spectra for the 15+ ion of cytochrome c compared with the best fitting EMG function. (A) Cytochrome c with no acetonitrile and (B-D) acetonitrile concentration at 5%, 10%, and 20%, respectively. The red dashed line (--) corresponds to the best fitting EMG in the control sample ($\tau = 12$), of which the peak position is shifted in B-D for comparison. The solid lines (—) in B-D correspond to the best fitting EMG with τ value of 14, 21, and 52, respectively. Note the second peak corresponds to one sodium addition, $[M + 14H + Na]^{15+}$.
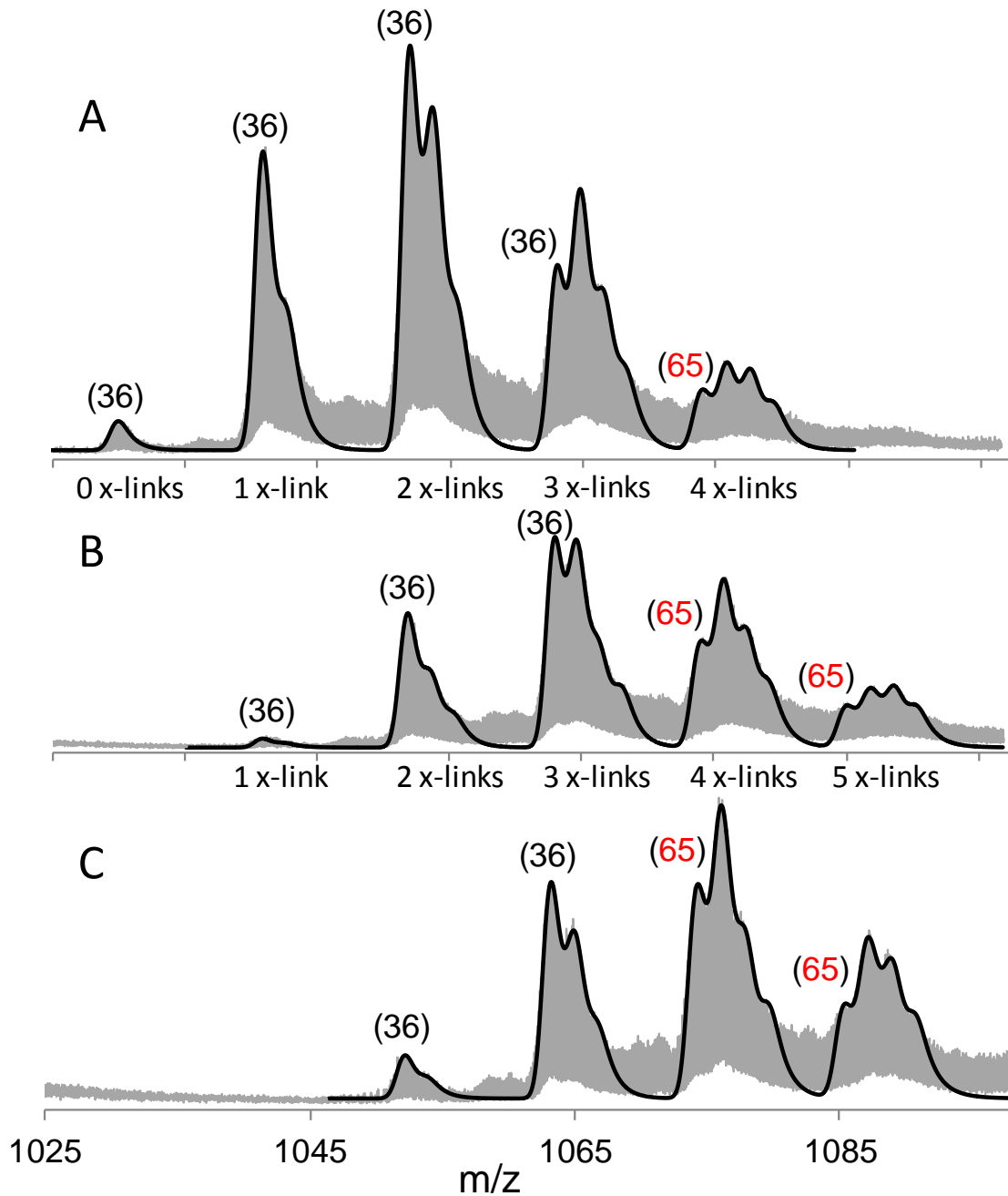
Figure 2. HDX/MS of cross-linked cytochrome c. The 12+ series of HD exchanged cytochrome c (gray line) after cross-linked with DSS (100:1 molar ratio) for (A) 10 minutes, (B) 30 minutes, and (C) 60 minutes compared with the best fitted EMG (solid line). Best fitting tau values are shown in parentheses.

Table 1. Number of identified cross-links. Summary for the number of unique cross-links identified by X!Link at various cross-linking reaction conditions.

Cα distance (Å)	25°C 20:1	45°C 100:1	45°C 500:1	45°C 2500:1	Total identified cross-links	Total possible cross-links
<18	12	16	19	17	28	66
18-23	2	1	3	7	8	36
23-27	0	1	2	9	10	38
> 27	0	0	0	0	0	13
Total	14	18	24	33	46	153

- Cα distances are between two cross-linked lysines and calculated from the x-ray crystallography structure PDB:2B4Z in the PDB database (*www.rcsb.org*).

- Total possible cross-links are the number of all the theoretical cross-links between two lysines based on the inter Cα distance ignoring any steric hindrance.
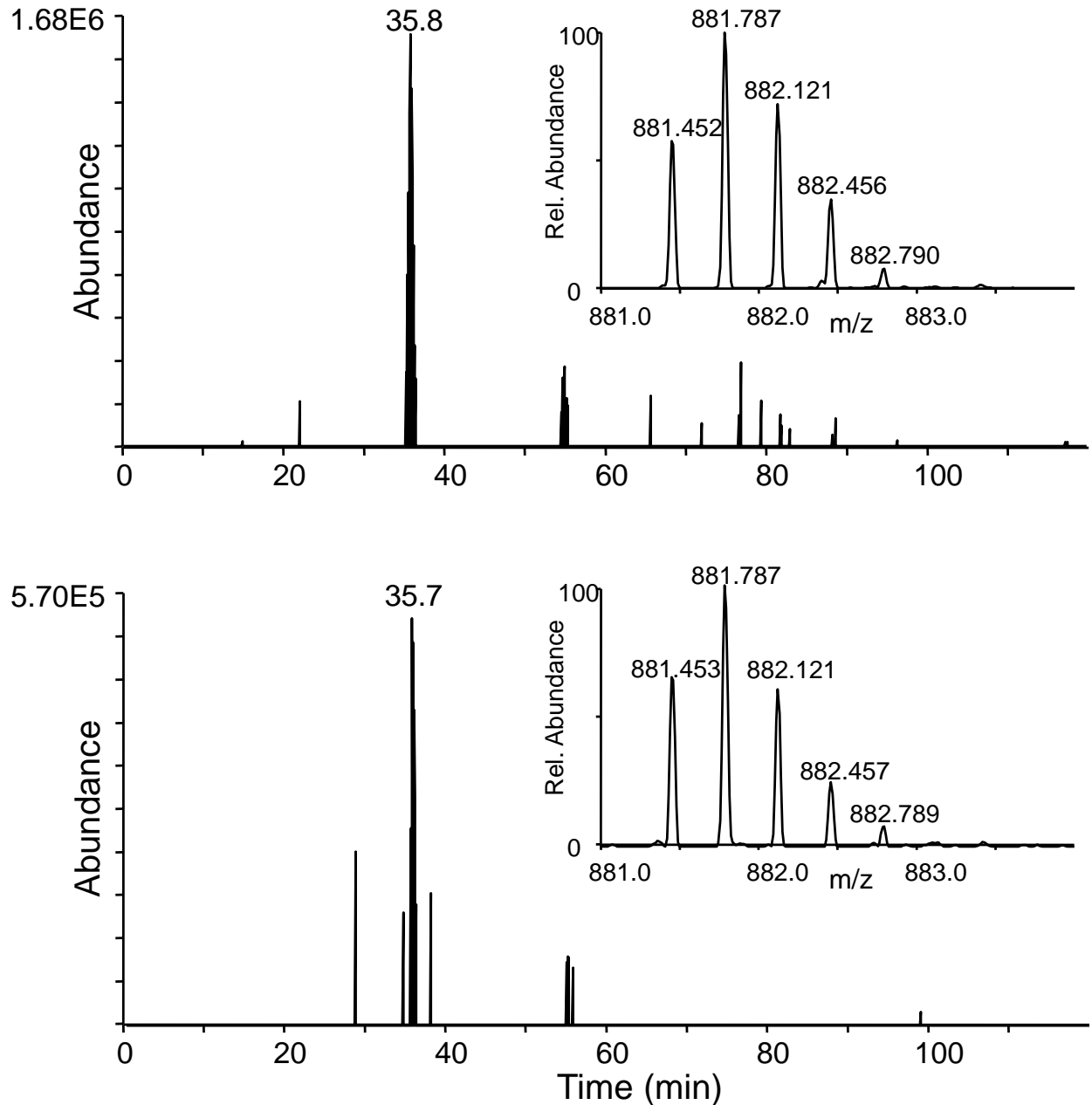
Figure 3. Extracted ion chromatogram of m/z 881.4535. XIC of m/z 881.4535 ([TGQAPGFSYTDAN**K**NK-**K**YIPGTK + 3H+]³⁺) for the data set at (A) 45°C/100:1 and (B) 25°C/20:1 reaction conditions. Mass tolerance of 0.004 Da was used for XIC generation. Zoomed-in inset precursor mass spectra were obtained averaging over the retention time of 35-36min.

Table 2. Additional cross-links identified at aggressive conditions. XIC peak areas of the cross-links with MS/MS validated at aggressive conditions.

Cross-link	XIC area at mild condition, %	XIC area at aggressive condition[a], %			Mild-to-Aggressive XIC Ratio, %		
		100:1	500:1	2500:1	100:1	500:1	2500:1
K22-K100	0.0401	0.225	0.267	0.0733	17.8	15.0	54.7
K8-K100	0.240	1.14	1.35	0.252	21.1	17.9	95.6
K53-K73	0.0211	0.145	0.228	0.0423	14.6	9.3	49.9
K88-K99	0.351	2.15	4.47	10.1	16.4	7.9	3.5
K39-K100[b]	**3.73×10^{-3}**	0.0242	0.0630	1.29	15.4	5.92	**0.290**
K25-K39[b]	**9.39×10^{-4}**	0.0125	0.0106	0.228	7.49	8.89	**0.412**
K79-K88	0.315	0.710	0.997	0.141	44.4	31.6	224
K39-K86[b]	**4.80×10^{-4}**	3.12×10^{-3}	4.70×10^{-3}	0.172	15.4	10.2	**0.279**

a. XIC peak areas were normalized by non-cross-linked peptide (TGPNLHGLFGR) to correct for run-to-run experimental variations.

b. Considered to have come from structurally distorted proteins and rejected.

Table 3. Additional cross-links at mild conditions. Additionally identified cross-links in the mild cross-linking condition by using XIC method.

| Cross-links	C$_\alpha$ dist. (Å)	Rt at mild condition	Rt at aggressive condition	E-values in the mild condition	Identified E-values
^{14}CAQCHTVE**K**GGK^{25}-^{100}**K**ATNE^{104}	12.2	60.8	61.3	--	0.0001, 0.001, 0.0003
^{8}**K**IFVQK^{13}-^{100}**K**ATNE^{104}	13.7	25.6	25.4	**0.032**, 0.027, **0.60**	0.0001, 0.0001, 0.23
^{40}TGQAPGFSYTDAN**K**NK^{55}-^{73}**K**YIPGTK^{79}	15.3	35.7	35.8	--	0.0002, 0.21, 6.2E-07
^{88}**K**GER^{92}-^{93}EDLIAYL**K**K^{100}	16.4	53.4	50.3	**0.043**, 0.12, 0.02	0.016, 0.18, 0.0009
^{74}YIPGT**K**MIFAGIK^{86}-^{88}**K**GER^{92}	23.4	50.0	47.8	**0.457**, 0.10, **0.41**	0.005, 0.19, 0.02

- Bold K in cross-linked peptide sequences represent cross-linked lysine residue.
- Rt represents LC retention time in min. Rt at aggressive condition is for the data where MS/MS is identified with sufficiently low E-values.
- Evalues are in the order of E-value, E-alpha, and E-beta. Bold E-values in the mild condition are higher than the cutoff values (0.03 for E-value and 0.3 for E-alpha and E-beta) and rejected.
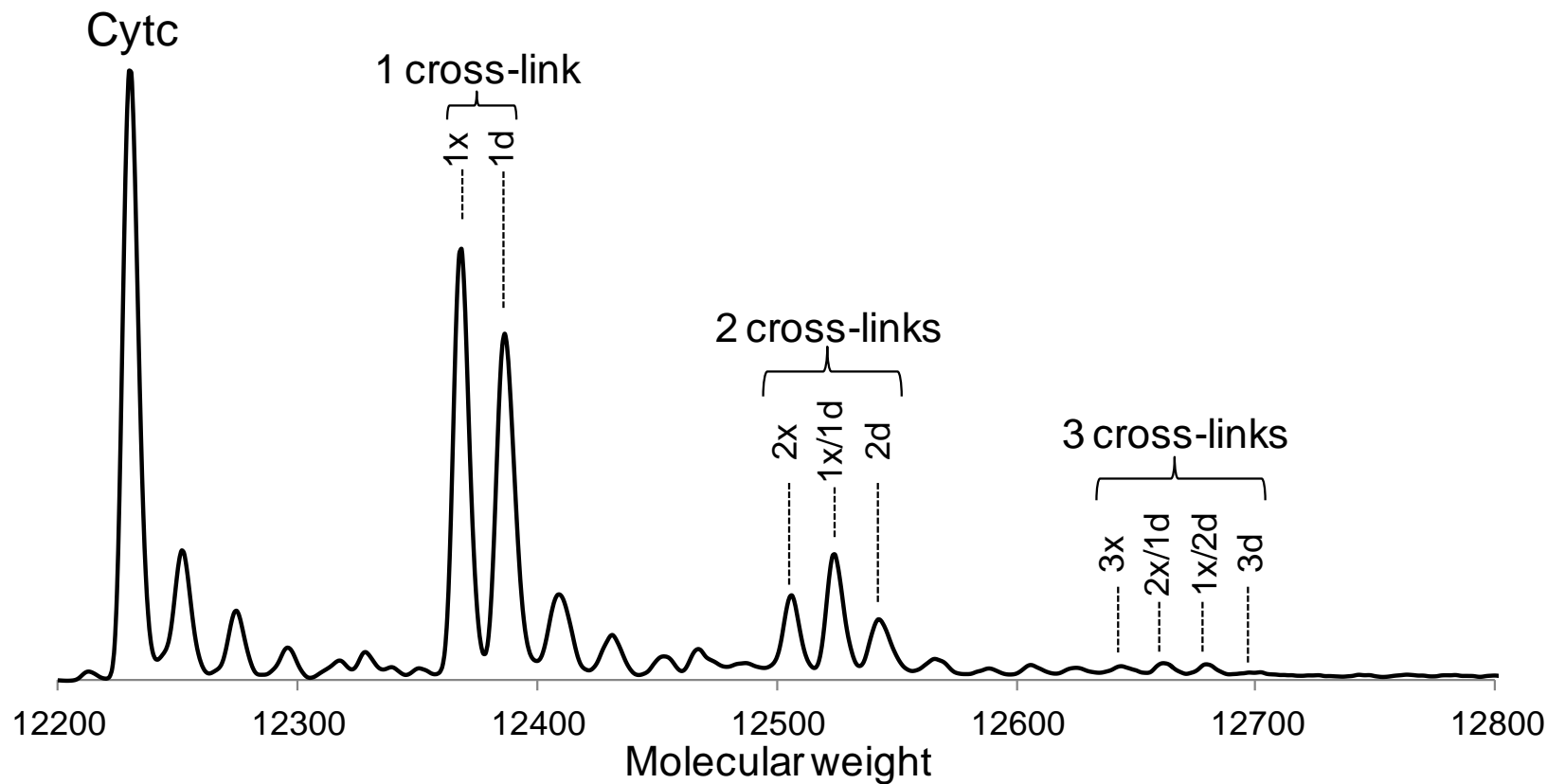
Figure S1. Deconvoluted cross-linked cytochrome c. Deconvoluted cytochrome c ESI mass spectrum obtained in the cross-linking condition that gives an average of one cross-link per protein. Two and three cross-links per protein are present at 35% and 6% of single cross-linked proteins, respectively. 'x' and 'd' represent intra-protein cross-links and dead-end modifications corresponding to +138Da and +156Da of mass shift, respectively.
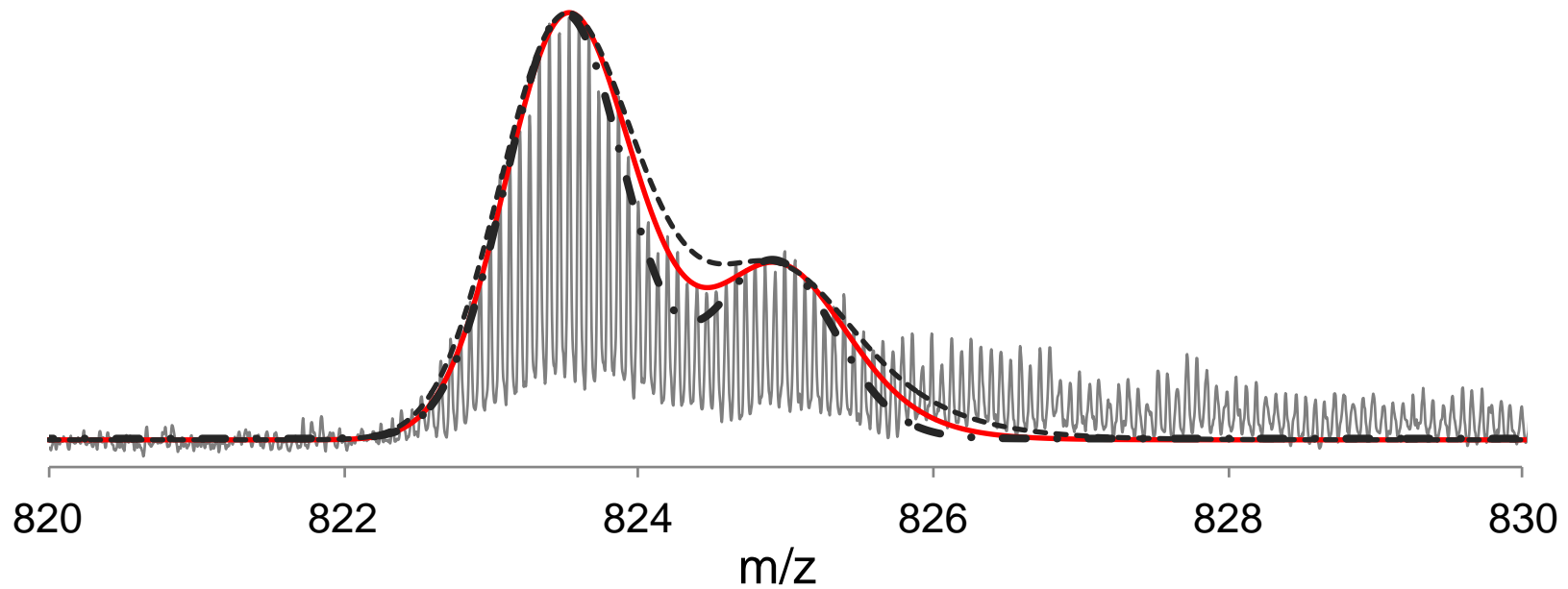
Figure S2. EMG fitting. The mass spectrum for the 15+ ion of cytochrome c (gray line) compared with best fitting EMG profile ($\tau=12$, red line) and EMG profiles with 10% narrower ($\tau=4.5$, dot-line) and 10% broader FWHM ($\tau=19.5$, dashed line).
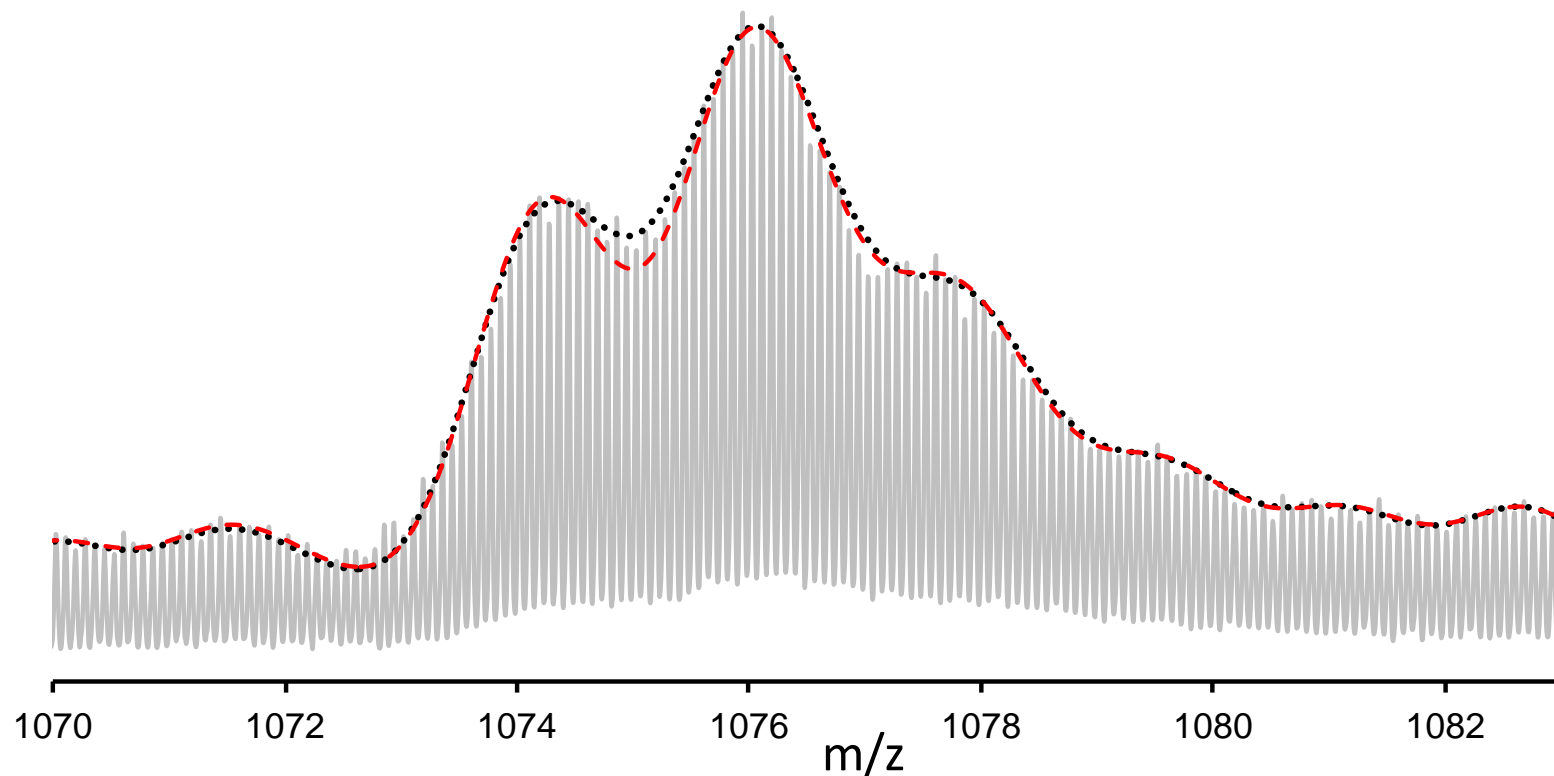
Figure S3. EMG fitting of all peaks. Zoomed-in mass spectrum of the 4 cross-link series (gray solid line) in Figure 2C compared with EMG profiles using tau values of 36 (red dashed line) and 65 (black dotted line). These EMG profiles were generated using a tau value of 36 to take into account all of the contaminant peaks leading up to the peaks of interest.
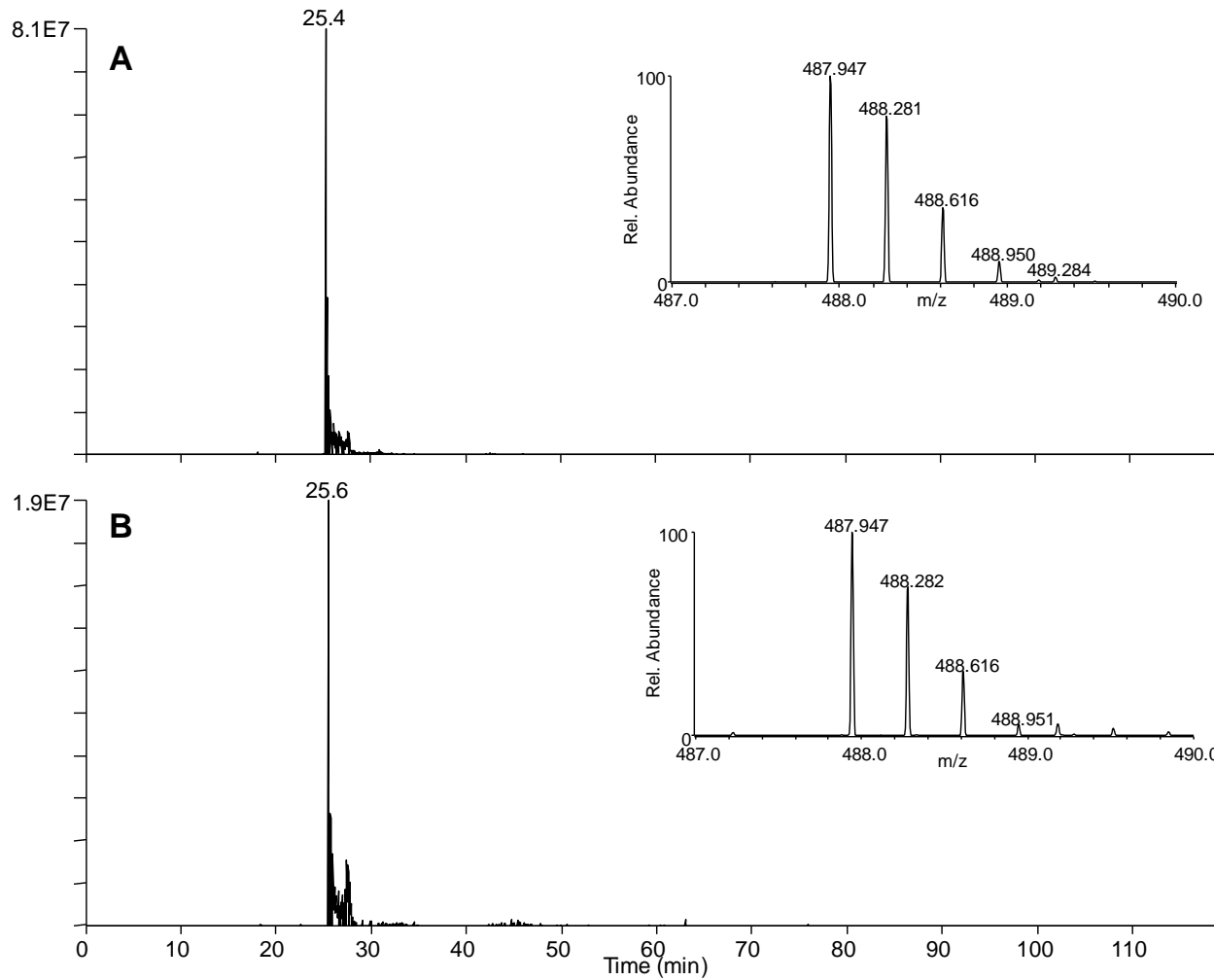
Figure S4. Extracted ion chromatogram of m/z 487.9486. Comparison of the XIC of m/z 487.9486, which corresponds to the 3+ ion of [8]KIFVQK[13]-[100]KATNE[104] cross-linked peptides at reaction conditions of A) 45°C/500:1 and B) 25°C/20:1. Zoomed-in inset precursor mass spectra were obtained averaging over the retention time of 25-26 min.

Supplementary Table 1A. Identified cross-linked peptide ions in the mild cross-linking condition (25°C/20:1 molar ratio).

K1	K2	Peptide1	Peptide2	Z	E-values
5	8	KIFVQK	AcGDVEKGK	3	7.36E-10, 0.0007, 5.19E-06
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	4	2.05E-11, 0.0009, 2.59E-06
5	13	AcGDVEKGK	KIFVQKCAQCHTVEK	5	0.007, 0.185, 0.011
5	13	AcGDVEKGKK	IFVQKCAQCHTVEK	4	0.002, 0.034, 0.020
5/7	13	AcGDVEKGKK	IFVQKCAQCHTVEK	5	0.010, 0.144, 0.014
5	87	KK	AcGDVEKGKK	3	0.0002, 0.053, 0.002
5	87	KKGER	AcGDVEKGKK	3	1.97E-07, 5.81E-05, 0.001
5	87/88	AcGDVEKGK	KKGEREDLIAYLK	4	0.0004, 0.002, 0.002
5	87/88	KKGER	AcGDVEKGK	3	1.01E-08, 0.011, 2.90E-05
5	87/88	KKGER	AcGDVEKGKK	4	1.26E-05, 6.75E-05, 0.002
5	88	KGER	AcGDVEKGK	3	1.32E-07, 0.005, 4.57E-07
5/7	88	AcGDVEKGKK	KGEREDLIAYLK	4	0.002, 0.045, 0.002
7	27	GKK	HKTGPNLHGLFGR	3	7.65E-11, 0.091, 2.43E-11
7	27	GKK	HKTGPNLHGLFGR	4	2.84E-10, 0.009, 3.45E-08
7	88	KGER	AcGDVEKGKK	3	4.48E-06, 0.021, 1.63E-07
8	87	KK	KIFVQK	2	6.60E-06, 0.265, 0.002

Supplementary Table 1A continued

8	88	KKGER	KIFVQK	3	4.84E-08, 0.019, 1.77E-05
25	100	GGKHK	KATNE	3	0.023, 0.259, 0.002
53	79	YIPGTKMIFAGIK	TGQAPGFSYTDANKNK	4	0.018, 0.227, 0.027
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	3	4.01E-09, 0.151, 3.28E-11
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	4	5.77E-05, 0.165, 5.83E-07
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	3	3.33E-08, 0.005, 1.15E-11
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	4	6.38E-09, 0.0003, 1.45E-07
73	86	KYIPGTK	MIFAGIKK	3	5.01E-10, 8.96E-05, 4.08E-06
73	86	KYIPGTK	MIFAGIKK	4	4.50E-06, 0.002, 1.97E-05
73	88	KYIPGTK	KGEREDLIAYLK	4	0.0007, 0.0007, 0.001

- If the same cross-linked ions were detected in multiple MS/MS spectra, only the best E-value is listed.
- K1 and K2 denote cross-linked lysine residue positions in cytochrome c protein sequence.
- E-values corresponding to cross-link, peptide1, and peptide2 are listed for the cross-links that passed maximum filtering values of 0.03, 0.3, and 0.3, respectively.

Supplementary Table 1B. Identified cross-linked peptide ions in the aggressive cross-linking condition of 45°C/100:1 molar ratio.

K1	K2	Peptide1	Peptide2	Z	E-values
5	8	KIFVQK	AcGDVEKGK	3	2.32E-10, 4.55E-06, 9.21E-06
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	4	1.25E-06, 6.19E-05, 0.008
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	5	6.18E-05, 2.73E-05, 0.018
5	13	AcGDVEKGK	KIFVQKCAQCHTVEK	4	0.0007, 0.002, 0.010
5/7	13	AcGDVEKGKK	IFVQKCAQCHTVEK	5	1.27E-07, 0.005, 5.30E-07
5	87	KKGER	AcGDVEKGKK	3	1.09E-06, 0.0003, 6.70E-07
5/7	87	KK	AcGDVEKGKK	3	3.06E-07, 0.073, 3.96E-05
5/7	87	KKGER	AcGDVEKGKK	4	4.44E-05, 0.003, 0.0007
5	88	KGER	AcGDVEKGK	3	2.91E-06, 0.003, 2.33E-05
5	88	KKGER	AcGDVEKGK	3	1.80E-08, 0.001, 0.0002
5/7	88	KGER	AcGDVEKGKK	3	1.07E-06, 0.0002, 0.0004
7	13	AcGDVEKGKK	IFVQKCAQCHTVEK	4	3.38E-06, 5.24E-05, 0.004
7	27	GKK	HKTGPNLHGLFGR	3	2.76E-12, 0.003, 8.78E-11
7	27	GKK	HKTGPNLHGLFGR	4	1.41E-09, 0.0006, 6.68E-08
7	73	GKK	KYIPGTK	3	0.001, 0.043, 0.014
8	87	KK	KIFVQK	2	0.0004, 0.24, 0.014
8	88	KIFVQK	KKGEREDLIAYLK	4	2.63E-06, 2.01E-05, 3.46E-06
8	88	KKGER	KIFVQK	3	1.60E-09, 0.010, 6.51E-07
8	100	KATNE	KIFVQK	3	0.010, 0.013, 0.066

Supplementary Table 1B continued

13	86	MIFAGIKK	IFVQKCAQCHTVEK	4	0.001, 0.0002, 0.250
13	86	MIFAGIKK	IFVQKCAQCHTVEK	5	2.13E-05, 3.57E-05, 0.077
22	100	KATNE	CAQCHTVEKGGK	3	0.004, 0.0009, 00.014
22	100	KATNE	CAQCHTVEKGGK	4	0.0001, 0.0004, 0.013
53	73	KYIPGTK	TGQAPGFSYTDANKNK	3	0.0002, 0.206, 6.25E-07
53	79	YIPGTKMIFAGIK	TGQAPGFSYTDANKNK	3	4.75E-08, 0.019, 1.59E-07
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	3	4.33E-12, 0.0001, 3.02E-09
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	4	7.62E-08, 0.037, 2.94E-08
72	79	YIPGTKMIFAGIK	GITWGEETLMEYLENPKK	4	3.58E-07, 8.41E-05, 1.56E-05
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	3	2.05E-08, 0.017, 7.78E-11
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	4	2.07E-11, 0.0003, 1.81E-11
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	5	0.0003, 0.068, 5.10E-06
73	86	KYIPGTK	MIFAGIKK	3	3.65E-12, 0.0008, 7.37E-06
73	86	KYIPGTK	MIFAGIKK	4	7.66E-05, 0.010, 9.52E-06
73	88	KGER	KYIPGTK	3	2.34E-10, 0.001, 2.71E-07
73	88	KYIPGTK	KGEREDLIAYLK	4	0.006, 1.37E-05, 0.018

Supplementary Table 1C. Identified cross-link ions in the aggressive cross-linking condition of 45°C/500:1 molar ratio.

K1	K2	Peptide1	Peptide2	Z	E-values
5	8	AcGDVEKGK	KIFVQKCAQCHTVEK	3	0.0005, 0.019, 0.0008
5	8	KIFVQK	AcGDVEKGK	3	1.82E-14, 3.24E-06, 5.42E-06
5	8/13	AcGDVEKGK	KIFVQKCAQCHTVEK	4	0.0004, 0.002, 0.007
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	3	0.0002, 0.290, 0.002
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	4	1.54E-06, 0.002, 0.0005
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	5	0.0006, 0.0001, 0.039
5	13	AcGDVEKGKK	IFVQKCAQCHTVEK	4	1.29E-05, 0.016, 0.0005
5	87	AcGDVEKGK	KKGEREDLIAYLK	4	0.003, 0.02, 5.25E-05
5	87/88	KKGER	AcGDVEKGK	3	1.94E-08, 0.0003, 1.97E-05
5	87/88	KKGER	AcGDVEKGKK	4	3.23E-06, 7.37E-06, 1.37E-05
5/7	87	KKGER	AcGDVEKGKK	3	5.50E-08, 2.23E-05, 1.49E-06
5	88	KGER	AcGDVEKGK	3	7.65E-07, 0.004, 7.37E-07
5/7	88	AcGDVEKGKK	KGEREDLIAYLK	4	0.013, 0.012, 6.29E-05
5/7	88	KGER	AcGDVEKGKK	3	3.67E-06, 0.006, 1.10E-05
7	13	AcGDVEKGKK	IFVQKCAQCHTVEK	5	1.39E-05, 0.029, 4.93E-06
7	27	GKK	HKTGPNLHGLFGR	3	1.44E-15, 0.005, 3.39E-11
7	27	GKK	HKTGPNLHGLFGR	4	1.01E-08, 0.075, 1.99E-08
7	87	KK	AcGDVEKGKK	3	5.42E-06, 0.028, 0.006
8	87/88	KIFVQK	KKGEREDLIAYLK	3	5.18E-05, 0.0002, 7.54E-06
8	87/88	KIFVQK	KKGEREDLIAYLK	4	0.002, 0.0004, 0.014
8	87/88	KKGER	KIFVQK	3	3.13E-13, 0.003, 7.43E-07
8	88	KIFVQK	KKGEREDLIAYLK	5	0.020, 0.019, 0.034
8/13	88	KGER	KIFVQKCAQCHTVEK	5	0.007, 0.128, 0.003
8	100	KATNE	KIFVQK	3	0.0001, 0.0001, 0.228
13	86	MIFAGIKK	IFVQKCAQCHTVEK	4	1.31E-08, 1.06E-08, 0.0006

Supplementary Table 1C continued

13	86	MIFAGIKK	IFVQKCAQCHTVEK	5	3.08E-07, 3.28E-06, 0.007
13	87	KKGER	IFVQKCAQCHTVEK	4	0.0005, 0.009, 0.004
13	87	KKGER	IFVQKCAQCHTVEK	5	9.68E-05, 0.103, 0.0001
22	100	KATNE	CAQCHTVEKGGK	3	0.0001, 0.001, 0.0003
25	39	GGKHK	KTGQAPGFSYTDANK	3	1.04E-10, 0.054, 7.69E-13
39	88	KKGER	KTGQAPGFSYTDANK	4	0.027, 0.297, 0.028
39	100	KATNE	KTGQAPGFSYTDANK	3	4.62E-09, 0.002, 1.81E-07
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	3	3.67E-06, 0.187, 3.69E-09
72	79	YIPGTKMIFAGIK	GITWGEETLMEYLENPKK	3	0.021, 0.040, 0.017
72	79	YIPGTKMIFAGIK	GITWGEETLMEYLENPKK	4	9.21E-07, 0.002, 2.57E-06
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	3	4.03E-08, 0.047, 9.33E-12
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	4	3.94E-11, 0.0003, 3.91E-09
72	86	MIFAGIKK	GITWGEETLMEYLENPKK	5	0.011, 0.245, 7.81E-05
73	86	KYIPGTK	MIFAGIKK	3	4.16E-09, 5.76E-05, 1.82E-07
73	86	KYIPGTK	MIFAGIKK	4	2.01E-07, 0.0008, 1.81E-05
73	87	KK	KYIPGTK	3	1.65E-06, 0.042, 0.0003
73	88	KGER	KYIPGTK	3	3.43E-08, 0.002, 1.77E-06
73	88	KYIPGTK	KGEREDLIAYLK	4	8.20E-06, 3.08E-05, 4.61E-05
79	88	KGER	YIPGTKMIFAGIK	4	0.005, 0.188, 0.020
86	88	MIFAGIKK	KGEREDLIAYLK	5	4.00E-05, 6.14E-06, 0.023
88	99	KGER	EDLIAYLKK	3	0.016, 0.180, 0.0009

Supplementary Table 1D. Identified cross-linked peptide ions in the aggressive cross-linking condition of 45°C/2500:1 molar ratio.

K1	K2	Peptide1	Peptide2	Z	Evalues
5	8	KIFVQK	AcGDVEKGK	3	4.36E-12, 8.38E-05, 7.23E-06
5	13	AcGDVEKGK	IFVQKCAQCHTVEK	4	0.013, 0.009, 0.023
5	13	AcGDVEKGKK	IFVQKCAQCHTVEK	5	0.0007, 0.0664, 0.0003
5	25	GGKHK	AcGDVEKGK	3	0.0002, 0.016, 0.001
5	27	AcGDVEKGK	HKTGPNLHGLFGR	4	2.83E-06, 0.036, 9.06E-07
5	73	AcGDVEKGK	KYIPGTK	3	0.002, 0.004, 0.019
5	87/88	AcGDVEKGK	KKGEREDLIAYLK	4	0.002, 0.002, 0.0003
5	88	AcGDVEKGKK	KGEREDLIAYLK	4	0.013, 0.016, 0.0005
5	88	KGER	AcGDVEKGK	3	2.62E-06, 0.044, 2.23E-06
5	88	KKGER	AcGDVEKGK	3	7.29E-06, 0.001, 0.0003
7	27	GKK	HKTGPNLHGLFGR	4	3.90E-08, 0.209, 2.51E-07
7	73	GKK	KYIPGTK	3	0.002, 0.089, 0.002
7	86	GKK	MIFAGIKK	3	0.009, 0.078, 0.008
7	87	KK	AcGDVEKGKK	3	8.48E-05, 0.064, 0.002
7	88	KGER	AcGDVEKGKK	3	0.003, 0.221, 0.0004
13	86	MIFAGIKK	IFVQKCAQCHTVEK	4	0.002, 0.0002, 0.292
13	86	MIFAGIKK	IFVQKCAQCHTVEK	5	0.003, 0.0003, 0.069
13	87	KKGER	IFVQKCAQCHTVEK	5	0.007, 0.112, 0.006
27	39	HKTGPNLHGLFGR	KTGQAPGFSYTDANK	4	0.015, 0.247, 0.0003
27	39	HKTGPNLHGLFGR	KTGQAPGFSYTDANK	5	0.008, 0.292, 9.31E-05

Supplementary Figure 1D continued

27	100	KATNE	HKTGPNLHGLFGR	3	1.51E-07, 0.0001, 7.18E-08
39	72	KTGQAPGFSYTDANK	GITWGEETLMEYLENPKK	4	0.029, 0.087, 0.005
39	73	KYIPGTK	KTGQAPGFSYTDANK	3	3.63E-06, 0.099, 4.24E-07
39	86	MIFAGIKK	KTGQAPGFSYTDANK	4	0.004, 0.213, 7.13E-05
39	88	KGER	KTGQAPGFSYTDANK	3	1.99E-07, .0078, 8.38E-12
39	99	GEREDLIAYLKK	KTGQAPGFSYTDANK	5	0.002, 0.212, 0.0001
39	100	KATNE	KTGQAPGFSYTDANK	3	5.77E-15, 0.0008, 1.71E-11
55	73	KYIPGTK	NKGITWGEETLMEYLENPK	4	0.011, 0.085, 0.002
72	79	YIPGTKMIFAGIK	GITWGEETLMEYLENPKK	4	0.006, 0.006, 0.0006
72	100	KATNE	GITWGEETLMEYLENPKK	3	4.26E-07, 0.078, 3.24E-10
73	86	KYIPGTK	MIFAGIKK	3	9.60E-07, 0.008, 0.0003
73	86	KYIPGTK	MIFAGIKK	4	0.0003, 0.26, 0.0007
73	100	KATNE	KYIPGTK	3	3.20E-06, 6.70E-05, 0.106
79	100	KATNE	YIPGTKMIFAGIK	3	0.006, 0.002, 8.86E-05
86	87	KK	MIFAGIKK	3	0.023, 0.044, 0.040
86	88	MIFAGIKK	KGEREDLIAYLK	3	6.32E-09, 4.28E-07, 3.30E-07
86	88	MIFAGIKK	KGEREDLIAYLK	5	0.0003, 0.0005, 0.013
86	100	KATNE	MIFAGIKK	3	0.005, 0.075, 0.002
88	99	KGER	EDLIAYLKK	3	0.018, 0.158, 0.0009
88	100	KATNE	KGEREDLIAYLK	3	0.002, 0.005, 0.049
88	100	KATNE	KKGEREDLIAYLK	3	0.002, 0.045, 0.008

CHAPTER 3

GENOTYPE AND DEVELOPMENTAL STAGE–DEPENDENT ASSYMETRIC LIPID DISTRIBUTIONS IN MAIZE THYLAKOIDS STUDIED BY MASS SPECTROMETRY IMAGING.

Adam T. Klein, and Young Jin Lee

Abstract

Maize has Kranz anatomy, where photosynthetic functions are asymmetrically distributed among two different photosynthetic cell types. We have recently shown differential localization of phosphatidylglycerols depending on fatty acid side chains between mesophyll and bundle sheath cells of B73 leaf tissue sections using high-spatial resolution mass spectrometry imaging. Here we expand the previous work, and compare the localizations and relative abundances of two major acidic lipids in thylakoid, sulfoquinovosyldiacylglycerols (SQDG) and phosphatidylglycerols (PG), at three different points across B73 and Mo17 inbred maize leaves. SQDGs were equally distributed in both the photosynthetic cells regardless of developmental stage, fatty acid, and genotype. However, the quantitative difference was noted between the two genotypes, especially the increase of SQDG 32:0 in Mo17 and decrease in B73. Similarly, PG 32:0 increases in Mo17 as plants develop whereas 16:1 containing PGs increase in B73. Most interestingly, 16:1 containing PGs contribute to thylakoid membranes in the mesophyll cells of B73 while bundle sheath cells are mostly composed of 16:0

51

containing PGs. In contrast, as 16:1 fatty acid containing PGs are deficient in Mo17, PG 32:0 is used to build thylakoid membranes of both mesophyll and bundle sheath cells.

Introduction

Thylakoid membranes play an important role in maintaining structural integrity of photosystem complexes[88]. Nishihara et al. studied the lipid molecular species of thylakoid membranes and showed significant differences for their fatty acyl chains depending on the lipid types[89]. Monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) are dominated by 18:3/18:3 fatty acyl chains, whereas the major species in sulfoquinovosyldiacylglycerol (SQDG) is 16:0/18:3 followed by 16:0/18:2. Phosphatidylglycerol (PG) is most unique in that its molecular species are vastly different depending on the plant species and even between mesophyll and bundle sheath of maize. Specifically, PGs with a 16:1 fatty acid at the sn-2 position, such as PG 32:1 and PG 34:4, were highly enriched in the mesophyll and PG 32:0 is mostly localized in the bundle sheath[89]. PG is also known for its correlation with chilling sensitivity; i.e., the plants with high level of disaturated thylakoid PGs are sensitive to chilling[90].

Mass spectrometry imaging (MSI) has become a valuable tool for analyzing molecular distributions of a wide range of compounds directly on plant or animal tissues[56-61]. Several different ionization techniques have been adopted for MSI; however, matrix assisted laser desorption ionization (MALDI) is the most attractive in terms of high spatial resolution, sensitivity, and chemical versatility, which is essential

for on-tissue single cell resolution imaging. The spatial resolution of MALDI-MSI has become routinely available down to the size of 20-30 μm and demonstrated as low as 2.5 μm[91].

Recently, we have visualized various metabolite distributions in the cross-section of a maize leaf using MALDI-MSI[38]. The cell-type specific asymmetric distribution of PG obtained by Nishihara in bulk chloroplast analysis was confirmed in our single cell level direct analysis. Minor differences for PG 34:2 and PG 34:3 were attributed to the high abundance of the 16:1 fatty acyl group in the older plant, which favors mesophyll, as we used 14 day old seedlings and they used 45 day old mature leaves. It is in accord with Roughan who reported the absence of 16:1 fatty acyl group in PG from the emerging portion of maize leaves and the increase toward the distal end[90].

Based on these precedents, in this work we explored quantitative fatty acyl distributions of PG and SQDG at three different locations across the maize leaves of two inbreds, B73 and Mo17, to understand the role of fatty acid compositions in the development and asymmetric distributions of thylakoid membranes of photosynthetic cells.

Experimental

Plant growth and sample preparation

Maize seeds (Zea *mays*, L. inbred B73 and inbred Mo17) were planted in soil and grown in a controlled environment in a greenhouse, with daily watering. Plant seedlings were harvested 13 days after planting for B73, and 15 days after planting for Mo17

during daylight hours (~1 PM). B73 has a faster growth rate and the difference in harvest time was to achieve similar growth heights, where the leaves were 15-20 cm long from ligule 2 to the tip of the leaf. Sections of the third true leaf were collected from three plants of each genotype at the proximal end of the exposed leaf (~2 cm above ligule 2), at the midpoint of the leaf, and at the distal end of the leaf (2-3 cm from the leaf tip).

Tissue samples were cryosectioned as described previously[54]. Briefly, plant tissues were submerged in gelatin and immediately frozen with liquid nitrogen. Transverse plant tissue sections were collected using a cryostat (Leica, CM1850, Leica Microsystems, Buffalo Grove, IL, USA) at a thickness of 10 μm. Due to the fragility of the tissue sections, the cryosectioned tissues were attached to adhesive tape windows to prevent metabolite delocalization. The leaf sections were dried and gradually warmed under moderate vacuum (~100 mtorr). Optical images of the dried tissues were collected before the application of matrix. 1,5-diaminonaphthalene (DAN, 97%, Sigma-Aldrich, St. Louis, MO, USA) was used as a matrix and applied by sublimation (~50 mtorr, 140°C, 4 minute heating time). The sublimation procedure is described in more detail elsewhere[92].

Mass spectrometry analysis

Data was collected on a MALDI-linear ion trap-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA, USA). The instrument was modified to use an external 355 nm frequency tripled Nd:YAG laser (UVFQ; Elforlight Ltd., Daventry, UK). Laser

energy was 4-5 µJ/pulse at a 60 Hz repetition rate. A 10x beam expander (Thorlabs, Newton, NJ, USA) was incorporated in the laser beam path prior to introduction into the mass spectrometer, which reduced the laser spot size to 9-10 µm. Tune Plus and Xcalibur (Thermo Scientific) were used to set the imaging acquisition parameters and data acquisition. Images were acquired using a 10 µm raster step size and 10 laser shots per step. Spectra were acquired in negative mode using the Orbitrap mass analyzer with a resolution of 30,000 and a m/z scan range of 100-1000.

MS/MS imaging was performed for selected ions in the same condition as above, but with an ion trap analyzer and a m/z scan range of 200-800. An isolation width of 2.0 Da and normalized collision energy of 35 were used. All images were generated using MSiReader[93] with a mass window of +/- 5 ppm and normalized to the total ion count (TIC) for MS images, and a mass window of +/-0.2 Da and no normalization for MS/MS images.

Additional analysis of selected tissues was performed using multiplex imaging with a 4 step spiral raster[94]. The first step collected a high resolution mass spectrum with the Orbitrap analyzer. MS/MS spectra were collected in steps 2-4 using the ion trap analyzer for the three most intense ions from the pre-loaded precursor mass list of common molecules of interest. Dynamic exclusion was used with a repeat count of 2, a repeat duration of 30 s, and an exclusion duration of 180 s.

Quantitative comparison was made by averaging metabolite intensities over a region of the maize leaf that contained both mesophyll and bundle sheath cells. Six

replicate regions of the same plant were analyzed, and the average of the six regions was reported.

Results and Discussion

The third true leaves from B73 and Mo17 inbreds were harvested and cryosectioned at three points along the leaf. The first section was collected at the point the leaf began unfolding and is exposed to light, herein described as the proximal end. The second section was collected at the midpoint between section 1 and the distal end of the leaf. The third section was collected ~2 cm from the distal end of the leaf.

The chemical images of various species of PG and SQDG are compared at the three different leaf cross-sections of B73 and Mo17 as shown in Suppl. Figures 1 and 2 along with three major metabolites closely related with photosynthesis (chlorophyll a, plastoquinone/plastoquinol, and carotene). Three representative lipids (PG 32:0, PG 32:1, and SQDG 34:3) are shown in Figure 1. The acidic lipid compounds of the thylakoid membrane (SQDG and PG) are most abundant at the midpoint of both inbreds. As we have previously reported[38], PG 32:0 (Figures 1D-F) and PG 32:1 (Figures 1G-I) are mostly localized in the bundle sheath cells and mesophyll cells, respectively, in the case of B73, whereas SQDG 34:3 is present in both the photosynthetic cells (Figures 1J-L). Mo17, however, shows slight differences with B73. SQDG has the same localization, but a lower abundance. PG 32:1 is nearly absent in Mo17 (Figure 1H') and PG 32:0 has a broader distribution (Figure 1E'), as discussed in more detail later.

Sulfoquinovosyldiacylglycerol distribution

SQDG is a chloroplast membrane lipid and is known to be present in both mesophyll and bundle sheath cells[38, 95]. Kenrick has reported that the majority of fatty acids that make up SQDGs are 16:0 and 18:3[96]. This is consistent with SQDGs detected in this study, as the most abundant SQDGs contain combinations of these two fatty acids; i.e., SQDG 32:0, 34:3, and SQDG 36:6 (See Suppl. Fig. 3 for MS/MS). Quantitative analysis was performed by averaging mass spectra in each MALDI-MSI dataset, then comparison was made between the same or similar class of compounds. Figure 2 shows the relative abundances of individual SQDGs with different fatty acid constituents at the three different locations of the leaf. SQDG 34:3 is most abundant across the entire leaf for both B73 and Mo17, and makes up over 60% of total SQDGs. Its relative abundance is least at the midpoint of B73, 63% compared to 74 or 79% at the proximal or distal end, mostly due to the increase of SQDG 36:6. The abundance of SQDG 32:0 is of interest, as the relative abundance increases from the proximal to distal ends in Mo17, but decreases in B73.

Phosphatidyl glycerol distribution

Roughan has shown that the most abundant PGs in younger parts of maize leaves predominantly have fatty acid compositions of 16:0, 18:2, and 18:3[90]. Figure 3 shows the relative abundance of PGs with different fatty acid constituents at each position along the leaf. The majority of PGs at the proximal side of the leaf are PG 32:0,

PG 34:2, or PG 34:3 for both B73 and Mo17. This observation agrees with that of Roughan as we confirmed their fatty acid compositions as 16:0/16:0, 16:0/18:2 and 16:0/18:3, respectively (Suppl. Figure 4). These three PGs are still the most abundant at the midpoint of the leaf (Figure 3B); however, the relative abundance of PG 32:0 is nearly doubled in both B73 and Mo17. This is accompanied by a significant decrease of PG 34:2 in B73 and PG 34:3 in Mo17.

At the distal end of the Mo17 leaf, PG 32:0 is enhanced even more extensively, which was accompanied by the decrease of PG 34:2 and 34:3. Overall, PGs with unsaturated fatty acyl chains decrease with leaf development while the disaturated PG increases, correlating well with the chilling sensitivity of Mo17[96-97]. The trend of PG 32:0 in Mo17 also correlates well with SQDG (Fig. 2); i.e., the increase of SQDG 32:0 from the proximal end to the distal end. In contrast, PG 32:0 at the distal end of B73 is decreased to the level of the proximal end along with the increase in PG 34:3 and 34:4. This is an opposite trend with Mo17, in agreement that B73 has lower chilling sensitivity than Mo17[98]. Interestingly, PG 32:1 is further increased from the level at the midpoint, which is in good correlation with the high levels of PG 34:3 and 34:4 compared to high levels of PG 34:2 and 34:3 at the proximal end. Namely, the 16:1 fatty acid component of PG is increasing in B73 as plants mature, similar to the 'Golden bantam' inbred in Roughan's work, contributing to the increase of PG 32:1, PG 34:3, and PG 34:4.

To further verify that 16:1 fatty acyl component of PG is increasing in B73 and 16:0 is increasing in Mo17, we performed MS/MS of PG 34:3 for all three places of B73 and

Mo17 leaves, and their relative composition of fatty acyl chains is summarized in Figure 4. There are two possible fatty acyl compositions in PG 34:3, 18:3/16:0 or 18:2/16:1. The portion of PG 18:2/16:1 is much more enhanced at the midpoint and distal part of B73, but PG 18:3/16:0 remains dominant in Mo17. This is also visualized in MS/MS imaging of PG 34:3 at the midpoint of leaves as shown in Figure 5, demonstrating 16:0 is dominant in Mo17 whereas 16:1 is quite comparable with 16:0 in B73.

The lack of 16:1 fatty acid in PG of Mo17 raises a question about the cellular localization of the lipid in mesophyll vs bundle sheath. Namely, if the increase of 16:1 component in PG is a general trend as plants grow, as shown for maize 'Golden bantam' and a few C3 plants by Roughan and maize B73 in this work, and also if 16:1 fatty acid of PG plays an important role in the mesophyll thylakoid membranes, as shown for maize 'Honey buntum' and a few C3 plants by Nishihara et al and B73 in this work (PG 32:1 of Fig. 1), what would be the fatty acid composition of PG in mesophylls of Mo17 where 16:1 is deficient? To answer this question, we carefully compared the images of PG 32:0, PG 32:1, PG 34:2, and PG 34:3 overlayed with the optical image at the midpoint of B73 and Mo17 leaves as shown in Figure 6. In B73, PG 32:1 is localized in the mesophyll, whereas the other three PGs are mostly localized to the bundle sheath. PG 34:3 is mostly localized to the bundle sheath in spite of the fact that it contains significant amount of PG 18:2/16:1. This is in contrast to Nishihara et al., where PG 18:2/16:1 is more abundant in the mesophyll, and attributed to the difference in genotype (B73 vs 'Honey buntum') or age (13 days vs 45 days). In Mo17, PG 32:1 is almost absent as shown in Figure 6B, consistent with the relative quantifications in Figure 3B. To

accommodate this deficiency, PG 32:0 is present in the mesophyll as well as bundle

sheath (Figure 6D). The other two PGs, PG 34:2 and PG 34:3, are localized in the bundle

sheath in Mo17, the same localizations as in B73.

Conclusions

Fatty acid compositions in membrane lipids have various important roles,

including special cellular or subcellular compartmentalization. In this study, we have

assessed the quantitative and localization change of major acidic lipids in the thylakoid

membranes of maize leaves as plants develop, and highlighted the difference between

the two genotypes. Overall, Mo17 shows the increase of 16:0 containing PGs and SQDGs

as plants develop, most notably 16:0/16:0, an opposite trend with B73. The 16:1

containing PGs play an important role in B73 as plants develop, similar to what is

known for other maize or some C3 plants. More specifically, PG 32:1 is a major

thylakoid membrane lipid in the mesophyll of B73, whereas other PGs constitute

bundle sheath cells. In contrast, the lack of PG 32:1 in Mo17 resulted in having PG 32:0

constitute mesophyll as well as bundle sheath cells. MALDI-MSI was chosen as the tool

of choice in this study and we clearly demonstrate the power of this technology for the

sophisticated analysis of single cell level high spatial resolution localization of lipids and

metabolites.

Acknowledgements
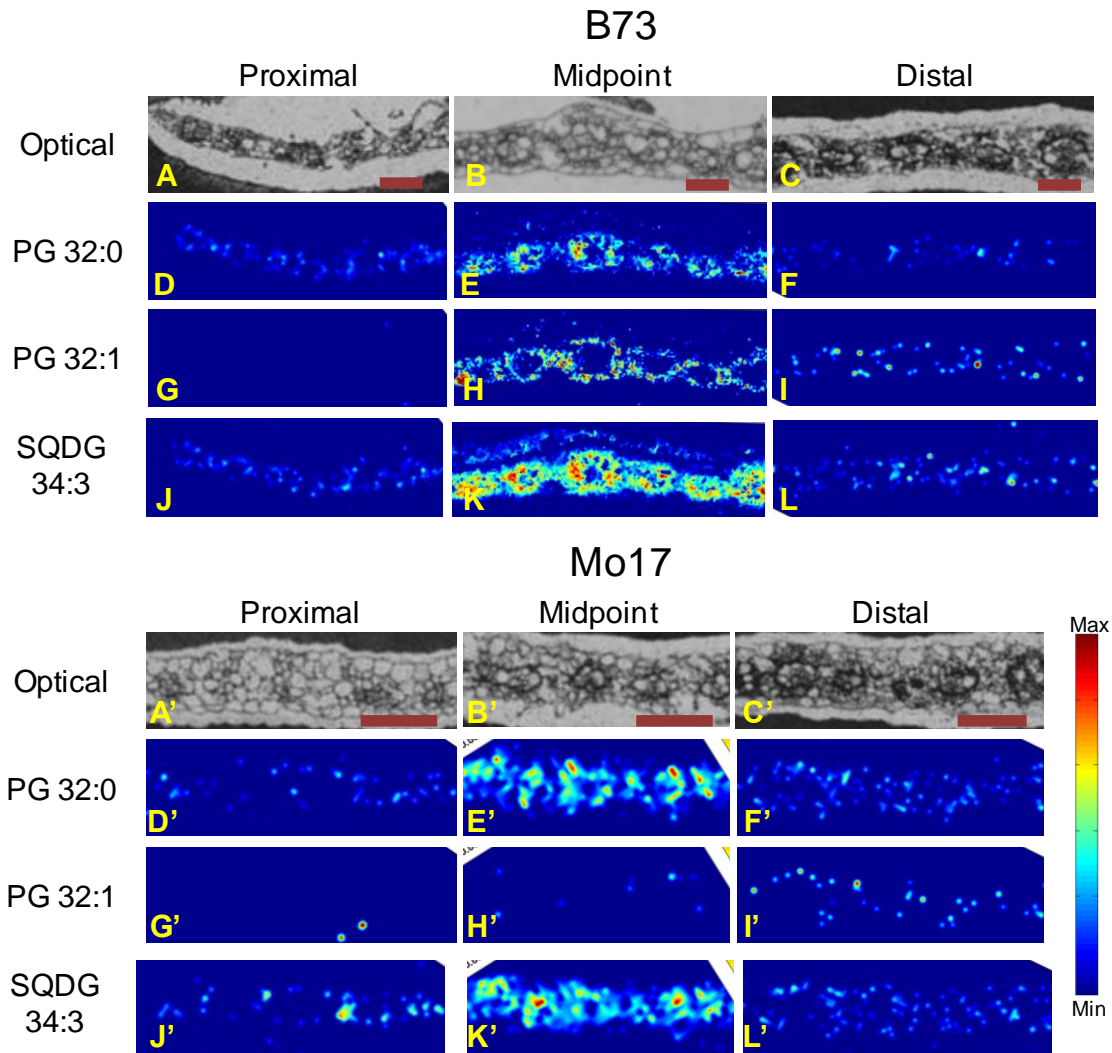
Figure 1. MALDI-MSI of selected metabolites. MALDI-MSI of the proximal, midpoint, and distal sections of maize leaf from inbred B73 (Top panel) and Mo17 (Bottom panel). Images shown are the optical images (A-C, A'-C') and chemical images of (D-F, D'-F') PG 32:0 (m/z 721.503; max = 3000), (G-I, G'-I') PG 32:1 (m/z 719.487; max = 1200), and (J-L, J'-L') SQDG 34:3 (m/z 815.501 max = 3000). Each analyte ion signal is normalized by the total ion count at each pixel, and the max values for color scheme were adjusted for the best comparison across the sections and between the genotypes. Scale bars represent 100 μm for each optical image.
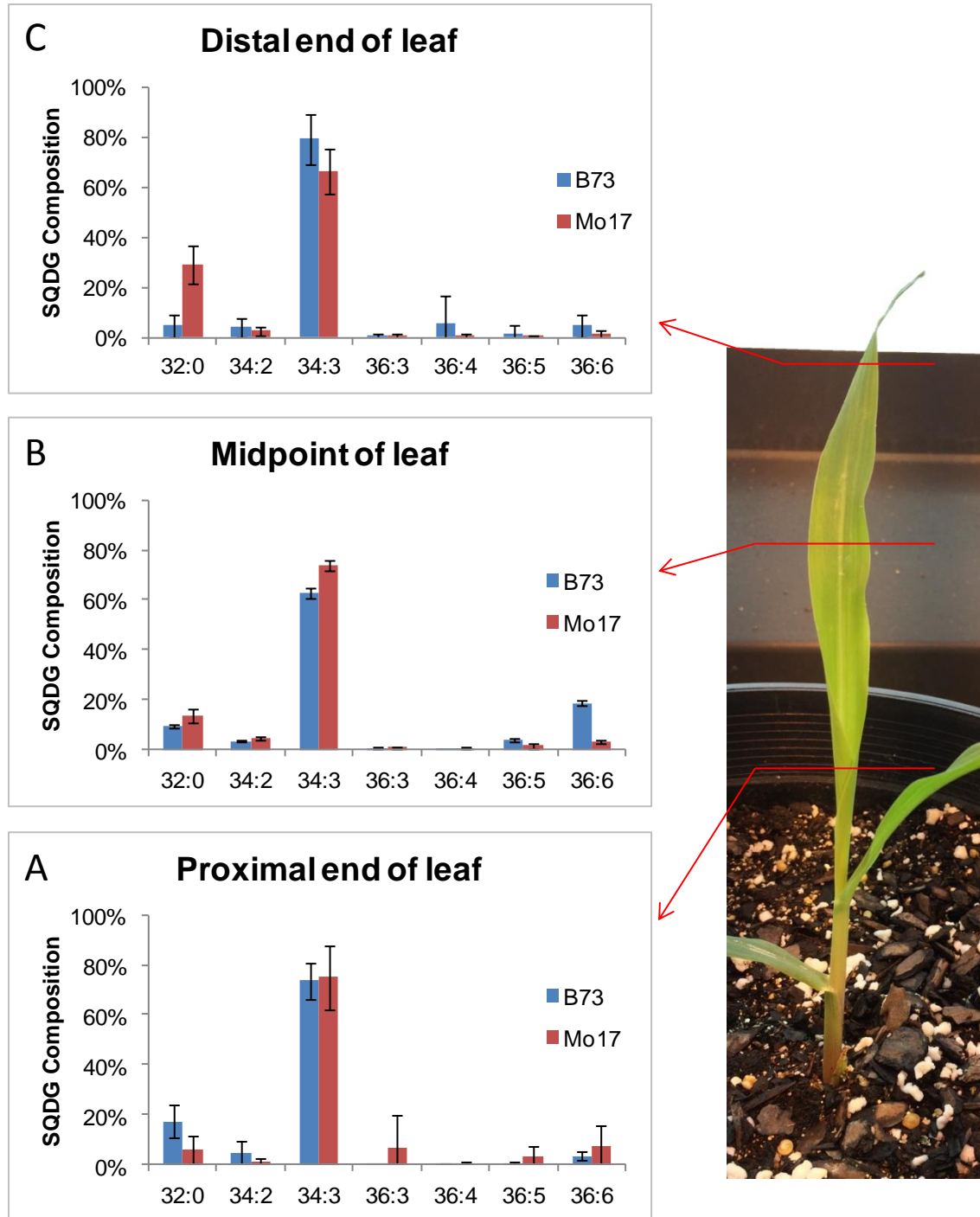
Figure 2. SQDG composition in maize leaves. The composition of each SQDG with different fatty acid constituents at (A) the proximal end, (B) midpoint, and (C) distal end of maize leaf for B73 and Mo17 inbreds.
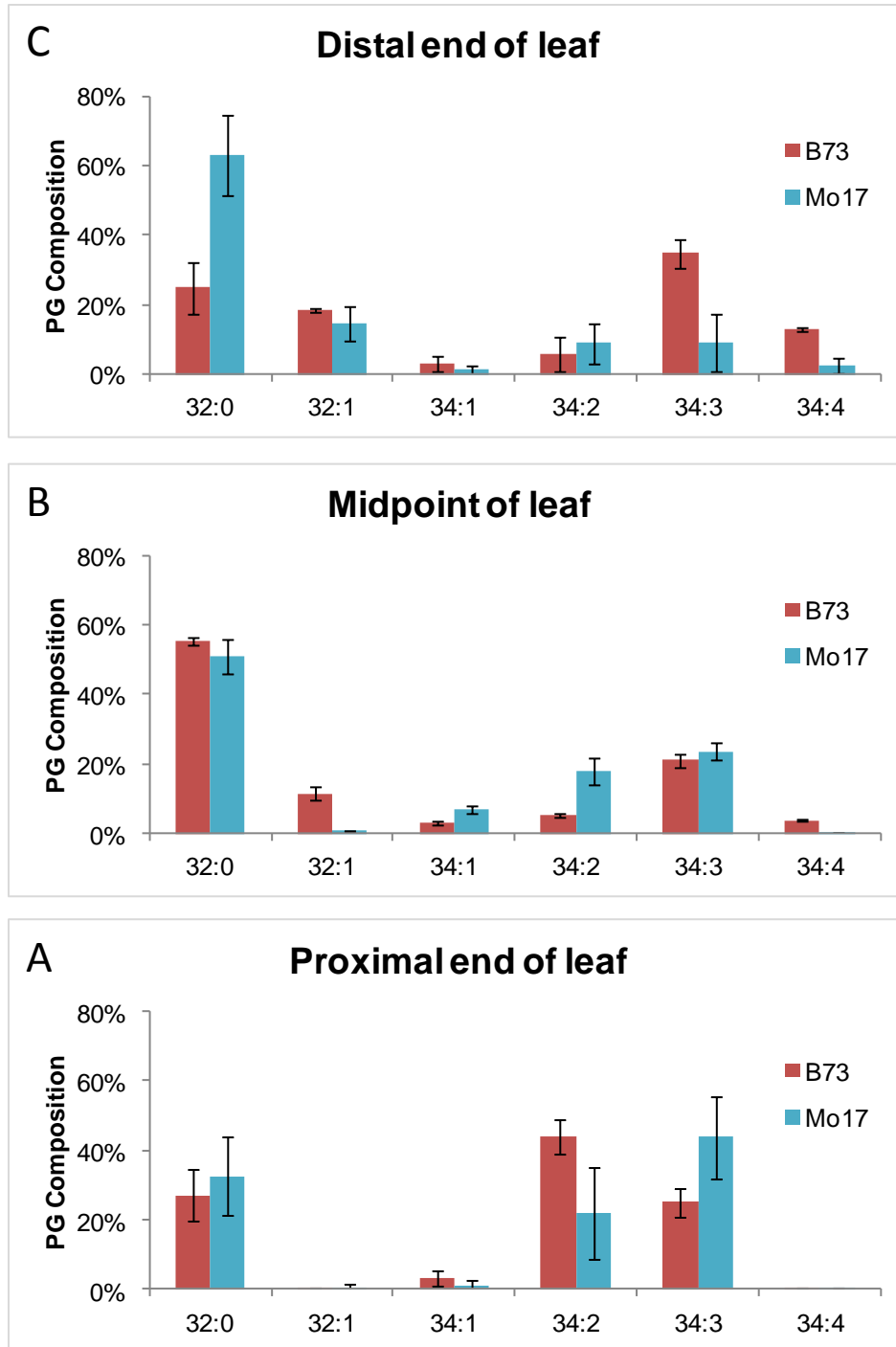
Figure 3. PG composition in maize leaves. The % of total PGs detected for individual PGs with different fatty acid constituents at (A) the proximal end, (B) midpoint, and (C) distal end of maize leaf for B73 and Mo17 inbreds.
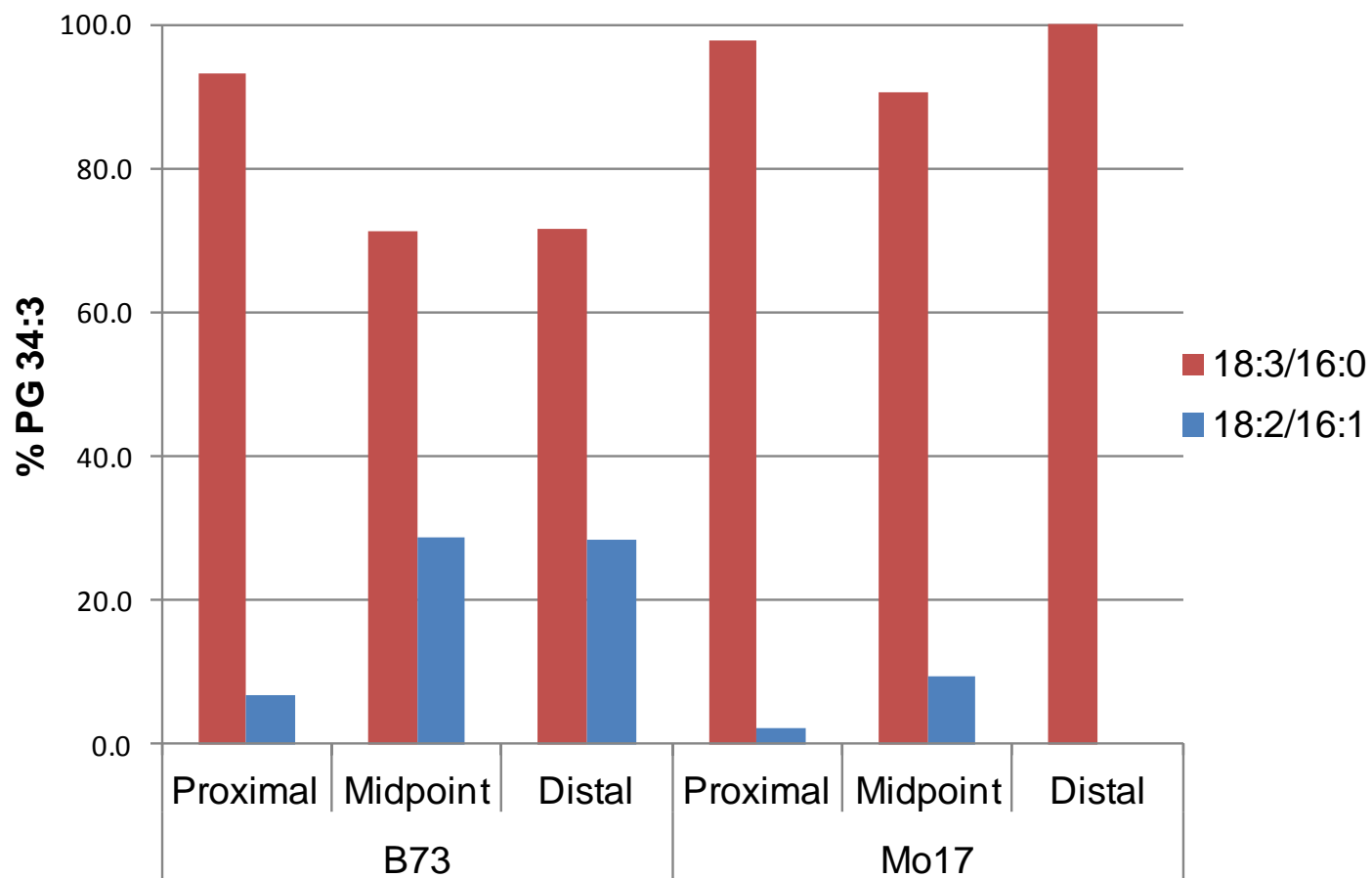
Figure 4. Relative abundance for MS/MS of PG 34:3. Relative abundance of 16:0 and 16:1 fragments from MS/MS of PG 34:3, *m/z* 743.49, in B73 and Mo17 maize leaves.
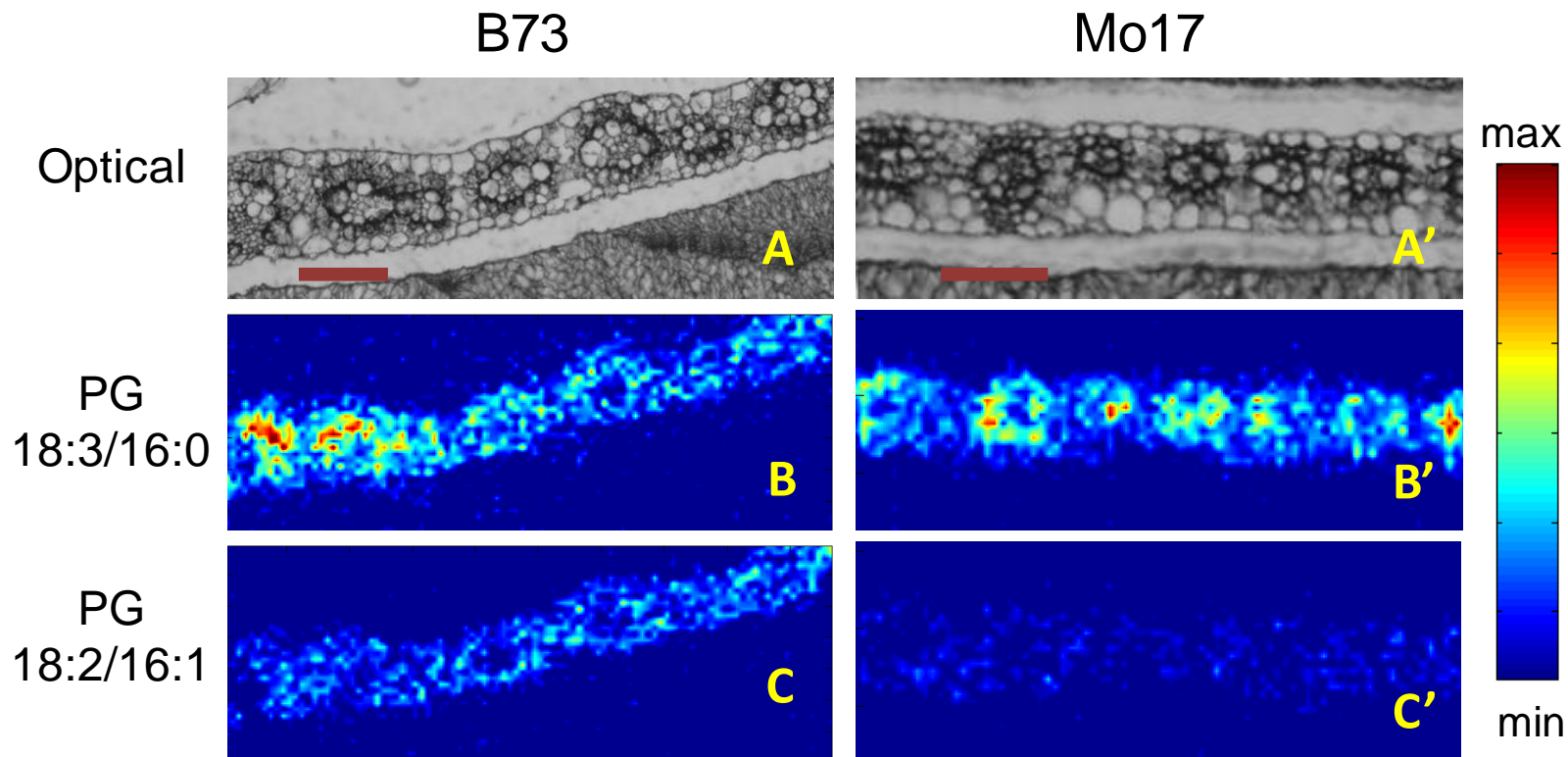
Figure 5. MS/MS images of PG 34:3. Chemical images of structural isomers of PG 34:3 at the midpoint of B73 and Mo17 leaves and corresponding optical images (A). The images for structural isomers were constructed from MS/MS images of the sum of major fragment ions corresponding to the fatty acid chains; (B) PG 18:3/16:0 (m/z 487.4 for [M−16:0]⁻, m/z 505.2 for [M-(16:0 − H$_2$O)]⁻, m/z 413.4 for [M-(16:0+glycerol)]⁻, and m/z 255.5 for 16:0), and (C) PG 18:2/16:1 (m/z 489.4 for [M−16:1]⁻, m/z 507.2 for [M-(16:1-H$_2$O)]⁻, m/z 415.4 for [M-(16:1+glycerol)]⁻, and m/z 253.3 for 16:1). Scale bar corresponds to 100 μm in each optical image.
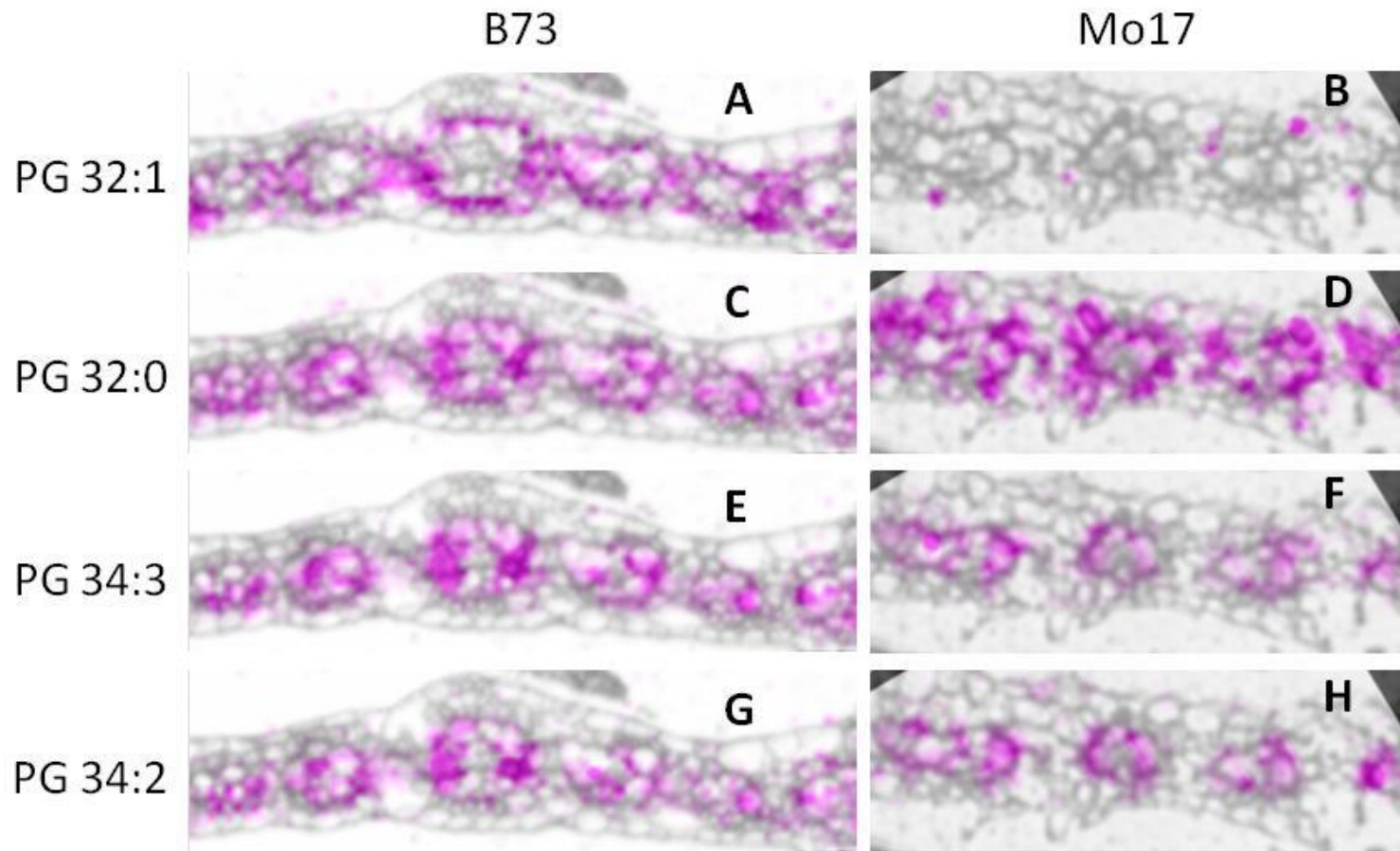
Figure 6. Overlay of optical image with PGs. Overlay of optical image with (A and B) PG 32:1, (C and D) PG 32:0, (E and F) PG 34:3, and (G and H) PG 34:2 at the midpoint of B73 and Mo17 maize leaves.
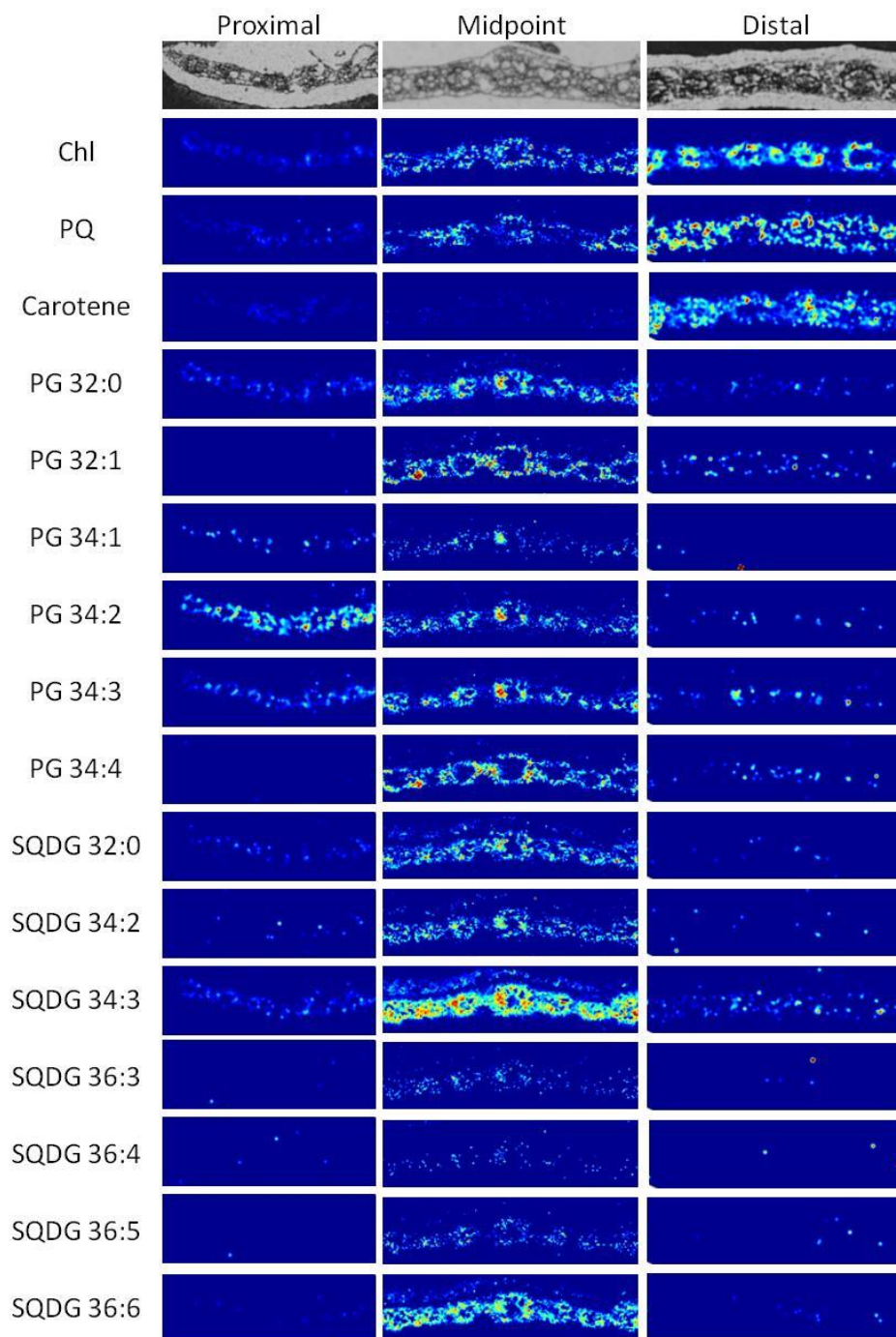
Figure S1. MSI of B73 maize leaf metabolites. MS images of PGs, SQDGs, and three photosynthesis related metabolites at the proximal, midpoint, and distal sections of B73 inbred maize leaf. Each analyte ion signal is normalized by the total ion count at each pixel, and the max values for color scheme were adjusted for the best comparison across the sections and between the genotypes. Chl: chlorophyll *a* (sum of fragments at m/z 592.267 (pheophorbide *a*), 613.232 (chlorophyllide *a*), and 870.566 (pheophytin *a*)). PQ: plastoquinone/plastoquinol.
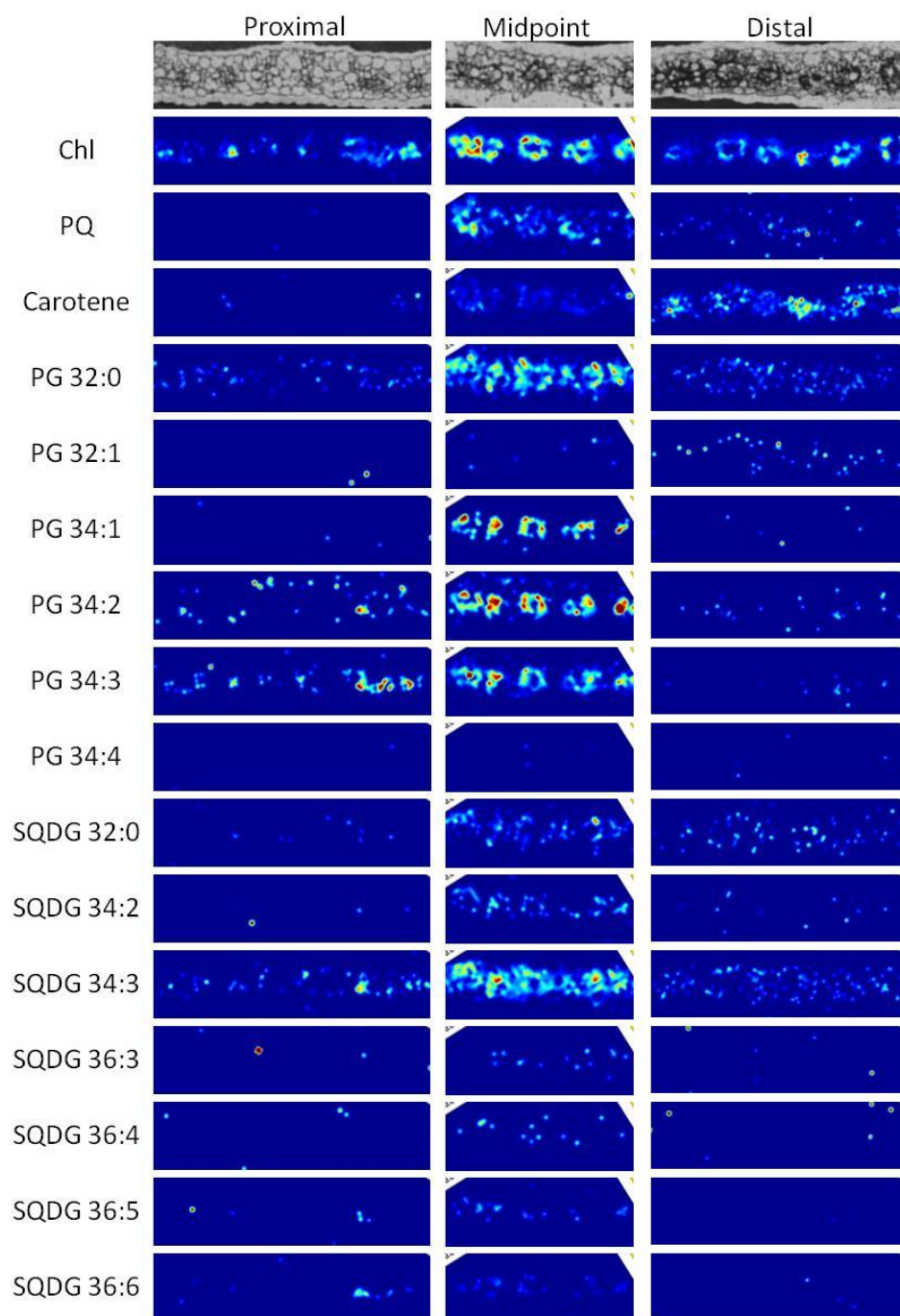
Figure S2. MSI of Mo17 maize leaf metabolites. MS images of PGs, SQDGs, and three photosynthesis related metabolites at the proximal, midpoint, and distal sections of Mo17 inbred maize leaf.
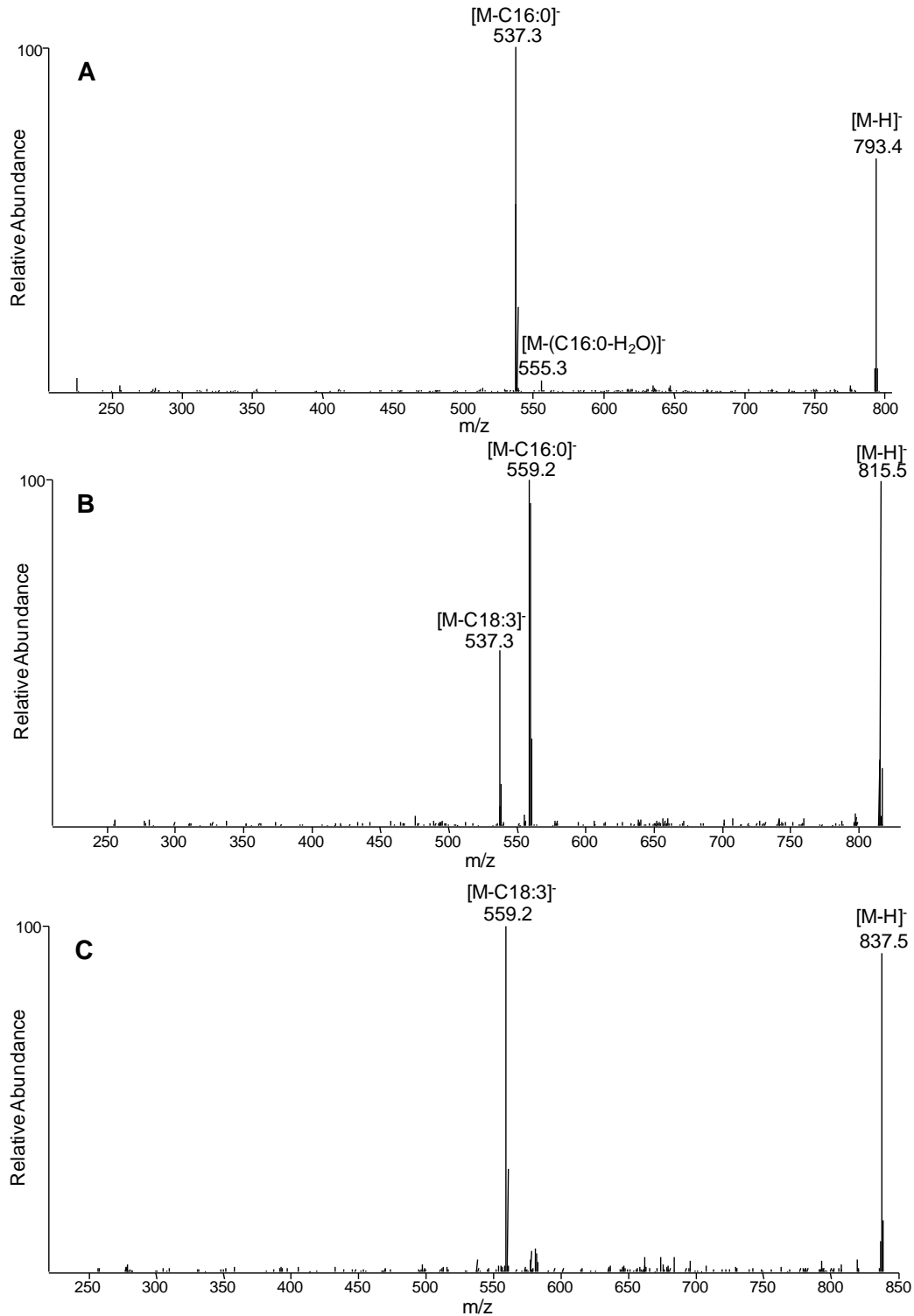
Figure S3. MS/MS of SQDG 32:2, 34:3, and 36:6. MS/MS spectra of (A) SQDG 32:0, (B) SQDG 34:3, and (C) SQDG 36:6 measured at the midpoint of B73 maize leaf. Major fragment ions are the loss of the 16:0 and 18:3 fatty acid side chains.
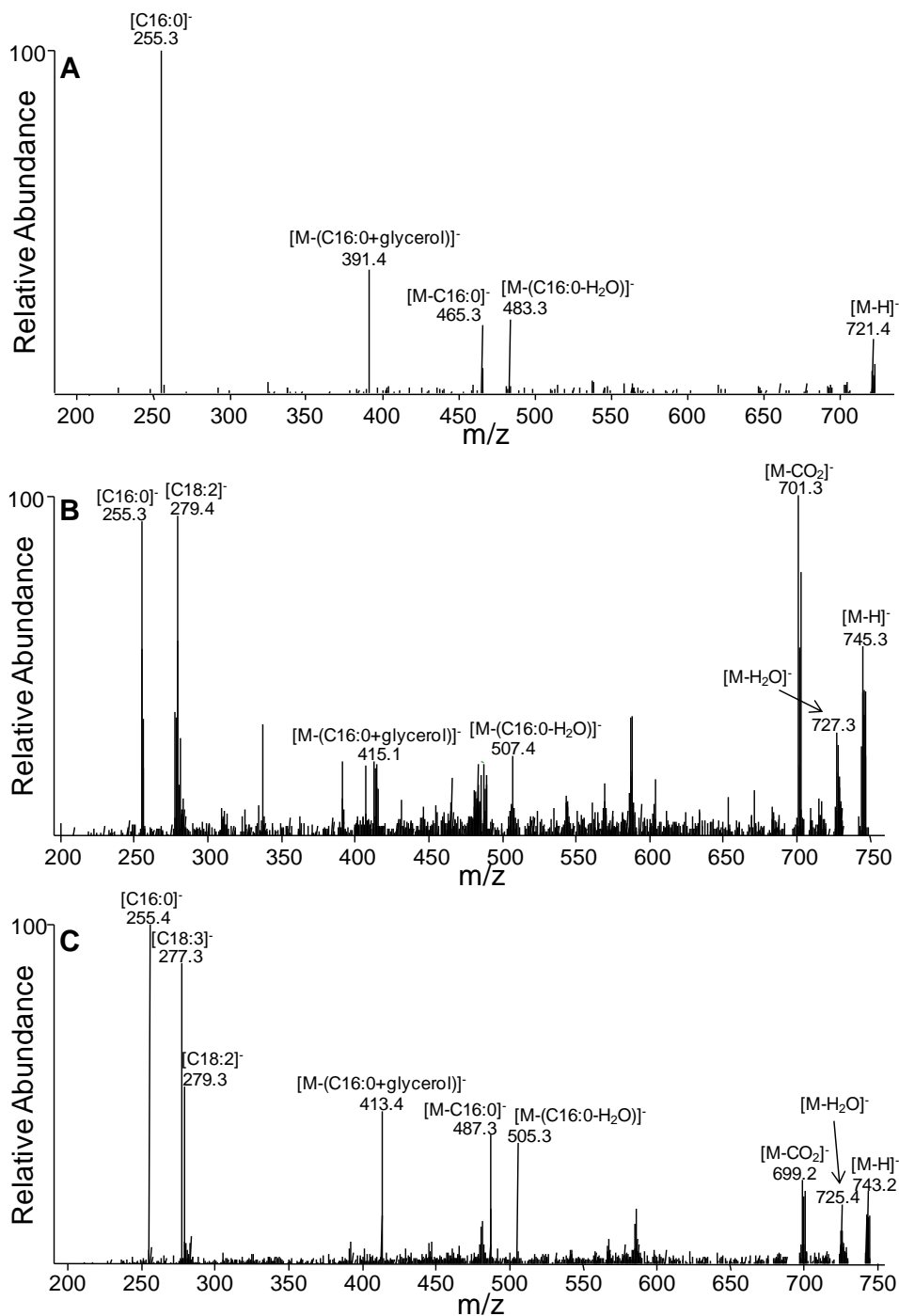
Figure S4. MS/MS of PG 34:3 and PG 34:2. MS/MS spectra of (A) PG 32:0, (B) PG 34:2, and (C) PG 34:3 from the proximal section of Mo17 maize leaf. Although the FA 18:2 anion was detected in PG 34:3, this is most likely due to the presence of other phospholipid species (i.e. PE 36:2) within the isolation width used for MS/MS. PGs are more likely to fragment at the sn-2 position[99], which is the location of 16:0 and 16:1 FAs[100-101], and 16:1 would be a better indication for the presence or absence of PG 18:2/16:1.

CHAPTER 4

INVESTIGATION OF THE CHEMICAL INTERFACE IN THE SOYBEAN-APHID INTERACTION USING MALDI-MASS SPECTROMETRY IMAGING

Adam T. Klein, Gargey B. Yagnik, Jessica D. Hohenstein, Zhiyuan Ji, Jiachen Zi, Malinda D. Reichert, Gustavo C. MacIntosh, Bing Yang, Reuben J. Peters, Javier Vela, and Young Jin Lee

Abstract

Mass spectrometry imaging (MSI) is an emerging technology for high-resolution plant biology. It has been utilized to study plant-pest interactions but is limited to the surface interfaces. Here we expand the technology to explore the chemical interactions occurring inside the plant tissues. Two sample preparation methods, imprinting and fracturing, were developed and applied, for the first time, to visualize internal metabolites of leaves in matrix-assisted laser desorption ionization (MALDI)-MSI. This is also the first time nanoparticle-based ionization was implemented to ionize diterpenoid phytochemicals that were difficult to analyze with traditional organic matrices. The interactions between rice-bacterium and soybean-aphid were investigated as two model systems to demonstrate the capability of high-resolution MSI based on MALDI. Localized molecular information of various plant- or pest-derived compounds provided valuable insight for the molecular processes occurring during the

plant-pest interactions. Specifically, salicylic acid and isoflavone based resistance was visualized in the soybean-aphid system, and antibiotic diterpenoids in rice-bacterium interactions.

Introduction

In nature, plants encounter diverse pests including pathogens and insects; thus the plant-pest interface is a battlefield of continuous chemical attacks and defenses. Plants possess an innate immune system that detects conserved pathogen-derived molecules and trigger an array of defense responses. Conversely, these pests deploy effector molecules to suppress innate defenses and to manipulate the host metabolism for nutritional benefit.[102-105] Most current work on the chemical interactions between plants and pests involves *ex situ* extraction of metabolic or genetic materials with homogenization of various tissues and cell types.[106-109] This approach has many limitations. Firstly and most importantly, it fails to provide critical information on the spatial distribution of compounds. Secondly, the homogenization process significantly dilutes analytical signals to levels that are sometimes indistinguishable from the background (false negative). Lastly, the lack of spatial information may cause chemical responses to non-target stimuli to be misattributed to plant-pest interactions (false positive).

Laser capture microdissection (LCM) has been applied for metabolic profiling of plant tissues.[110-111] This method can circumvent the above limitations by precisely cutting out specific cells or tissues of interest. However, it has its own limitations

including potential delocalization of small molecules during fixation and embedding. Most of all, in a typical LCM analysis, tens or hundreds of cells need to be collected, resulting in extended sample preparation time and the inability to measure asymmetric metabolic distributions among the same cell types. Liquid extraction surface analysis mass spectrometry (LESA-MS) is another method recently developed to overcome the lack of spatial information in traditional metabolic profiling.[112] In this approach, a robotic arm automatically extracts soluble materials from each localized area using a micropipette and introduces the sample directly into the mass spectrometer. This method minimizes sample preparation and allows imaging experiments[113]; however, its spatial resolution is limited to only about one millimeter size.

In situ imaging can provide fine details with high-spatial resolution. However, molecular imaging in plants is mostly limited to fluorescence,[114] which provides subcellular distributions but only those of macromolecules labeled as targets. Accumulation of small molecules is one of the most direct evidences of plant-pest interactions but the technology for their in situ chemical imaging is nearly nonexistent. Mass spectrometry imaging (MSI) is an excellent tool of choice for imaging small molecules.[115] It provides extremely high sensitivity, down to attomole (10^{-18} mole) levels[116], a high-degree of in situ chemical mixture separation (up to thousands of molecules in m/z space), and the ability to characterize novel molecules; thus, it is extremely well-suited for metabolite imaging. MSI has recently been used extensively for imaging plant metabolites[60, 117-118], including our own work.[62, 119-120]

Most MSI applications for plant materials are made through direct interrogation on plant surfaces or thin sections obtained by cryosectioning.[105, 117] Unfortunately, neither method is applicable for visualization of the internal metabolite distributions of thin leaves or flower petals along the lateral dimension. Imprinting internal plant molecules onto Teflon or silicate surfaces has been proven to be an efficient sample preparation method for this purpose[121] and is extensively utilized for desorption electrospray ionization (DESI).[117] However, DESI is limited to polar compounds analysis and its spatial resolution is typically ~200 μm[122], although <50 μm resolutions have been reported[123-124]. NanoDESI enables up to 10 μm spatial resolution,[125] but is not commercially available. MSI using matrix-assisted laser desorption ionization (MALDI) is a versatile alternative as it allows much higher spatial resolution and analysis of wide classes of compounds. Combining cell-transfer by contact printing and carbon-substrate-based laser desorption ionization, Li et al. profiled and imaged trichome specialized metabolites of *Solanum habrochaites*.[65] Plant metabolites associated with pest defense have been investigated using MALDI-MSI[126-127], however, no study has been made for MSI to interrogate internal metabolites induced by plant-pest interactions. Hamm et al. utilized MSI to image the chemical responses of *Cabernet Sauvignon* after infected by *Plasmorpara viticola*.[128] This work analyzed only a few target molecules, resveratrol and pterostilbene, on the surface of the leaf by direct laser desorption ionization.

Two biological systems are used to study the chemical interfaces in plant-pest interactions in the current work: soybean-soybean aphids (*Glycine max)* colonized with

Aphis glycines and rice-bacterium (*Oryza sativa)* infected with *Xanthomonas oryzae* pv. *oryzae* (Xoo). Soybean aphids are native to Asia and are a recently introduced pest affecting soybean plants in North America. Soybean aphid infestation results in changes in the metabolome of soybean plants. Some of these changes include the accumulation of phytohormones such as salicylic acid, ethylene, and jasmonic acid that can mediate defense responses,[109, 129] changes in amino acids that may modify the nutritional capacity of the plant,[130] and changes in chlorophyll content.[131] However, our knowledge of the soybean metabolite changes in response to aphid colonization is limited, and their spatial distribution is unknown. Understanding how the metabolites change during plant-pest interactions can be crucial to understanding a plant's resistance or susceptibility to attack.

Here we demonstrate the utility of MALDI-MSI as a tool to investigate plant-pest interactions that is not limited to the surface analysis, but includes the interior of the plant tissues where metabolism is significantly altered by the interactions. Specifically, the plant responses to two distinct pest classes, insect (aphid) and bacterium (Xoo), were investigated. To expose internal molecules for interrogation by the laser beam, we developed and applied 'imprinting' and 'fracturing' methods. The imprinting technique is used to squeeze out internal metabolites with minimal delocalization during the process. Fracturing is a simple method often used in optical microscopy to expose and visualize internal plant tissues, but here we applied this technique, for the first time, for MS imaging. Additionally, we successfully applied nanoparticle-assisted laser

desorption ionization (NALDI),[136-137] for characterization of diterpenoids. The rice bacteria interaction portion is included in the dissertation of Gargey B. Yagnik.

Experimental

Materials

Methanol (LCMS Chromasolv) and 1,5-diaminonaphthalene (DAN, 97%) were purchased from Sigma Aldrich (St. Louis, MO, USA). Dihydroxybenzoic acid (DHB, 99.0%) was purchased for Acros Organics (Pittsburgh, PA, USA). Porous polytetrafluoroethylene (PTFE) sheets were 45 to 50% porous with 7-14 μm pore size (Porex Corporation, Fairburn, GA, USA).

Soybean plant growth

Aphid-susceptible soybean plants (*Glycine max* (L.) Merr.) cv. SD01-76R were grown in a growth chamber at a constant temperature of 25°C with a 16 light: 8 dark photoperiod. Seeds were surface sterilized overnight using chlorine gas as previously described.[138] In each pot, two seeds were planted in steam sterilized Metro-Mix® 900 soil (Sun Gro Horticulture, Vancouver, BC, Canada) and after one week, seedlings were thinned to one per pot. For the duration of the experiment, plants were watered twice weekly and additionally fertilized once per week with a 1:1 mixture of all-purpose Miracle-Gro Excel (21-5-20, The Scott's Company LLC, Marysville, Ohio, USA) and Miracle-Gro Professional (15-5-15, The Scott's Co.) applied at a 12.5 mL L^{-1} in water. Plants were infested at V2 stage.

Insect material and aphid infestation

Soybean aphids [*Aphis glycines* Matsumura (Hemiptera: Aphididae); biotype 1] were obtained from a laboratory colony maintained on SD01-76R plants at Iowa State University. Experimental plants were infested by transferring 30 mixed-age apterous aphids to the abaxial side of the center leaflet using a small paintbrush. To prevent movement, aphids were confined using clip cages (BioQuip products, Rancho Dominguez, CA, USA). Clip cages were fastened on both experimental and control plants to mimic any environmental changes caused by the cage. Aphids were allowed to feed and reproduce on the leaflet for seven days.

Sublimation of organic matrix

DHB was applied on the PTFE sheet by sublimation, as described by Hankin et al.[139], and subjected to MALDI-MSI. Sublimation was performed by placing 500 mg DHB into the bottom of the apparatus and placing it on a hot plate or heating mantle. The imprinted sample was adhered to a microscope slide with double-sided tape, which was attached to the condenser part of the apparatus with tape. The apparatus was sealed and evacuated to ~150 mtorr. Once the pressure stabilized, ice water was added to the condenser and the heating mantle was turned on to 230 °C. Heat was applied for 5-7 minutes, until matrix sublimation was visually observed on the sides of the apparatus and the bottom of the condenser. The apparatus was removed from the heating mantle and allowed to cool. The ice bath was warmed to room temperature by adding more water to prevent condensation when the apparatus was opened. The

matrix-coated PTFE membrane was removed and subjected to MALDI-MSI data acquisition.

Matrix application with oscillating capillary nebulizer

We have previously demonstrated diaminonaphthalene (DAN) is a good matrix for small molecules in negative mode[54]. For the negative ion mode of soybean-aphids interactions, DAN was dissolved in methanol at 10 mM concentration and 300 μL of solution was sprayed on the imprinted sample using a home-made oscillating capillary nebulizer (OCN).[62] The procedure for matrix solution application was as follows: (1) fill up the 500 μL syringe with matrix solution, (2) place a blank target plate 8-10 cm below the tip of the OCN, (3) start the nebulizing gas (N_2) flow and adjust the pressure to ~40 psi, (4) set the flow rate of the syringe pump at 50 μL/min and start the flow, (5) monitor the blank target plate surface to ensure matrix is being applied and wetting is minimal, (6) stop the syringe pump and place the PTFE membrane below the tip of the OCN, and (7) turn on the syringe pump and spray 500 μL of matrix solution that covers a ~2 cm sample uniformly.

Imprinting leaf metabolites to porous PTFE

Aphids were removed from the leaf using forceps, a brush, and air duster. The leaf was removed from the plant and promptly included in the imprinting setup. The imprinting of the plant leaf on a porous polytetrafluoroethylene (PTFE) sheet was performed as described by Thunig et al.[121] As shown in supplementary information

(Figure S1) the plant leaf or leaf section was placed with the abaxial side toward the porous PTFE sheet. The leaf was covered with three laboratory wipes in order to absorb any excess liquid extracted during imprinting. Three additional laboratory wipes were placed behind the porous PTFE sheet to absorb any liquid that is squeezed through the PTFE sheet. This was then placed between two metal plates and inserted in a vise. The vise was hand tightened and kept under pressure for 3-4 minutes. Under pressure the wipes absorb excess water and leaf material released during the imprinting process. After the defined time interval the porous PTFE sheet and leaf were removed from the vise. Control samples were also imprinted to account for any artifacts that may occur during sample preparation.

MSI data acquisition and data processing

A linear ion trap-Orbitrap mass spectrometer with MALDI ion source (MALDI LTQ-Orbitrap Discovery; Thermo Scientific, San Jose, CA, USA) was used for the current study. The instrument was modified to use an external frequency tripled, diode pumped Nd:YAG laser operating at 355 nm (UVFQ; Elforlight Ltd., Daventry, UK). Laser energy of about 4-5μJ/pulse and 5-7μJ/pulse were used for MSI of soybean leaf and rice leaf, respectively, at 60 Hz repetition rate. The laser spot size was estimated to be 30~40 μm as determined from laser burn marks on a thin film of α-cyanohydroxycinnamic acid. A raster step size of 100 μm and 150 μm was used for MSI of rice leaf and soybean leaf, respectively, unless otherwise noted. Each spectrum was collected with 10 laser shots

per scan and one scan per raster step. Orbitrap scans were acquired over the *m/z* range of 100 to 1200.

ImageQuest software (Thermo Scientific) was used to produce chemical images from MS imaging data sets. A mass tolerance of ±0.003 Da was used for generating Orbitrap MS images. All the MS images were normalized against total ion count at each pixel and the maximum scale was arbitrarily adjusted to produce the best quality images as indicated in the figure caption. All the image features shown in this manuscript were reproduced in at least two or three replicate experiments.

Results and Discussion

Imprinting of plant leaves for MALDI-MSI

The imprinting of plant leaves, as far as we are aware, has not been used for MALDI-MSI to visualize internal metabolites. DESI-MSI has been widely used for this purpose but MALDI-MSI is expected to provide superior spatial resolution. Its high spatial resolution, however, would be useful only if the localization of plant metabolites is retained during the brute force squeezing process. To study the extent of metabolite delocalization during the imprinting procedure, we have obtained MALDI-MSI of a soybean leaf with a 30 μm raster step after imprinting to a PTFE surface. Spatial resolution of 20-30 μm can be routinely obtained in most commercial MALDI-MSI instruments, while most DESI-MSI is performed at the size of 200-300 μm. Figure 1 shows MS images of a few representative metabolites. High image quality suggests the localization information is mostly retained during the imprinting.

The ion intensity profiles of choline and phosphocholine are constructed across the x-axis at a fixed y-position of Figure 1, as shown in Figure 2. Some half bandwidths are as narrow as ~50 μm (e.g. red arrow) and some rising or falling half widths are as narrow as ~20 μm (e.g. purple arrow). Considering inherent metabolite distributions, the broadening originating from the sample preparation and mass spectrometric measurement must be narrower than 50 μm, including the delocalization of metabolites during imprinting and the laser beam sampling size (~30 μm for this particular data set). Hence, we estimate a spatial resolution as high as 20-30 μm is potentially achievable using the imprinting sample preparation in MALDI-MSI. The raster step size of 150 μm was used for the rest of the imprinting experiments because of the wide imaging area that had to be covered, which is still better than that of typical DESI-MSI. It should be noted that we cannot discard the possibility that there might be more delocalization of metabolites on the thicker part of tissues such as midrib due to the greater pressing force.

Application of imprinting method for soybean-aphid interactions

The imprinting method developed above was applied to study the chemical interactions between soybean and aphids. Soybean leaves had been infested with aphids for 7 days prior to MALDI-MSI analysis. The aphids were kept inside a clip-cage on the center of the leaf to restrict aphid movement and identify the region where feeding occurred. At 7 days, aphids were removed and leaf metabolites were transferred to a porous PTFE membrane by the imprinting method described in the

experimental section. Uneven leaf surfaces are often difficult to be directly probed by a laser beam because surface roughness may result in the loss of the depth of field and insufficient laser flux for some local areas. The flat surface of the membrane provides uniform laser energy over the entire sampling area regardless of original surface roughness due to veins or trichomes. Sample handling is also simplified with imprinting by eliminating the long drying process required for a large leaf and minimizing cracking and curling during the drying process.

The PTFE imprints were analyzed with dihydroxybenzoic acid (DHB) as a matrix in positive ion mode and with 1,5-diaminonaphthalene (DAN) as a matrix in negative ion mode. Figure 3 shows positive ion *m/z* images of various compounds. The MALDI-linear ion trap-Orbitrap mass spectrometer used in this experiment provides accurate mass information in the Orbitrap and structural information by MS/MS in the linear ion trap, allowing positive identifications for many compounds directly on tissue. The compound assignments were based on Metlin metabolite database searches (http://metlin.scripps.edu/) with mass tolerance of 3 ppm. MS/MS experiments were performed using the ion trap to further support the assignments as shown in the supplementary data for oligosaccharides, phosphocholine, and arginine (Figure S2 and S3). Other ions either did not have sufficient signals for MS/MS or produced almost no fragments.

Hexose sugars ranging from one to four monomer units were detected as potassiated ions ($[H_2O(C_6H_{10}O_5)_n+K]^+$, n=1-4) and localized to the caged aphid region of the leaf, as seen in Figure 3B-3E. We hypothesize that the images of these sugars

correspond to honeydew deposited on the surface of the soybean leaf. Although soybean plant cells contain all these hexose sugars including raffinose (tri-hexose) and stachyose (tetra-hexose), raffinose and stachyose are normally not detected in soybean leaves.[140-141] It has been shown that aphids are able to reduce the osmotic pressure in their gut through the activity of a sucrase-transglucosidase enzyme[142] that transforms the glucose moiety of sucrose into oligosaccharides that are then excreted in the honeydew together with glucose and sucrose.[143-144] Glucose, fructose, sucrose, maltose and erlose have previously been found in soybean aphid honeydew.[145] To evaluate this hypothesis, we briefly rinsed off the soybean leaf surface with water prior to imprinting. As shown in Figure S4, the hexose sugars mostly disappeared after the rinse consistent with this explanation. Only di-hexose, most likely sucrose which is abundant in plant leaves, is barely visible in spite of ten times higher intensity scales than in Figure 3.

Figures 3G-3L show that the nucleobases (adenine, guanine, cytosine), phosphocholine, and amino acids (glutamine, arginine) are co-localized but do not share the spatial distribution of honeydew. Instead, their pattern aligns with the dark spots in the optical image (Figure 3A), which likely correspond to dead cells. Interaction of aphid saliva effectors with defense-related plant phenolics or reactive oxygen species (ROS) may result in leaf tissue browning and chlorosis, and eventually necrosis.[146] This cell death could lead to metabolite leakage from cells that would explain the pattern observed. Alternatively, these patterns of metabolite accumulation could be related to metabolic modifications induced by aphids. In compatible interactions, aphids can

84

modify nitrogen metabolism and amino acid accumulation, to enhance the nutritional value of colonized leaves.[130, 147] Further study is necessary to better understand the obtained metabolite distributions.

Figure 4 shows a comparison of negative mode ion images for soybean leaves with and without aphids. The compound assignments were based on Metlin database searches with 3ppm mass tolerance and literature searches. MS/MS were performed directly on the PTFE surface for the shown compounds; however, some of them had significant interference within ±0.5 Da and could not obtain meaningful MS/MS spectra (Figure S5). Figure 4B and 4B' shows kaempferol-rhamnoside-glucoside with and without aphids (see MS/MS in Figure S5). Kaempferol glycosides normally accumulate in mature soybean leaves, and their major role is UV protection. Accordingly, the same distributions were observed across the entire leaf, regardless of the presence of aphids, suggesting that most housekeeping metabolites are not affected by aphids.

Figure 4C, *m/z* 128.071 tentatively assigned as deprotonated pipecolic acid (Pip), shows a similar image to that of the protonated form in positive ion mode (Figure 3F) in that both are distributed in the non-vein area within the cage. This assignment is based on the accurate mass and the recent report of Pip in *Arabidopsis* and tobacco as a positive regulator for systemic acquired resistance.[148-149] Pip is a catabolite of lysine and can be easily protonated or deprotonated, consistent with its observation in both positive and negative ion mode. It was consistently detected in the aphid region but was always absent in the control (Figure 4C') for several replicate experiments, confirming it is a result of aphid-infestation, and not an experimental artifact.

Another supporting evidence for Pip is the accumulation of salicylic acid (SA) in the caged region of the leaf infested by aphids (Figure 4D) but not in the control (Figure 4D'). The assignment of SA is supported by MS/MS for its characteristic CO_2-loss (Figure S5). Unlike Pip, SA was observed only in negative ion mode, as expected from its low proton affinity. SA-mediated plant defense response is well known and reported for the soybean-aphids system.[109, 129] In Figure 4D and 4D', however, SA is also observed in the veins on the proximal side in both leaf samples. While more work is needed to understand this observation, it is important to note that both control and aphid-infested plants were grown in close proximity in the same growth chamber. This experimental setup may have been conducive to the development of priming in control plants, manifested as an accumulation of SA in the phloem. It was shown that the main soybean volatile induced by soybean aphid infestations is Methyl-SA,[150] and this compound is a strong inducer of priming.[151] Moreover, evidence of priming in the experimental setup used here has been described before.[109]

Figures 4E and 4F show m/z 267.066 and 253.050, tentatively assigned as formononetin and dihydroxyflavone (e.g. daidzein), respectively; isoflavones known to be induced in soybean leaves by insects.[152] They are both localized to the caged region on the leaf with aphids, and are not detected in control plants (Figures 4E' and 4F'). Several soybean aphid resistance quantitative trait loci (QTL) have been mapped to regions that also contain QTL for high isoflavonoid content, and higher content of isoflavonoids, especially daidzein, has been found for soybean leaves with aphids.[153] Quantitative analysis is currently being performed to confirm the change of isoflavone

levels with the feeding of aphids (a separate manuscript is in preparation as a part of systematic soybean-aphid interaction study), and preliminary data are consistent with the current result. It is interesting to note that, except SA, Pip and isoflavone-responses are all localized to the feeding sites, and does not occur in the veins, suggesting they are specific local responses.

Conclusions

In this work, we have demonstrated that MSI can be a very useful tool to study the chemical interfaces in plant-pest interactions. Sample preparation is a critical bottleneck because it is almost impossible to cross-section thin leaves in the planar dimension. For the visualization of metabolites across the leaf, the imprinting and fracturing methods have been developed to expose internal metabolites of plant leaves. The two sample preparation methods are complementary to each other. Imprinting is efficient for the analysis of cytoplasmic hydrophilic compounds that can be squeezed out. The greatest advantage of this method is its versatility. It can be applicable to almost any plant leaves, regardless of their size or surface roughness. This method, however, has a critical limitation that membrane-bound molecules, membrane lipids, or other hydrophobic compounds cannot be analyzed. Another limitation is the possible loss of spatial resolution during the imprinting process; however, this seems to be relatively minimal for most tissues. The fracturing method is effective for compounds that are present on the fracture-opened surface including phospholipids and chlorophylls that cannot be analyzed with the imprinting technique. One limitation is

that we cannot control which layer will fracture-open. In the case of rice leaf, the mesophyll layers surrounding the veins seem to be most fragile and vulnerable to fracturing. Another limitation is that it is difficult to apply to a large size leaf, such as soybean leaf, because it tends to partially crack before it is completely dried. In-parallel control experiment with an uninfected leaf would be important in such experiments due to the possible metabolic turnover during sample processing.
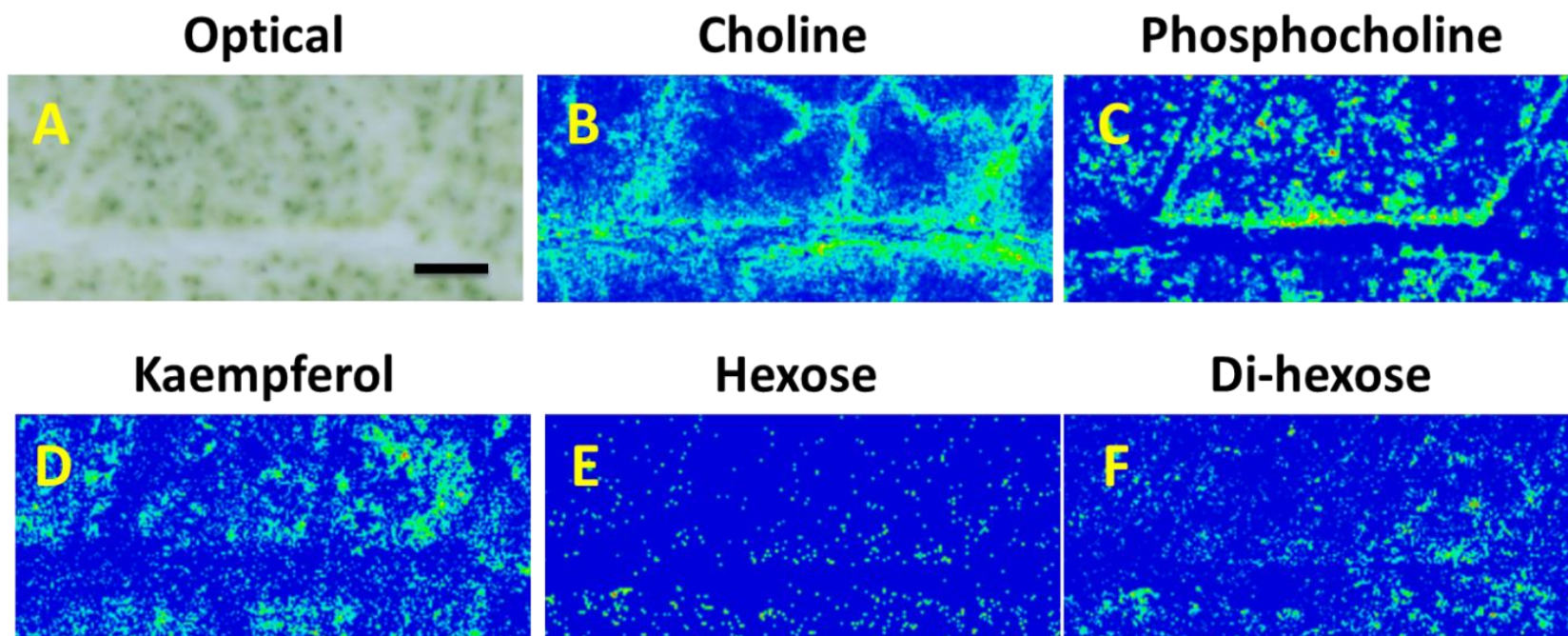
Efficient ionization is often a critical obstacle in MSI depending on what types of analytes are the target molecules of interest. As is well known, the classes of compounds that can be ionized by MALDI are greatly affected by the matrix of choice. We have successfully developed and demonstrated the use of Fe_3O_4 and TiO_2 nanoparticles to efficiently ionize diterpenoid phytoalexin compounds that cannot otherwise be analyzed. We are currently screening various nanoparticles for a wide range of plant metabolites, to further understand ionization mechanisms and apply this to analyses of diverse classes of plant metabolites.

The demonstrated approach of using MSI to study the chemical interfaces of plant-pest interactions has many advantages compared to traditional metabolic profiling, especially in providing precise localization information with high sensitivity. The current work demonstrates the power of this technology and we envision that application of this approach will open new opportunities in plant pathology.

Figure 1. Imprinting MSI with 30 μm raster size. Imprinting MALDI-MSI of a soybean leaf with 30 μm raster step size obtained in positive ion mode. (A) Optical image of imprinted PTFE surface, and chemical images for (B) choline (*m/z* 104.107), (C) phosphocholine (*m/z* 184.073), (D) kaempferol (*m/z* 287.055), (E) hexose (*m/z* 219.027), and (F) di-hexose (*m/z* 381.079). Each analyte ion signal was normalized by the total ion count at each pixel and the max values were arbitrarily adjusted for the best image quality. The scale bar corresponds to 1 mm. Choline and phosphocholine were detected as the molecular cation (M$^+$), kaempferol as the protonated form ([M+H]$^+$), and hexose and di-hexose as the potassiated form ([M+K]$^+$).
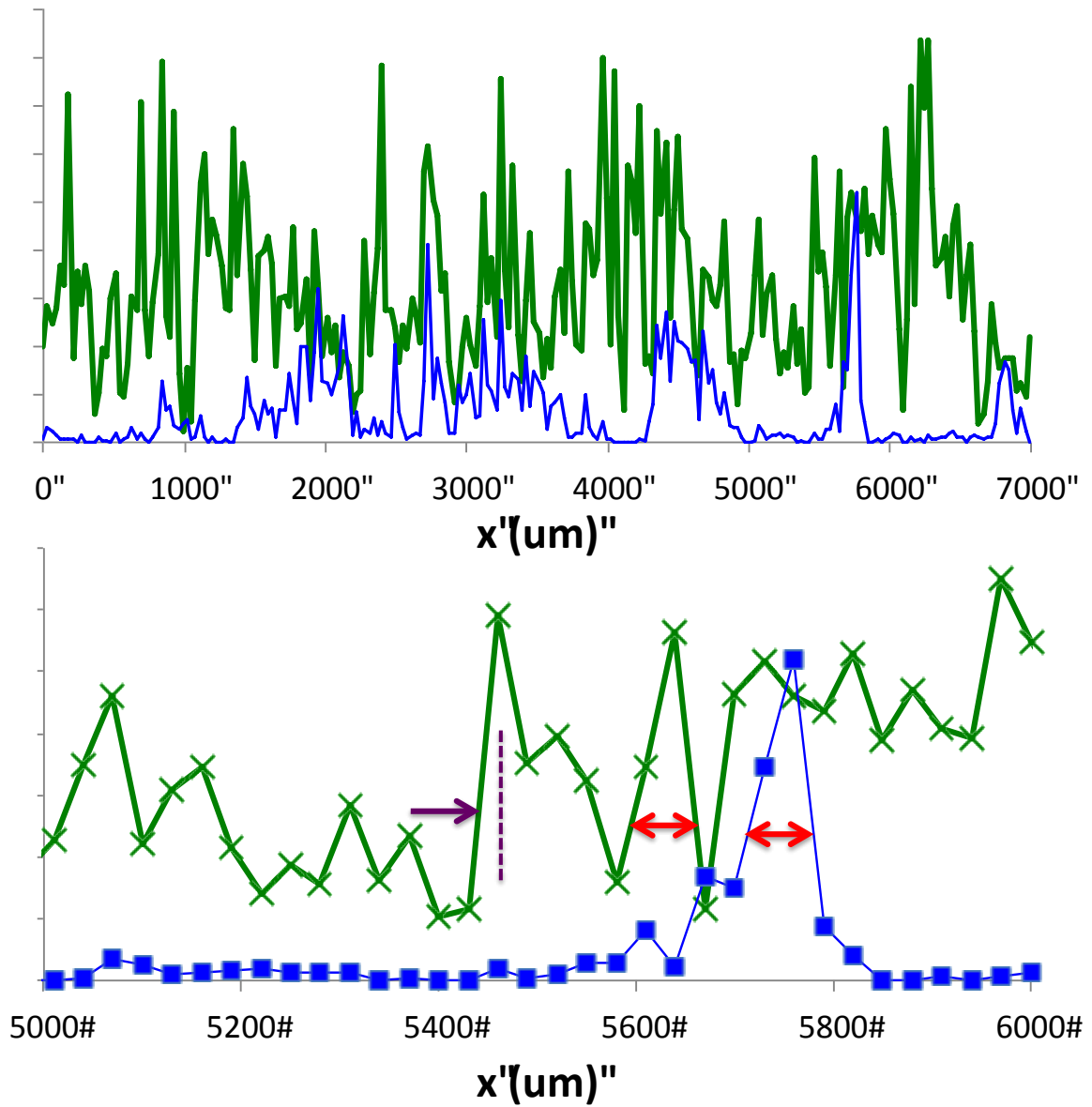
Figure 2. Choline and phosphocholine ion profiles. Top: Selected ion profiles for choline (thick green) and phosphocholine (thin blue) at the y-position of 1,320 μm in Figure 1. **Bottom:** Zoomed-in profiles for x-range of 5,000-6,000 μm.
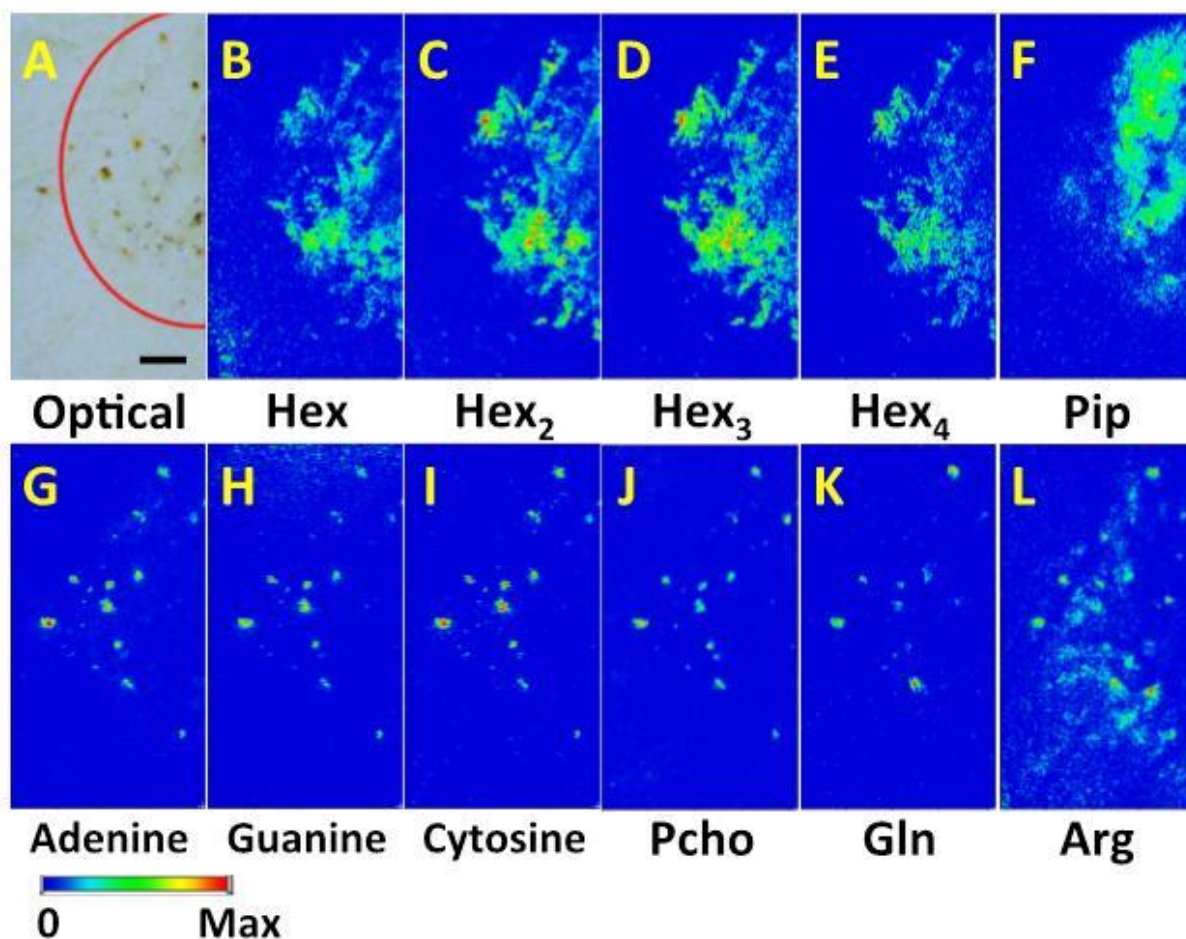
Figure 3. Positive mode MALDI-MSI of imprinted soybean leaf. (A) Optical image of imprinted PTFE surface with the red circle representing the caged aphid region. Chemical images for (B) hexose (Hex, *m/z* 219.027, 1.3ppm; 1×10^{-3}), (C) di-hexose (Hex$_2$, *m/z* 381.079, 0.8ppm; 1×10^{-2}), (D) tri-hexose (Hex$_3$, *m/z* 543.131, 2.8ppm; 5×10^{-3}), (E) tetra-hexose (Hex$_4$, *m/z* 705.184, 2.2ppm; 2×10^{-3}), (F) pipecolic acid (Pip, *m/z* 130.086, 0.2ppm; 7×10^{-4}), (G) adenine (*m/z* 136.062, 1.8ppm; 4×10^{-3}), (H) guanine (*m/z* 152.057, 0.9ppm; 2×10^{3}), (I) cytosine (*m/z* 112.051, 0.1ppm; 8×10^{-4}), (J) phosphocholine (PCho, *m/z* 184.073, 2.6ppm; 4×10^{-2}), (K) glutamine (Gln, *m/z* 147.076, 1.7ppm; 6×10^{-4}), and (L) arginine (Arg, *m/z* 175.119, 2.0ppm; 1×10^{-3}). Each analyte ion signal is normalized by the total ion count at each pixel and the max values used in generating images are shown within parenthesis along with the experimental *m/z* values and mass errors. The amino acids, pipecolic acid, and nucleobases were detected as protonated ions ([M+H]$^+$), the hexoses as potassiated ions ([M+K]$^+$), and phosphocholine as a molecular cation (M$^+$). The scale bar corresponds to 5mm.
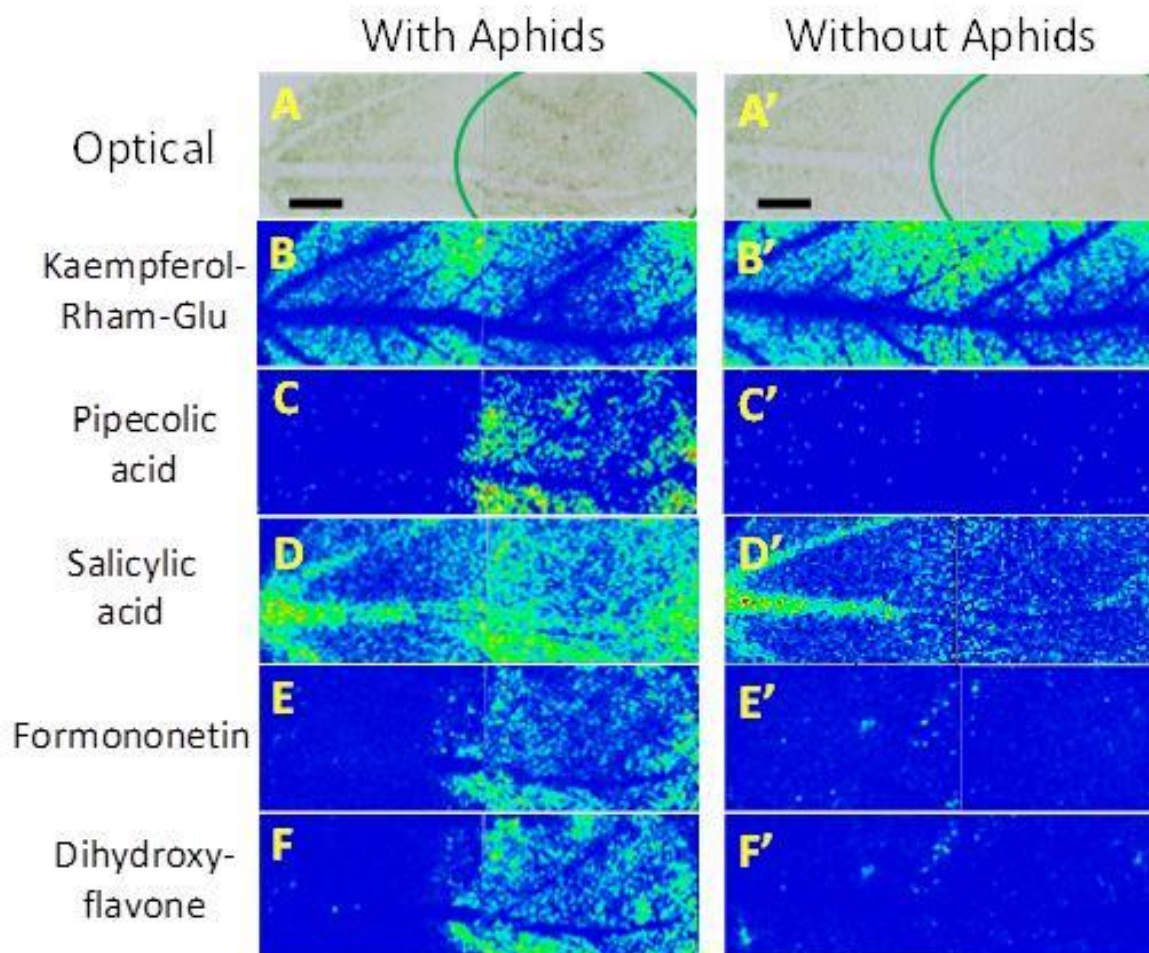
Figure 4. Negative mode MALDI-MSI of imprinted soybean leaf with and without aphids. (A) Optical image of imprinted PTFE surface with the green circle representing caged aphid region. Chemical images for (B) kaempferol-rhamnoside-glucoside (m/z 593.150, 1.9ppm; 1.25×10^{-1}), (C) pipecolic acid (m/z 128.071, 0.4ppm; 8×10^{-4}), (D) salicylic acid (m/z 137.024, 0.2ppm; 4×10^{-3}), (E) formononetin (m/z 267.066, 2.5ppm; 1×10^{-2}), (F) dihydroxyflavone (m/z 253.050, 2.4ppm; 1×10^{-2}); all detected as deprotonated ions ([M-H]$^-$). A' to F' correspond to optical and chemical images in control without aphids. Each imprinted sample was analyzed in two sections to minimize the oxidation of DAN during long acquisition times. The two sections were merged together to create the images. The scale bar corresponds to 5mm.
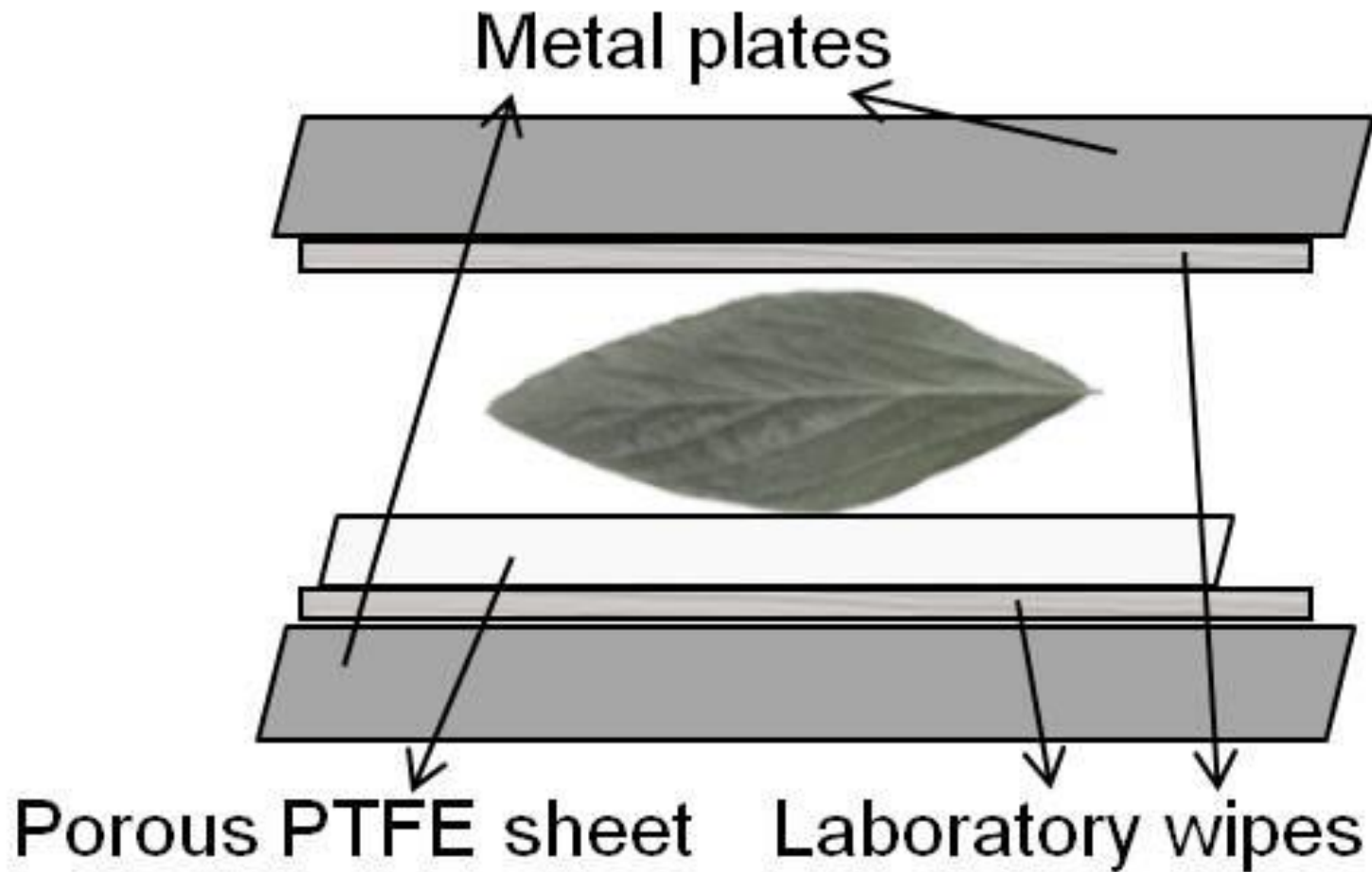
Figure S1. Soybean imprinting setup. The layers from top to bottom are: metal plate, laboratory wipes, soybean leaf, imprinting substrate, laboratory wipes, and metal plate.
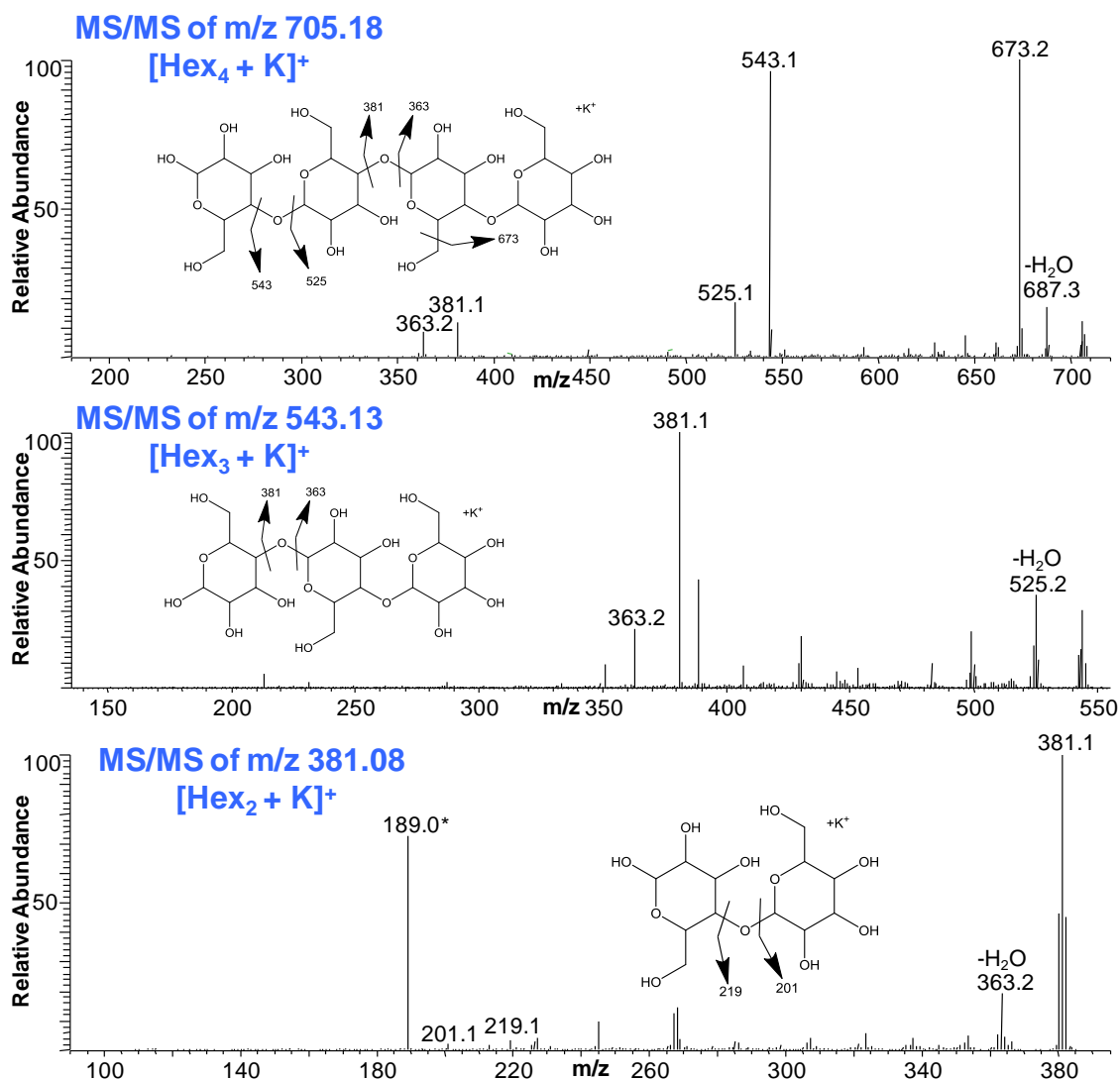
Figure S2. MS/MS of saccharides. MALDI-ion trap MS/MS spectra of three oligosaccharides shown in Figure 3 ([Hex2+K]+, [Hex3+K]+, [Hex4+K]+), obtained directly on the PTFE surface after the imprinting of an aphids-infested soybean leaf. Water-loss (-18 Da), methanol loss (-32 Da), and glucose monomer loss (-$C_6H_{10}O_5$ or –$C_6H_{12}O_6$, -162 Da/-180 Da) are common fragmentations of oligosacchardies. MS/MS of monomer ([Hex+K]+) does not have significant fragmentation and not shown. *: An unknown peak, most likely contamination from MS/MS of a near-by peak within 1 Da.
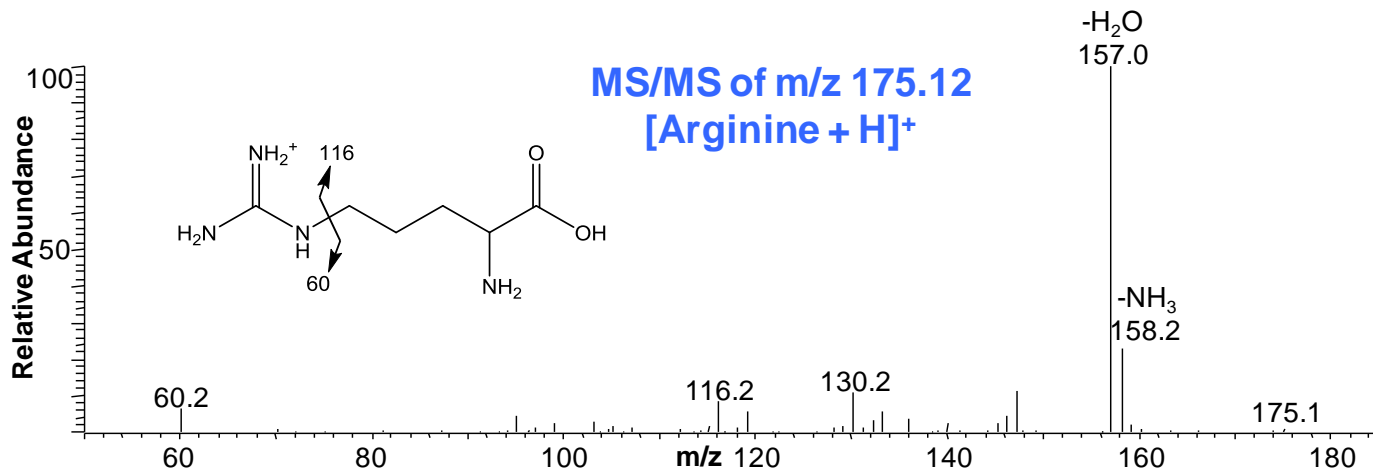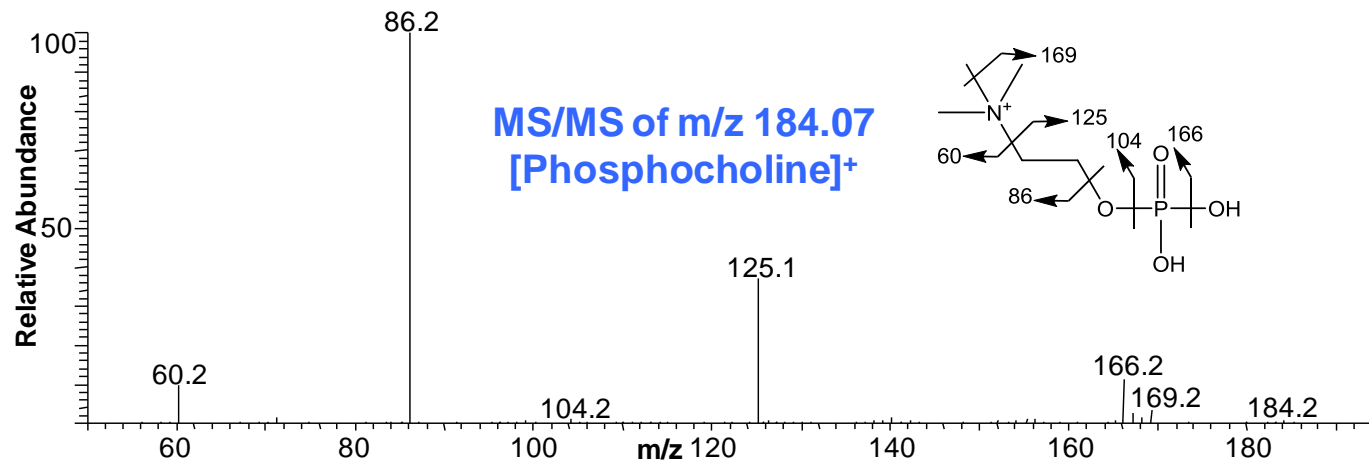
Figure S3. MS/MS of phosphocholine and arginine. MALDI-ion trap MS/MS spectra of phosphocholine and arginine shown in Figure 3, obtained directly on the PTFE surface after the imprinting of an aphids-infested soybean leaf. MS/MS spectra are matching with the standards.
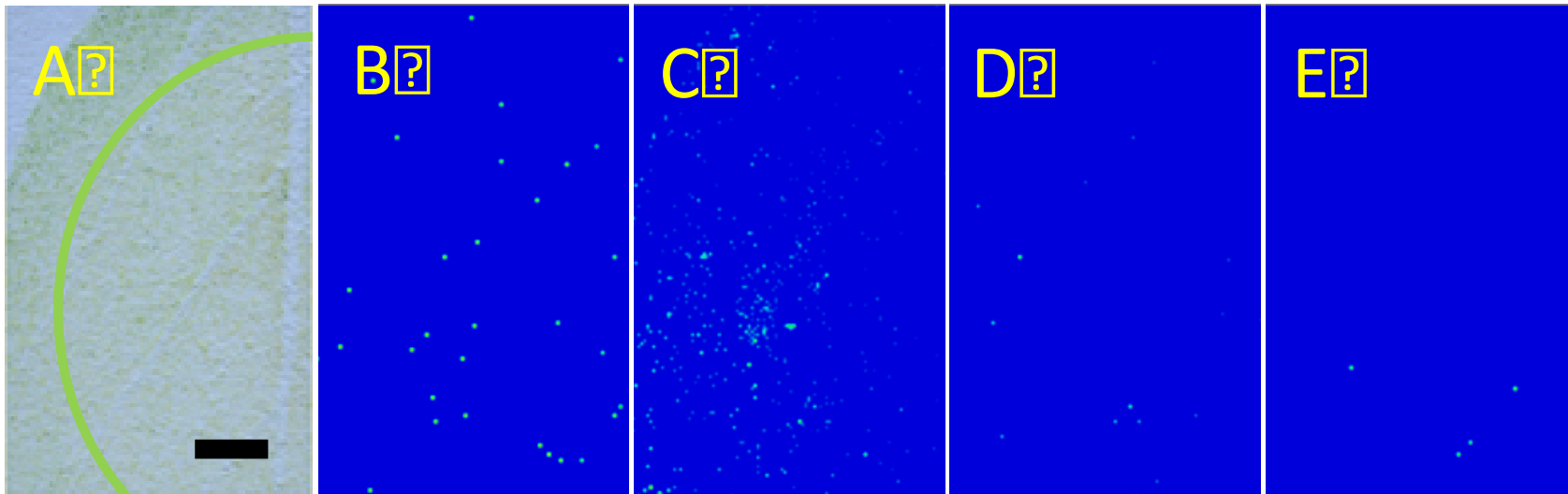
Figure S4. Imprint MSI after washing. Positive mode MALDI-MSI of the PTFE surface after washing and imprinting of an aphids-infested soybean leaf. (A) Optical image of imprinted PTFE surface with the green circle representing caged aphid region. Each analyte ion signal is normalized by the total ion count at each pixel and the molecular images are scaled with the max values shown in the parentheses for (B) hexose (1×10^{-4}), (C) di-hexose (1×10^{-3}), (D) tri-hexose (5×10^{-4}), (E) tetra-hexose (2×10^{-4}); all detected as potassiated ions ($[M+K]^+$). The scale bar corresponds to 5mm.
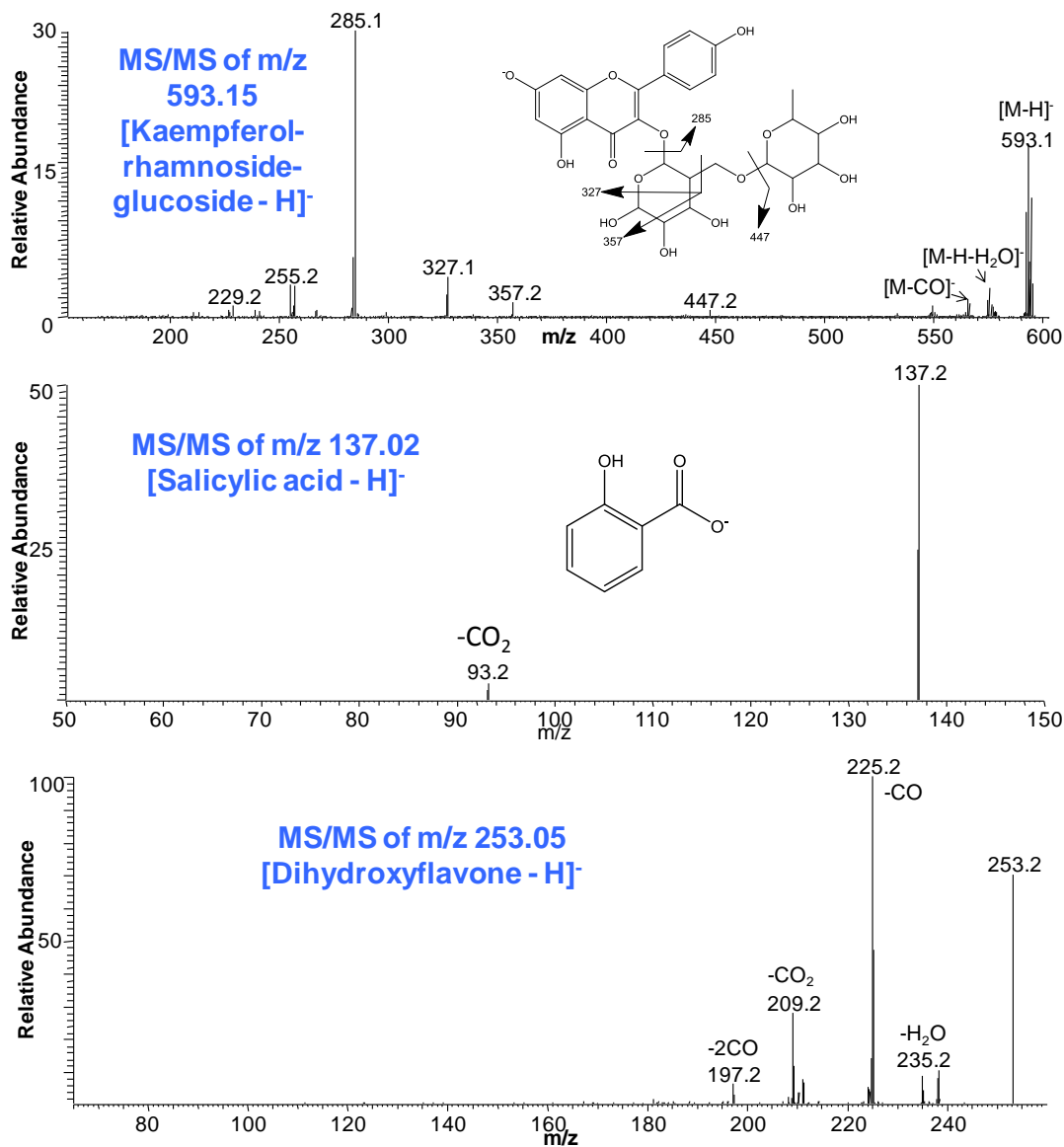
Figure S5. Negative mode MS/MS of soybean metabolites. MALDI-ion trap MS/MS spectra of kaempferol-rhamnoside-glucoside, salicylic acid and dihydroxyflavone, obtained directly on the PTFE surface after the imprinting of an aphids-infested soybean leaf. Kaempferol-rhamnoside-glucoside shows characteristic fragmentation of the side chain sugar units. CO_2-loss is a characteristic fragment ion of salicylic acid according to standard compounds analysis and Metlin. Major products in MS/MS of dihydroxyflavone are matching with those of daidzein in Kang et al. (Rapid Commun Mass Spectrom. 2007;21(6):857-68), and 2',6-dihydroxyflavone in MassBank (www.massbank.jp).

CHAPTER 5

QUANTITATIVE ANALYSIS OF QUATERNARY AMMONIUM COMPOUNDS IN COMMON BEAN SEEDS BY HYDROPHILIC INTERACTION LIQUID CHROMATOGRAPHY MASS SPECTROMETRY

A part of a paper being prepared by Michael Millican and Gwyn Beattie

Adam T. Klein, Michael Millican, Gwyn A. Beattie, and Young-Jin Lee

Abstract

Quaternary ammonium compounds (QACs) are present in plant tissues and can be used by pathogens for nutrition and osmoprotection. Much work has been done to study the uptake and catabolism of QACs in bacteria, such as *pseudomonas syringae* pv. syringae, as well as qualitative analysis of plant tissues for the presence of QACs. Here we developed a method to quantify QACs present in seed exudates of common bean (*Phaseolus vulgaris* L.). The method was targeted toward the detection and quantification of carnitine, which has only been detected in a handful of plant tissues. Quantitation of major plant QACs choline and phosphorylcholine was also performed on the seed exudates. The method has exceptional sensitivity for choline and carnitine, as they were able to be quantified down to 0.05 μM and 0.003 μM, respectively. This is the first report to be able to detect and quantify carnitine released during the germination of common bean seeds.

Introduction

Pseudomonas syringae pv. syringae infect common bean (*Phaseolus vulgaris* L.) and are able to benefit from the chemical environment of the plant. Exogenous compounds including quaternary ammonium compounds (QACs) can provide nutritional benefits as well as serve as an osmoprotectant for the bacteria[154]. Although *P. syringae* are capable of synthesizing osmolytes[155-157], it is much more energetically favorable to uptake QACs from host species[158]. Glycine betaine, which cannot be synthesized by *P. syringae*, is an effective osmoprotectant. It and its precursors in the betaine catabolic pathway can be readily taken up from the environment [159-160]. *P. syringae* has multiple transporters that can uptake choline and its ester phosphorylcholine. Choline and phosphorylcholine are abundant metabolites released in germinating seeds of *Phaseolus vulgaris* and have been shown to support *P. syringae* growth[161-162].

Carnitine is another QAC that certain strains of *P. syringae* are able to uptake and catabolize for osmoprotection[160, 162]. Only about half of the *P. syringae* strains are capable of utilizing carnitine; however, it is a significant promoter of bacterial growth in these situations[162]. *P. syringae* pv. syringae is one strain which can utilize carnitine, but no reports have been made to confirm the presence of carnitine in *Phaseolus vulgaris.* Carnitine is a common metabolite in animal tissue, but is rarely detected in plant tissue[162-163], possibly due to the lack of sensitive quantitative methods for the analysis of QACs in plants[164]. Even choline and phosphorylcholine,

100

which are known to be released during seed germination, have not been quantified during the germination process.

High performance liquid chromatography (HPLC) coupled with mass spectrometry (MS) has been proven to be a valuable technique for analyzing complex biological samples with sensitive detection[165-169]. HPLC simplifies complex samples by separating analytes of interest and minimizing endogenous matrix effects. Hydrophilic interaction liquid chromatography (HILIC) has grown in popularity recently because of its use of MS-compatible solvents and the ability to separate very polar molecules[170-173]. MS detection adds another dimension of separation, improving specificity for the desired analytes and allowing for high-throughput analysis. Quadrupole MS detectors are able to scan quickly, allowing multiple spectra to be collected during the elution of an analyte. This high sampling frequency is important for accurate quantitation of the eluting compounds.

In this study, we use HILIC MS to analyze QACs released by germinating seeds. Our focus is on developing a method sensitive enough to detect and quantify carnitine, as well as quantify the major QACs choline and phosphorylcholine throughout the seed germination process.

Experimental

Materials

Acetonitrile (LCMS Chromasolv), ammonium hydroxide (28%), formic acid (~98%), choline chloride (≥99%), and L-carnitine hydrochloride (≥98%) were

purchased from Sigma Aldrich (St. Louis, MO, USA). Phosphorylcholine chloride

sodium salt hydrate was purchased from Tokyo Chemical Industry Co. Ltd (TCI,

Philadelphia, PA, USA). Water (Optima, LCMS) was purchased from Fisher Scientific

(Waltman, MA, USA).

Mobile phase preparation

A 10 mM ammonium formate solution was prepared by combining 0.61 g of

ammonium hydroxide and 377 µL of formic acid in 1000 mL of water. Additional

formic acid was added to adjust the pH to 3.5.

Mobile phase A was prepared by combining 900 mL of the 10 mM

ammonium formate solution was combined with 100 mL of acetonitrile. Mobile

phase B was prepared by combining 100 mL of the 10 mM ammonium formate

solution with 900 mL acetonitrile. Both mobile phase solutions were sonicated and

degassed prior to use.

Standard preparation

Calibration standards were prepared by individually dissolving choline

chloride, L-carnitine chloride, and phosphorylcholine chloride magnesium salt in

water at 100 mM. Further dilutions were made using mobile phase B to prepare

standards from 0.05 – 10 µM for choline, 0.01 – 0.3 µM for carnitine, and 2.5 – 25 µM

for phosphorylcholine. The standards were prepared as a mixture of the three

metabolites, as described in Table 1, in order to minimize the number of standards needed to construct calibration curves.

Seed exudate germination

Bean seeds (Bush Blue Lake 274) were surface-sterilized by submersion in 70% ethanol for 1 min, in 10% bleach for 1 min, and in sterile water until the bleach smell was no longer detected. Surface-sterilized seeds were then placed in multi-well tissue culture plates, with one seed and 1 ml of sterile water per well. The plates were subjected to a 16-h light/8-h dark photoperiod under plant growth lights, and the seeds were allowed to germinate for 0, 1, 2, or 3 days. At various stages of seed germination and growth, the liquid around the germinated seed was removed and diluted with water to a final volume of 1 ml; this seed solution was washed over the germinated seed 3X to maximize recovery of the seed exudate.

LCMS analysis

Aqueous seed exudate samples were diluted 1:4 with mobile phase B prior to LCMS analysis. Samples were analyzed on an LCMS 2020 (Shimadzu, Columbia, MD, USA) quadrupole mass spectrometer with a Zorbax HILIC Plus, 2.1x50 mm, 3.5 μm column (Agilent, Wilmington, DE, USA). LabSolutions (Shimadzu) software was used to define the acquisition parameters. The autosampler was kept at 4°C. Flow rate was set at 0.5 mL/min with a gradient elution as follows: 90% mobile phase B at 0 minutes, ramped to 75% at 1 minute, 20% at 3 minutes, held at 20% until 5

minutes, and re-equilibrated at 90% from 5 to 7 minutes. Data were collected using positive mode electrospray ionization and two scan events: the first event collected selected ion monitoring (SIM) scans for choline (*m/z* 104.1), carnitine (162.1), and phosphorylcholine (184.1) and the second scan event collected a full MS scan from *m/z* 50-500. The first scan event was used for quantification of the metabolites of interest, whereas the second scan event was used for qualitative analysis of any possible interfering metabolites.

Integration was performed on the SIM scan chromatograms and the peak area was used for quantification. The signal-to-noise ratio (s/n) was also measured for each sample to establish a limit of quantitation (LOQ).

Results and Discussion

Standard analysis

HPLC separation was performed using a gradient and a hydrophilic interaction liquid chromatography (HILIC) column. The metabolites of interest are all small polar molecules, and are difficult to retain on a traditional reverse phase C18 column. Using a HILIC column, with a gradient that starts with a high percentage of organic solvent (90% acetonitrile) and finishes at 20% acetonitrile allows for separation of choline, carnitine, and phosphorylcholine. Figure 1 shows a chromatogram for Standard 4 (5 μM choline, 0.1 μM carnitine, and 15 μM phosphorylcholine). Choline, carnitine and phosphorylcholine have retention times of 1.5, 2.2, and 2.7 minutes, respectively. The three analytes of interest are baseline

resolved and have full width half maximum (FWHM) values of 0.09, 0.12, and 0.19, respectively. The broad peak for phosphorylcholine led to challenges in quantification at very low levels. Since phosphorylcholine is a major QAC released during germination and of secondary interest in this study, however, no further optimization was performed.

Calibration standards were analyzed prior to and after each set of seed exudate samples. Bracket standards used for normalization were also analyzed every ten samples. Although the samples were kept at 4°C during analysis, degradation did appear to occur when comparing the pre-run and post-run standards (data not shown). The degradation was most prevalent for phosphorylcholine, but choline and carnitine were also seen to degrade over long sample runs. As a result, only the initial standards were used for the calibration curves. Figure 2 shows the calibration curves for choline, carnitine, and phosphorylcholine. Choline has the largest dynamic range, with a linear range from 0.05 μM to 10 μM. The linear range for carnitine and phosphorylcholine only covered 1 order of magnitude; however, the range covered was suitable for analysis of seed exudates. The slope, y-intercept, and R^2 values for the calibration curves are listed in Table 2. R^2 values obtained for choline, carnitine, and phosphorylcholine were 0.97, 0.99, and 0.93, respectively.

Seed exudate samples

Exudates from *Phaseolus vulgaris* seeds were analyzed to determine the abundance of choline, carnitine, and phosphorylcholine released during germination. The germination stages span the stages of growth from initial seed imbibition to the formation of lateral roots. Four germination stages were defined based on visual inspection: an imbibed seed with no visible radicle (stage 1), a visible radicle but no visible greening (stage 2), visible greening without visible lateral roots (stage 3), and the presence of visible lateral roots (stage 4).

Figure 3 shows the SIM scans for choline, carnitine, and phosphorylcholine. Retention times match those of the standards. An additional peak was seen in the SIM scan for carnitine (m/z 162.1), but the additional peak elutes at 1.4 min and does not interfere with the elution of the carnitine peak at 2.2 minutes (Figure 3B). Qualitative analysis was performed with the full MS scan of *m/z* 50-500 to confirm no other significant metabolites or sample matrix compounds were co-eluting with the analytes of interest. Seed exudate chromatograms from all germination stages were investigated since metabolites will change during the germination of the seed. Average mass spectra that correlate with the retention times for choline (1.4-1.6 min), carnitine (2.1-2.3 min), and phosphorylcholine (2.6-2.8 min) from a seed exudate at germination stage 3 are shown in Figure 4. Choline is the major compound that elutes at 1.5 min (Figure 4A). The spectrum for the carnitine elution window has a base peak with *m/z* of 175.1, which most likely arises from the amino acid arginine. The peak at *m/z* 59.0 is also relatively abundant and is thought to

come from the aqueous mobile phase solvent, as it increases in abundance with the increase in percent aqueous phase. It is the base peak for the spectra correlating with the elution of phosphorylcholine. Besides the peak at m/z 59.0 from the mobile phase, no other significant peaks are seen to co-elute with phosphorylcholine.

Quantification of seed exudates

Bracket standards were used to correct for possible degradation during sample analysis. Standard 4 was analyzed every ten samples to provide bracket standards. This was performed using the following equation:

$$PA_{corrected} = PA_{measured} * \frac{CalStd}{{1}/{2}\,(BS1 + BS2)}$$

where $PA_{measured}$ is the measured peak area in the sample, CalStd is the peak area of the pre-run calibration Standard 4, and BS1 and BS2 are the bracket standards immediately before and after the sample. This corrects for the degradation of metabolites over time, as it normalizes the measured peak area to the initial standards analyzed at the beginning of the sequence. The concentration of the seed exudate samples was then calculated using the normalized area values and the calibration curves generated from all standard analyses. The dilution factor during sample preparation was also taken into account for the final concentration calculation. The limit of quantitation (LOQ) was determined using the peak area and s/n ratio of the seed exudate samples. The sample that had the largest integrated area with a s/n <10 was used to calculate the LOQ, which was calculated using the raw peak area and calibration curves. For carnitine and phosphorylcholine, the LOQ

was calculated to be 0.003 μM and 3.66 μM respectively. Choline, on the other hand, was not detected with a s/n <10, so the lowest standard concentration analyzed was used for the LOQ, which was 0.05 μM. It should be noted that the LOQ is reported later on at four times higher than mentioned here due to the dilution factor of the samples during analysis.

Figure 5 shows the individual and average QAC concentrations for each seed germination stage. As seen in Figure 5A, carnitine concentration is highest at the onset of germination and decreases as germination progresses. The same trend can be seen for choline (Figure 5B). It should be noted that even though choline and carnitine have similar trends throughout seed germination, the choline concentration is over 100 times greater than that of carnitine. The choline concentration was 24.3 μM at stage 1 of germination, where the carnitine value was only 0.14 μM. The phosphorylcholine abundance remained relatively constant throughout seed germination and was the most abundant of the three metabolites of interest, having a concentration between 25 and 40 μM during germination. The high levels of choline and phosphorylcholine can be explained by the fact they are used in the synthesis of phosphatidylcholines, which are major components of cell membranes, and as such are expected to be highly abundant in all tissues[174]. Carnitine on the other hand, is a transporter of fatty acids across cell membranes during fatty acid metabolism[175].

Conclusions

We were able to develop a quantitative method for the analysis of hydrophilic metabolites important to the nutrition and survival of *Pseudomonas syringae* pv. syringae. The method provides good chromatographic separation of choline, carnitine, and phosphorylcholine, and offers a working concentration range suitable for seed exudate samples. The method is sensitive to small quantities of choline and carnitine, with LOQs of 0.05 and 0.003 μM, respectively.

Acknowledgments

Table 1. Concentration of metabolites in calibration standards.

Calibration standard	Choline concentration (μM)	Carnitine concentration (μM)	Phosphorylcholine concentration (μM)
Standard 0	0	0	0
Standard 1	0.05	0.01	2.5
Standard 2	0.1	0.03	5
Standard 3	1	0.07	10
Standard 4	5	0.1	15
Standard 5	10	0.3	25

Table 2. Calibration curve slope, y-intercept, and R^2 values.

QAC	Slope	Y-intercept	R^2
Choline	357112	116831	0.97
Carnitine	1427184	24497	0.99
Phosphorylcholine	23414	-56175	0.93
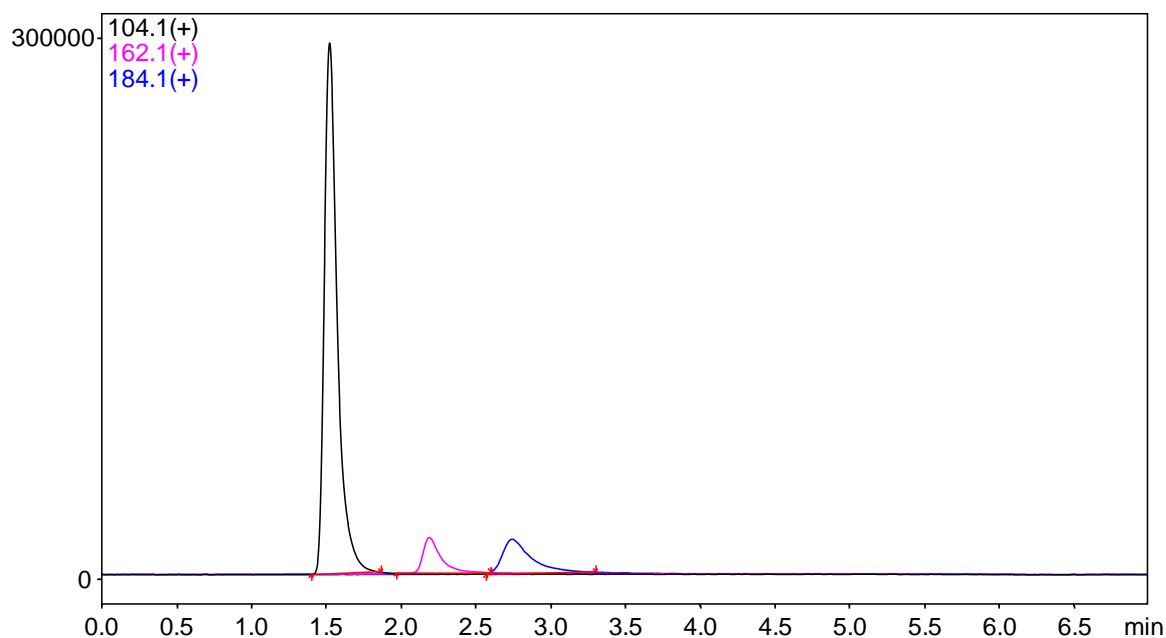
Figure 1. SIM scan LCMS chromatogram. Overlay of the SIM scans of choline, *m/z* 104.1 (black); carnitine, *m/z* 162.1 (pink); and phosphorylcholine, *m/z* 184.1 (blue) from a standard run.
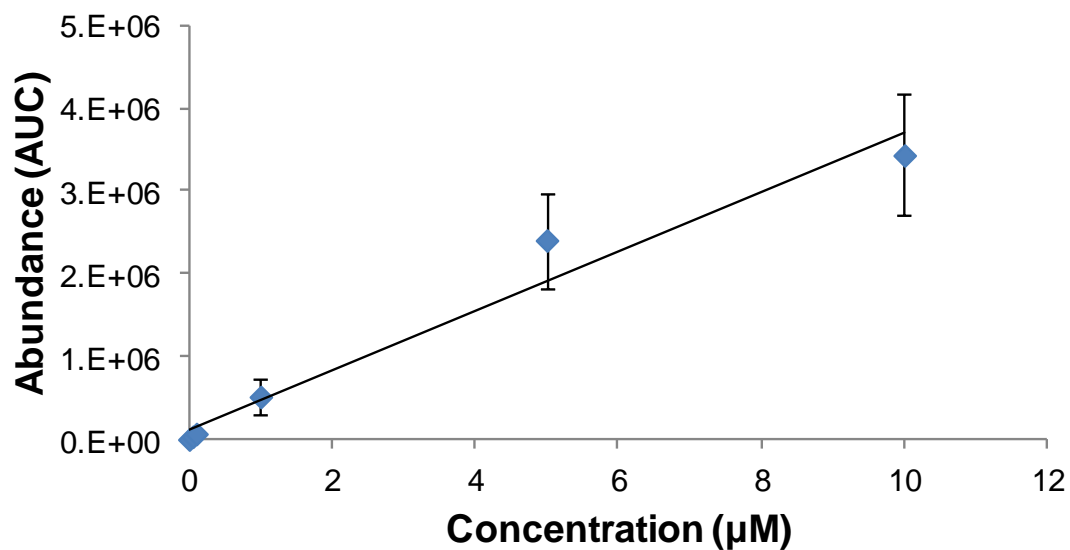
Figure 2A. Choline calibration curve. Calibration curve for choline using standards ranging in concentration from 0.05 µM to 10 µM.

111
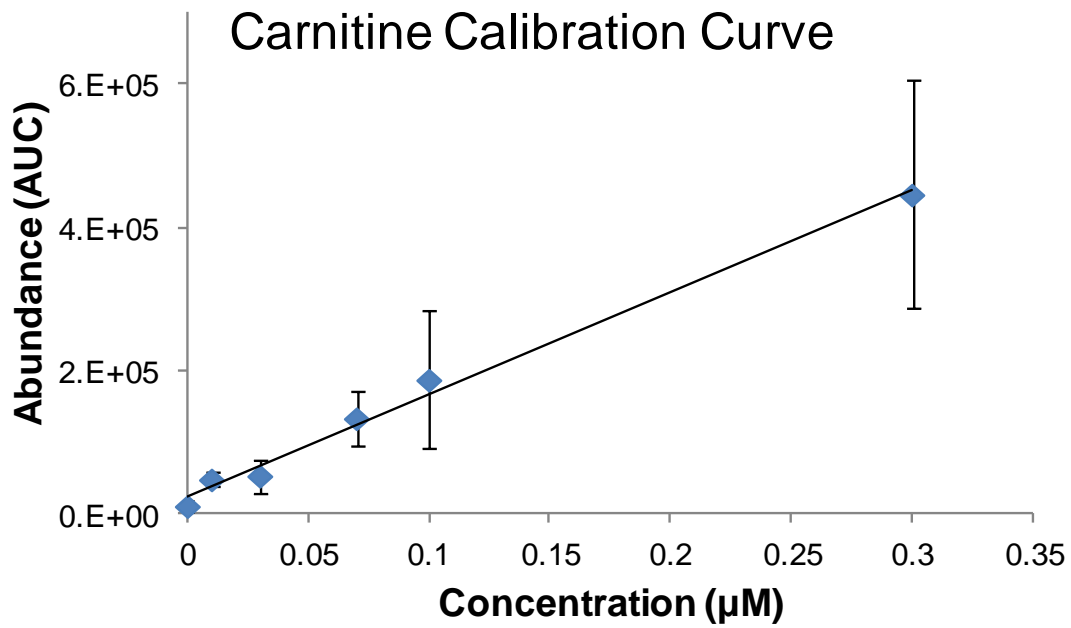
Figure 2B. Carnitine calibration curve. Calibration curve for carnitine using standards ranging in concentration from 0.01 μM to 0.3 μM.
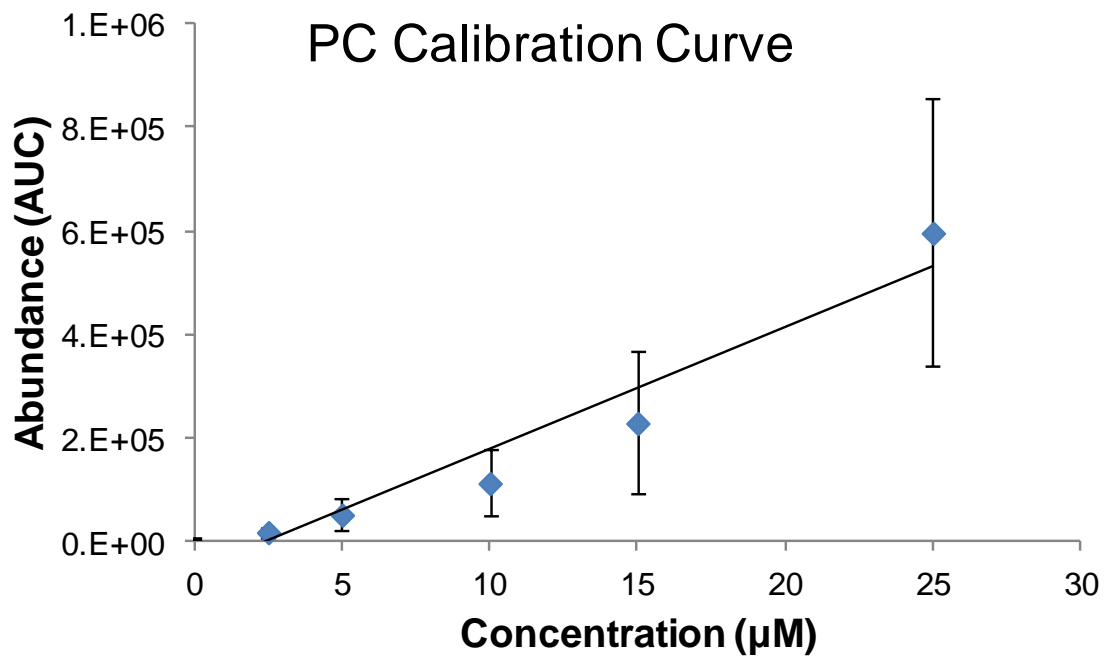
Figure 2C. Phosphorylcholine calibration curve. Calibration curve for phosphorylcholine using standards ranging in concentration from 2.5 μM to 25 μM.
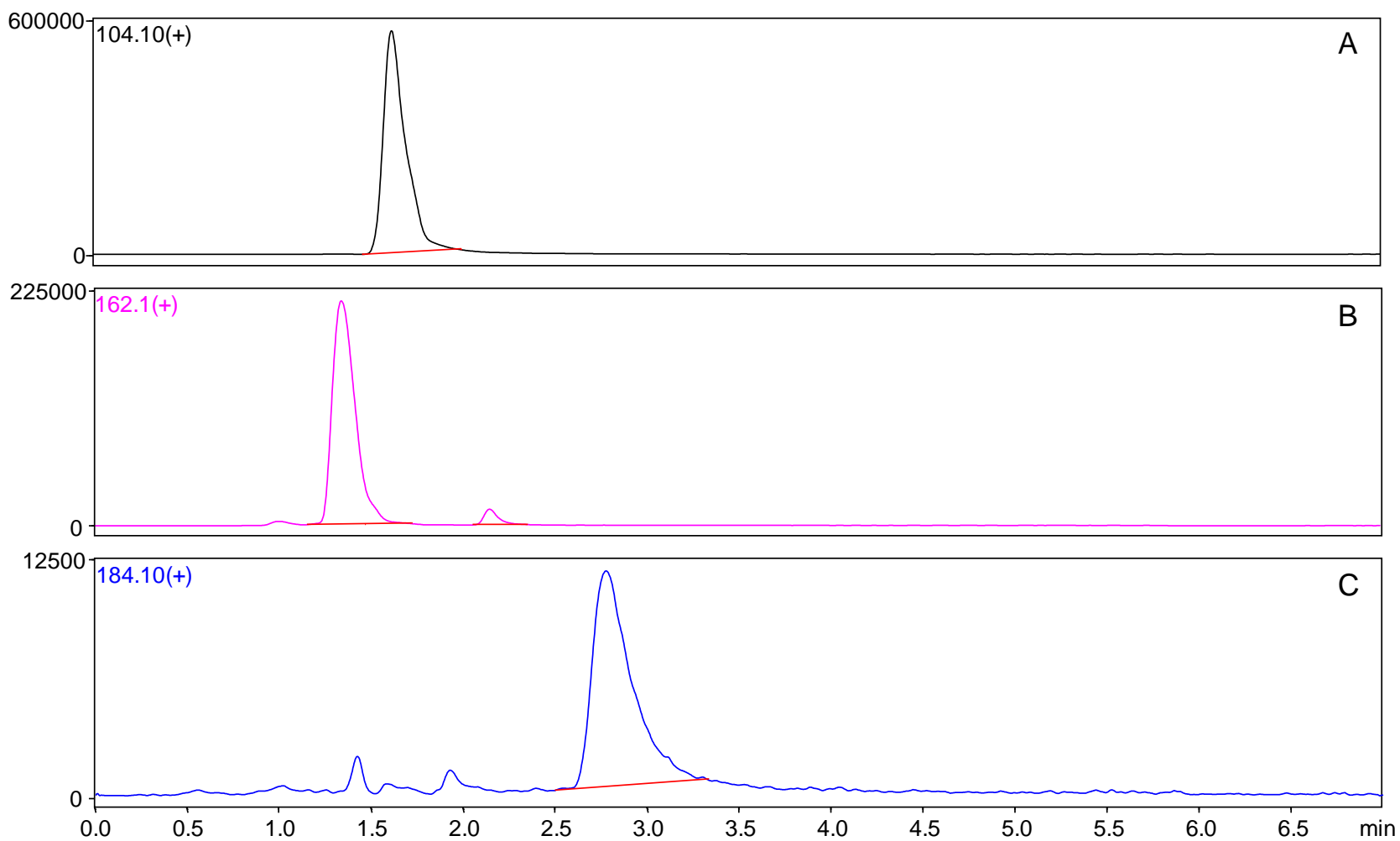
Figure 3. SIM scan of a seed exudate sample. Chromatograms of *m/z* values of (A) 104.1 (choline), (B) 162.1 (carnitine), and (C) 184.1 phosphorylcholine.
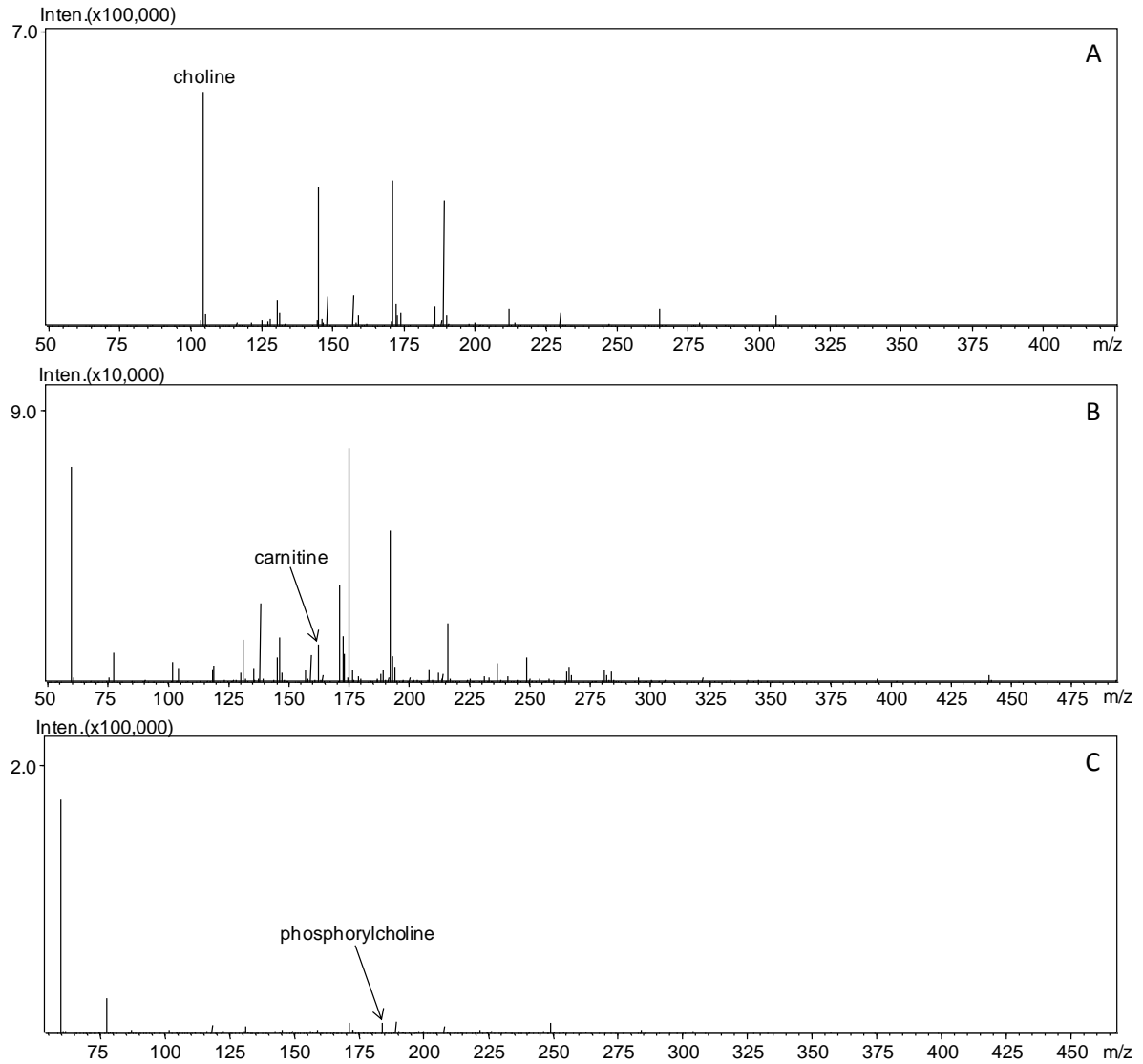
Figure 4. Seed exudate mass spectra. Average spectra at the elution times corresponding with the elution times of (A) choline, (B) carnitine, and (C) phosphorylcholine.
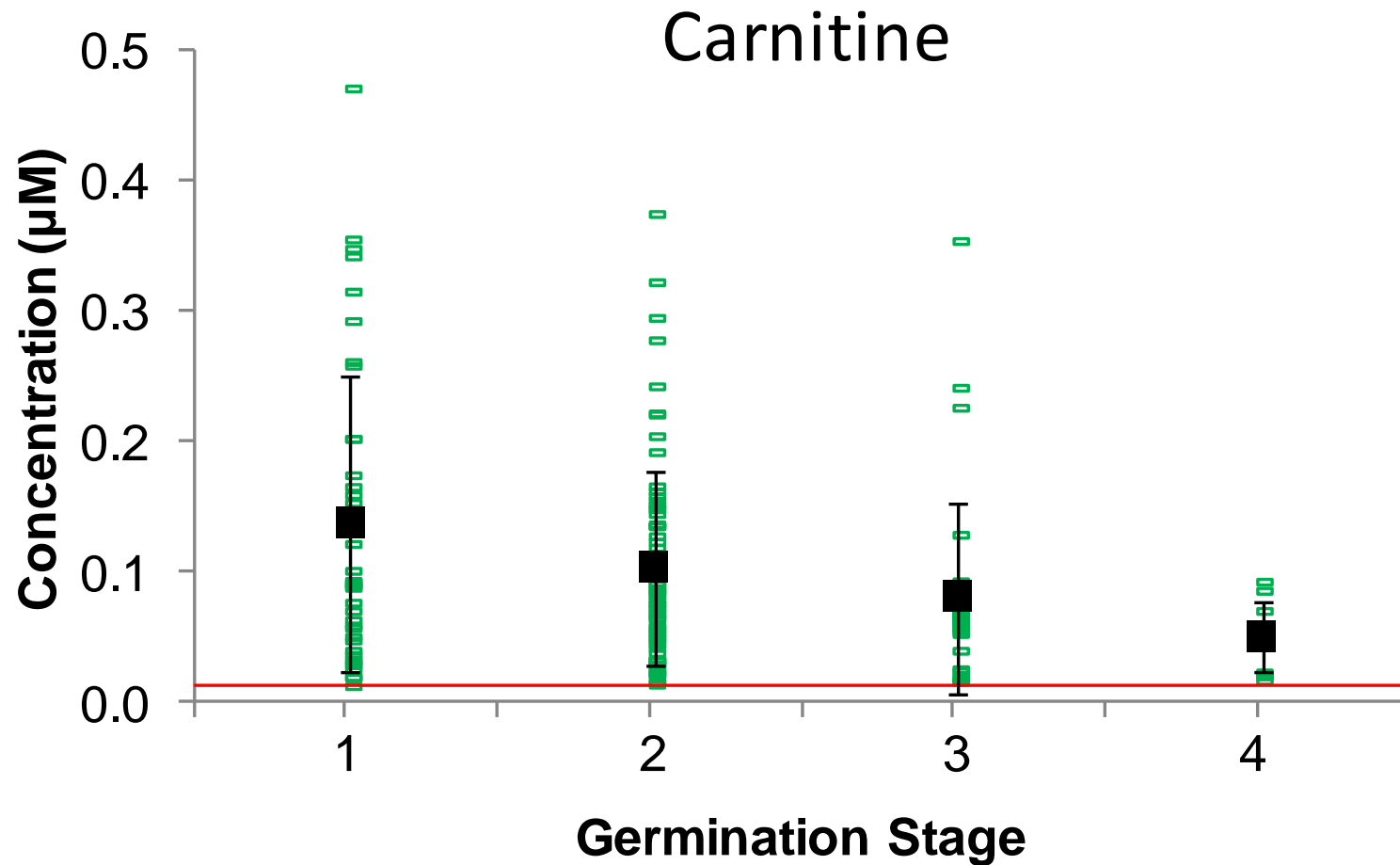
Figure 5A. Abundance of carnitine in seed exudates. Individual data points are shown in green, the average value is in black, and the LOQ is shown as a red line.
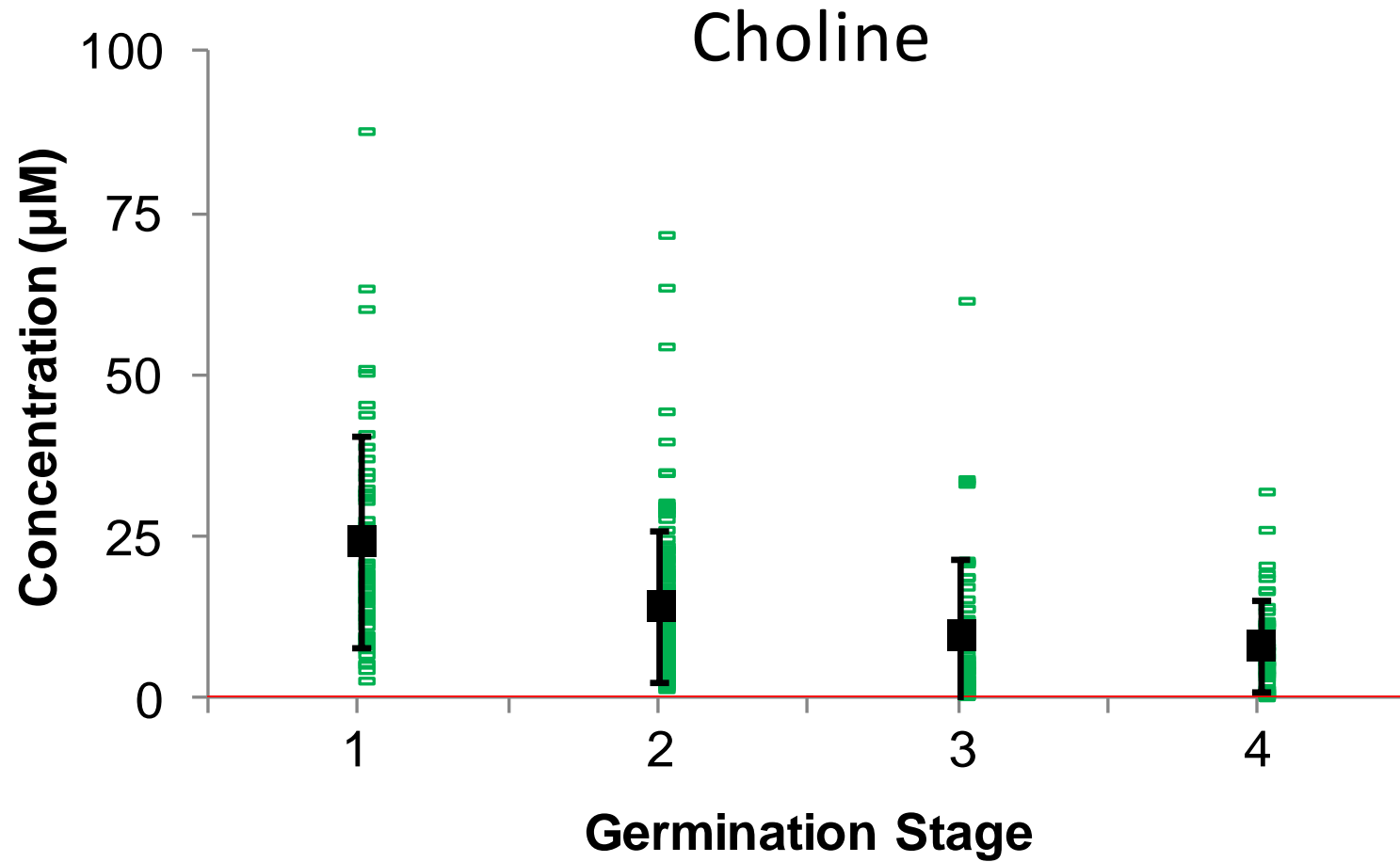
Figure 5B. Abundance of choline in seed exudates. Individual data points are shown in green, the average value is in black, and the LOQ is shown as a red line.
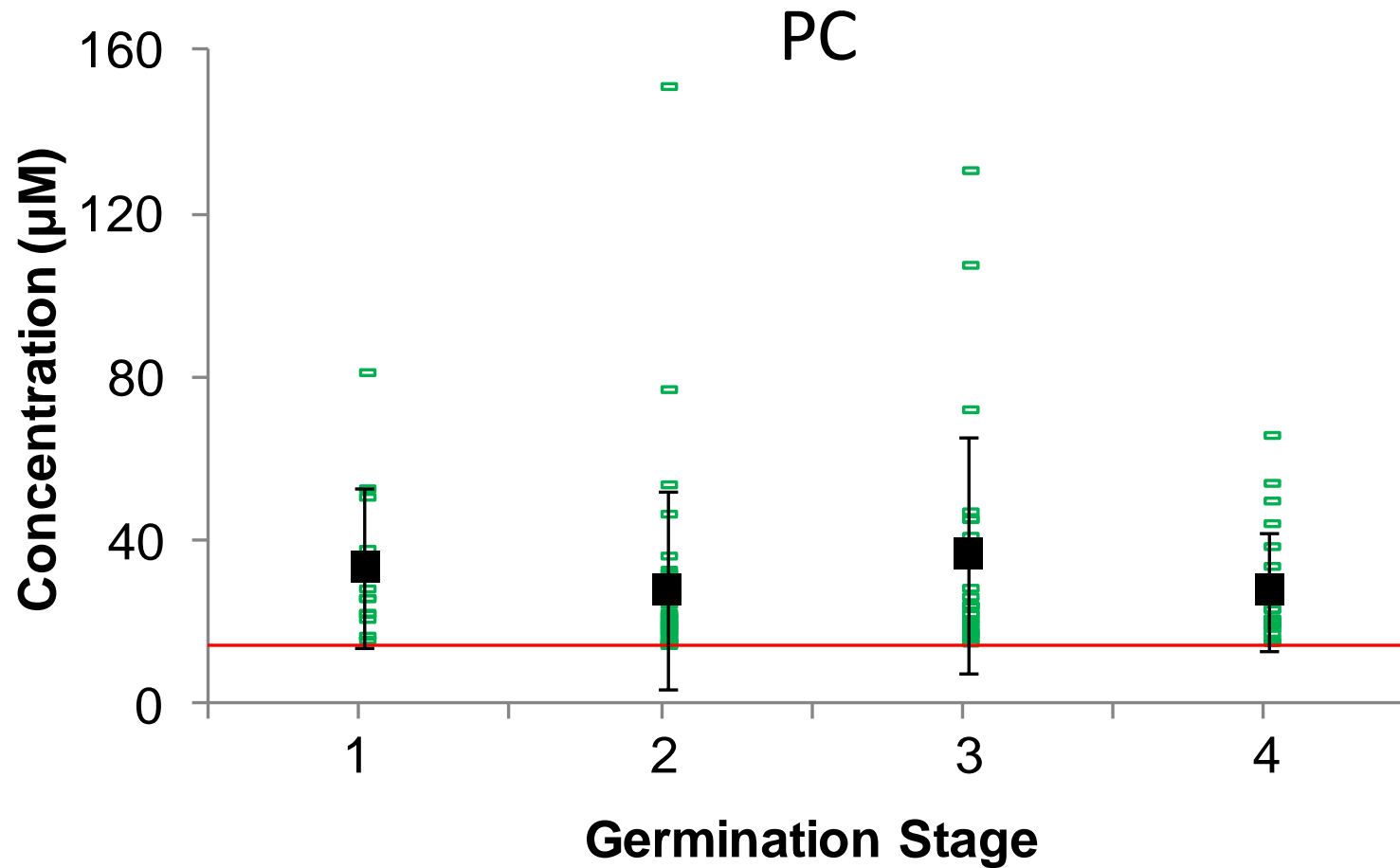
Figure 5C. Abundance of phosphorylcholine in seed exudates. Individual data points are shown in green, the average value is in black, and the LOQ is shown as a red line.

CHAPTER 6

SUMMARY AND OUTLOOK

Summary

The work outlined in this dissertation has sought to expand applications and methodologies for mass spectrometry of biological samples. New methodologies were used to determine when structural distortion is occurring due to cross-linking, as well as increasing the detection efficiency in chemical cross-linking MS. MALDI-MSI was applied to sections of maize leaves to better understand the distribution of photosynthetic related metabolites across a leaf tissue and between genotypes. Soybean-aphid interactions were also investigated with MALDI-MSI by applying imprinting, which is used primarily with DESI-MSI. Finally a quantitative method was developed using HILIC-MS to quantify QACs released during seed germination, with the ability to detect low abundance metabolites.

Outlook

Biological mass spectrometry is a broad field and has a wide variety of applications. Advancements have been made to instrumentation, sample preparation, sample analysis techniques, and software, and improvements continue to be made. More sensitive mass spectrometers, 3-dimensional imaging software, and new matrices and cross-linkers are some examples of the advancement of biological mass spectrometry.

The work presented in this dissertation, although only focused on a small slice of biological mass spectrometry can be beneficial to future experiments. The work in Chapter 2 can be combined with the other techniques, like enrichment strategies, to greatly improve detection efficiency. Applying the strategy mentioned in Chapter 2, where HDX is applied to CXMS experiments, to other proteins can confirm the cross-linking reaction is not distorting the protein conformation.

Understanding metabolite distributions in plants can help bridge the gap between metabolomics and genomics to provide a better overall picture of their biological systems. The MALDI-MSI work in Chapter 3 on maize leaves provides insights into different metabolite distributions between genotypes. This work can be expanded to other genotypes of maize leaves to better understand the differences in metabolite distribution. One genotype of interest is the intercross recombinant inbred Mo17xB73, which has been shown to produce larger ears of corn with bigger kernels than the inbred types. The application of MS/MS imaging of phospholipids also proves to be beneficial, especially if the differences in fatty acid saturation play a role in the construction of the C4 photosynthetic pathway.

The expansion of the plant-pest interaction studies from Chapter 4 is ongoing and has many potential applications. Studies to understand the metabolic distribution in secondary leaves of an infested plant will help identify the global changes in soybean plants during infestation. Work is also being performed to determine how defense compounds get distributed throughout the plant, and the affects increased levels of isoflavones have on infestation. This work can also be applied to other biological tissues that are prone to plant infestation.

Investigating additional QACs, as discussed in Chapter 5, can provide even more insight into how bacteria benefit from host plants. Using the developed method can help detect low abundance metabolites that can't be detected using other methods. As carnitine has only been reported in a few plant tissues, this quantitative method could be applied to other tissue types to possibly detect carnitine in the tissue.

The objective of this work was to advance methodology and applications for biological mass spectrometry. This work is a small part of biological mass spectrometry, however, these applications can be expanded to further the work mentioned here, or used to explore other biological systems.

REFERENCES

[1]   E. J. Boland, P. K. Nair, D. D. Lemon, J. S. Olson, J. D. Hellums, *An in vitro capillary system for studies on microcirculatory O2 transport, Vol. 62*, **1987**.

[2]   J. M. Berg, J. L. Tymoczko, L. Stryer, DNA polymerases require a template and a primer, **2002**.

[3]   J. A. Sibley, A. L. Lehninger, Determination of aldolase in animal tissues, *Journal of Biological Chemistry* **1949**, *177*, 859.

[4]   H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The protein data bank, *Nucleic acids research* **2000**, *28*, 235.

[5]   J. Kuriyan, S. Wilz, M. Karplus, G. A. Petsko, X-ray structure and refinement of carbon-monoxy (Fe II)-myoglobin at 1.5 Å resolution, *Journal of Molecular Biology* **1986**, *192*, 133.

[6]   B. J. Wylie, L. J. Sperling, A. J. Nieuwkoop, W. T. Franks, E. Oldfield, C. M. Rienstra, Ultrahigh resolution protein structures using NMR chemical shift tensors, *Proceedings of the National Academy of Sciences* **2011**, *108*, 16974.

[7]   C. A. Scarff, K. Thalassinos, G. R. Hilton, J. H. Scrivens, Travelling wave ion mobility mass spectrometry studies of protein structure: biological significance and comparison with X-ray crystallography and nuclear magnetic resonance spectroscopy measurements, *Rapid Communications in Mass Spectrometry* **2008**, *22*, 3297.

[8]   J. G. Kiselar, S. Maleknia, M. Sullivan, K. Downard, M. Chance, Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry, *International journal of radiation biology* **2002**, *78*, 101.

[9]   Z. Zhang, D. L. Smith, Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation, *Protein Science* **1993**, *2*, 522.

[10]   M. M. Young, N. Tang, J. C. Hempel, C. M. Oshiro, E. W. Taylor, I. D. Kuntz, B. W. Gibson, G. Dollinger, High throughput protein fold identification by using

experimental constraints derived from intramolecular cross-links and mass spectrometry, *Proceedings of the National Academy of Sciences* **2000**, *97*, 5802.

[11]   L. Konermann, X. Tong, Y. Pan, Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches, *Journal of Mass Spectrometry* **2008**, *43*, 1021.

[12]   A. Sinz, Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein–protein interactions, *Mass Spectrometry Reviews* **2006**, *25*, 663.

[13]   B. Bogdanov, R. D. Smith, Proteomics by FTICR mass spectrometry: top down and bottom up, *Mass Spectrometry Reviews* **2005**, *24*, 168.

[14]   T. W. Jaskolla, M. Karas, Compelling evidence for lucky survivor and gas phase protonation: the unified MALDI analyte protonation mechanism, *Journal of The American Society for Mass Spectrometry* **2011**, *22*, 976.

[15]   L. M. Mikesh, B. Ueberheide, A. Chi, J. J. Coon, J. E. Syka, J. Shabanowitz, D. F. Hunt, The utility of ETD mass spectrometry in proteomic analysis, *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2006**, *1764*, 1811.

[16]   V. H. Wysocki, G. Tsaprailis, L. L. Smith, L. A. Breci, Mobile and localized protons: a framework for understanding peptide dissociation, *Journal of Mass Spectrometry* **2000**, *35*, 1399.

[17]   J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, D. F. Hunt, Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry, *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 9528.

[18]   L. Elviri, *ETD and ECD mass spectrometry fragmentation for the characterization of protein post translational modifications*, INTECH Open Access Publisher, **2012**.

[19]   O. Rinner, J. Seebacher, T. Walzthoeni, L. Mueller, M. Beck, A. Schmidt, M. Mueller, R. Aebersold, Identification of cross-linked peptides from large sequence databases, *Nature methods* **2008**, *5*, 315.

[20]   A. Leitner, T. Walzthoeni, A. Kahraman, F. Herzog, O. Rinner, M. Beck, R. Aebersold, Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics, *Molecular & Cellular Proteomics* **2010**, *9*, 1634.

[21]   M. A. Nessen, G. Kramer, J. Back, J. M. Baskin, L. E. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, Selective enrichment of azide-containing peptides from complex mixtures, *Journal of Proteome Research* **2009**, *8*, 3702.

[22]   E. V. Petrotchenko, J. J. Serpa, C. H. Borchers, An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics, *Molecular & Cellular Proteomics* **2011**, *10*, M110. 001420.

[23]   A. Sinz, K. Wang, Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: application to nebulin-calmodulin complexes, *Biochemistry* **2001**, *40*, 7903.

[24]   E. V. Petrotchenko, V. K. Olkhovik, C. H. Borchers, Isotopically coded cleavable cross-linker for studying protein-protein interaction and protein complexes, *Molecular & Cellular Proteomics* **2005**, *4*, 1167.

[25]   K. L. Bennett, M. Kussmann, M. Mikkelsen, P. Roepstorff, P. Björk, M. Godzwon, P. Sörensen, Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—A novel approach to assess intermolecular protein contacts, *Protein Science* **2000**, *9*, 1503.

[26]   P. Singh, S. A. Shaffer, A. Scherl, C. Holman, R. A. Pfuetzner, T. J. Larson Freeman, S. I. Miller, P. Hernandez, R. D. Appel, D. R. Goodlett, Characterization of protein cross-links via mass spectrometry and an open-modification search strategy, *Analytical Chemistry* **2008**, *80*, 8799.

[27]   O. Rinner, J. Seebacher, T. Walzthoeni, L. Mueller, M. Beck, A. Schmidt, M. Mueller, R. Aebersold, Identification of cross-linked peptides from large sequence databases, *Nat Meth* **2008**, *5*, 315.

[28]   Y. J. Lee, L. L. Lackner, J. M. Nunnari, B. S. Phinney, Shotgun Cross-Linking Analysis for Studying Quaternary and Tertiary Protein Structures, *Journal of Proteome Research* **2007**, *6*, 3908.

123

[29] P. Singh, A. Panchaud, D. R. Goodlett, Chemical Cross-Linking and Mass Spectrometry As a Low-Resolution Protein Structure Determination Technique, *Analytical Chemistry* **2010**, *82*, 2636.

[30] M. Huse, J. Kuriyan, The conformational plasticity of protein kinases, *Cell* **2002**, *109*, 275.

[31] D. Houde, Y. Peng, S. A. Berkowitz, J. R. Engen, Post-translational modifications differentially affect IgG1 conformation and receptor binding, *Molecular & Cellular Proteomics* **2010**, *9*, 1716.

[32] B. E. Finn, J. Evenäs, T. Drakenberg, J. P. Waltho, E. Thulin, S. Forsén, Calcium-induced structural changes and domain autonomy in calmodulin, *Nature Structural & Molecular Biology* **1995**, *2*, 777.

[33] P. A. Ndakidemi, F. D. Dakora, Legume seed flavonoids and nitrogenous metabolites as signals and protectants in early seedling development, *Functional Plant Biology* **2003**, *30*, 729.

[34] C. E. Jeffree, 2 The fine structure of the plant cuticle, *Annual Plant Reviews, Biology of the plant cuticle* **2008**, *23*, 11.

[35] E. Melin, V. Rama Das, Influence of Root-Metabolites on the Growth of Tree Mycorrhizal Fungi, *Physiologia Plantarum* **1954**, *7*, 851.

[36] A. Besserer, V. Puech-Pagès, P. Kiefer, V. Gomez-Roldan, A. Jauneau, S. Roy, J.-C. Portais, C. Roux, G. Bécard, N. Séjalon-Delmas, Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria, *PLoS Biol* **2006**, *4*, e226.

[37] P. D. Coley, J. P. Bryant, F. S. Chapin III, Resource availability and plant antiherbivore defense, *Science(Washington)* **1985**, *230*, 895.

[38] A. R. Korte, M. D. Yandeau-Nelson, B. J. Nikolau, Y. J. Lee, Subcellular-level resolution MALDI-MS imaging of maize leaf metabolites by MALDI-linear ion trap-Orbitrap mass spectrometer, *Analytical and Bioanalytical Chemistry* **2015**, *407*, 2301.

[39] T. Taji, C. Ohsumi, S. Iuchi, M. Seki, M. Kasuga, M. Kobayashi, K. Yamaguchi-Shinozaki, K. Shinozaki, Important roles of drought-and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana, *The Plant Journal* **2002**, *29*, 417.

[40] D. Bartels, R. Sunkar, Drought and salt tolerance in plants, *Critical reviews in plant sciences* **2005**, *24*, 23.

[41] J. Krasensky, C. Jonak, Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks, *Journal of Experimental Botany* **2012**, *63*, 1593.

[42] C. Guy, F. Kaplan, J. Kopka, J. Selbig, D. K. Hincha, Metabolomics of temperature stress, *Physiologia Plantarum* **2008**, *132*, 220.

[43] K. Dettmer, P. A. Aronov, B. D. Hammock, Mass spectrometry-based metabolomics, *Mass Spectrometry Reviews* **2007**, *26*, 51.

[44] M. E. Kurczy, T. R. Northen, S. A. Trauger, G. Siuzdak, in *Mass Spectrometry Imaging of Small Molecules*, Springer, **2015**, pp. 141.

[45] D. Wang, S. Bodovitz, Single cell analysis: the new frontier in 'omics', *Trends in biotechnology* **2010**, *28*, 281.

[46] T. K. Dutta, S. Harayama, Time-of-flight mass spectrometric analysis of high-molecular-weight alkanes in crude oil by silver nitrate chemical ionization after laser desorption, *Analytical Chemistry* **2001**, *73*, 864.

[47] M. S. Kahr, C. L. Wilkins, Silver nitrate chemical ionization for analysis of hydrocarbon polymers by laser desorption Fourier transform mass spectrometry, *Journal of The American Society for Mass Spectrometry* **1993**, *4*, 453.

[48] A. C. Peterson, J. D. Russell, D. J. Bailey, M. S. Westphall, J. J. Coon, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics, *Molecular & Cellular Proteomics* **2012**, *11*, 1475.

[49] M. Mann, R. C. Hendrickson, A. Pandey, Analysis of proteins and proteomes by mass spectrometry, *Annual review of biochemistry* **2001**, *70*, 437.

[50]   H.-Y. J. Wang, S. N. Jackson, J. McEuen, A. S. Woods, Localization and analyses of small drug molecules in rat brain tissue sections, *Analytical Chemistry* **2005**, *77*, 6682.

[51]   M. D. Mohr, K. OlafBörnsen, H. M. Widmer, Matrix-assisted laser desorption/ionization mass spectrometry: Improved matrix for oligosaccharides, *Rapid Communications in Mass Spectrometry* **1995**, *9*, 809.

[52]   C. D. Cerruti, F. Benabdellah, O. Laprévote, D. Touboul, A. Brunelle, MALDI imaging and structural analysis of rat brain lipid negative ions with 9-aminoacridine matrix, *Analytical Chemistry* **2012**, *84*, 2164.

[53]   A. l. Thomas, J. L. Charbonneau, E. Fournaise, P. Chaurand, Sublimation of new matrix candidates for high spatial resolution imaging mass spectrometry of lipids: enhanced information in both positive and negative polarities after 1, 5-diaminonapthalene deposition, *Analytical Chemistry* **2012**, *84*, 2048.

[54]   A. R. Korte, Y. J. Lee, MALDI-MS analysis and imaging of small molecule metabolites with 1, 5-diaminonaphthalene (DAN), *Journal of Mass Spectrometry* **2014**, *49*, 737.

[55]   A. T. Klein, G. B. Yagnik, J. D. Hohenstein, Z. Ji, J. Zi, M. D. Reichert, G. C. MacIntosh, B. Yang, R. J. Peters, J. Vela, Investigation of the Chemical Interface in the Soybean-Aphid and Rice-Bacteria Interactions Using MALDI-Mass Spectrometry Imaging, *Analytical Chemistry* **2015**.

[56]   J. H. Jungmann, R. M. Heeren, Emerging technologies in mass spectrometry imaging, *Journal of Proteomics* **2012**, *75*, 5077.

[57]   E. R. A. van Hove, D. F. Smith, R. M. Heeren, A concise review of mass spectrometry imaging, *Journal of Chromatography A* **2010**, *1217*, 3946.

[58]   A. Svatoš, Mass spectrometric imaging of small molecules, *Trends in biotechnology* **2010**, *28*, 425.

[59]   N. Zaima, T. Hayasaka, N. Goto-Inoue, M. Setou, Matrix-assisted laser desorption/ionization imaging mass spectrometry, *International journal of molecular sciences* **2010**, *11*, 5040.

[60] Y. J. Lee, D. C. Perdian, Z. Song, E. S. Yeung, B. J. Nikolau, Use of mass spectrometry for imaging metabolites in plants, *The Plant Journal* **2012**, *70*, 81.

[61] P. M. Angel, R. M. Caprioli, Matrix-assisted laser desorption ionization imaging mass spectrometry: in situ molecular mapping, *Biochemistry* **2013**, *52*, 3818.

[62] J. H. Jun, Z. Song, Z. Liu, B. J. Nikolau, E. S. Yeung, Y. J. Lee, High-Spatial and High-Mass Resolution Imaging of Surface Metabolites of Arabidopsis thaliana by Laser Desorption-Ionization Mass Spectrometry Using Colloidal Silver, *Analytical Chemistry* **2010**, *82*, 3255.

[63] S. Robinson, K. Warburton, M. Seymour, M. Clench, J. Thomas-Oates, Localization of water-soluble carbohydrates in wheat stems using imaging matrix-assisted laser desorption ionization mass spectrometry, *New Phytologist* **2007**, *173*, 438.

[64] V. Vrkoslav, A. Muck, J. Cvačka, A. Svatoš, MALDI imaging of neutral cuticular lipids in insects and plants, *Journal of The American Society for Mass Spectrometry* **2010**, *21*, 220.

[65] C. Li, Z. Wang, A. D. Jones, Chemical imaging of trichome specialized metabolites using contact printing and laser desorption/ionization mass spectrometry, *Analytical and Bioanalytical Chemistry* **2014**, *406*, 171.

[66] B. C. Bohrer, S. I. Merenbloom, S. L. Koeniger, A. E. Hilderbrand, D. E. Clemmer, Biomolecule Analysis by Ion Mobility Spectrometry, *Annual Review of Analytical Chemistry* **2008**, *1*, 293.

[67] H. Wei, J. Mo, L. Tao, R. J. Russell, A. A. Tymiak, G. Chen, R. E. Iacob, J. R. Engen, Hydrogen/deuterium exchange mass spectrometry for probing higher order structure of protein therapeutics: methodology and applications, *Drug discovery today* **2014**, *19*, 95.

[68] L. Konermann, J. Pan, Y.-H. Liu, Hydrogen exchange mass spectrometry for studying protein structure and dynamics, *Chemical Society Reviews* **2011**, *40*, 1224.

[69] G. F. Pirrone, R. E. Iacob, J. R. Engen, Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014, *Analytical Chemistry* **2015**, *87*, 99.

[70] R. C. Huang, G. Chen, Higher order structure characterization of protein therapeutics by hydrogen/deuterium exchange mass spectrometry, *Analytical and Bioanalytical Chemistry* **2014**, *406*, 6541.

[71] K. D. Rand, S. D. Pringle, M. Morris, J. R. Engen, J. M. Brown, ETD in a traveling wave ion guide at tuned Z-spray ion source conditions allows for site-specific hydrogen/deuterium exchange measurements, *Journal of The American Society for Mass Spectrometry* **2011**, *22*, 1784.

[72] R. E. Iacob, G. Chen, J. Ahn, S. Houel, H. Wei, J. Mo, L. Tao, D. Cohen, D. Xie, Z. Lin, The Influence of Adnectin Binding on the Extracellular Domain of Epidermal Growth Factor Receptor, *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 2093.

[73] J. Pan, J. Han, C. H. Borchers, L. Konermann, Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein, *Journal of the American Chemical Society* **2009**, *131*, 12801.

[74] B. C. Gau, J. S. Sharp, D. L. Rempel, M. L. Gross, Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding, *Analytical Chemistry* **2009**, *81*, 6563.

[75] L. Konermann, B. B. Stocks, Y. Pan, X. Tong, Mass spectrometry combined with oxidative labeling for exploring protein structure and folding, *Mass Spectrometry Reviews* **2010**, *29*, 651.

[76] Y. Yan, G. Chen, H. Wei, R. Y.-C. Huang, J. Mo, D. L. Rempel, A. A. Tymiak, M. L. Gross, Fast photochemical oxidation of proteins (FPOP) maps the epitope of EGFR binding to adnectin, *Journal of The American Society for Mass Spectrometry* **2014**, *25*, 2084.

[77] Y. J. Lee, Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes, *Molecular BioSystems* **2008**, *4*, 816.

[78] X. Tang, G. R. Munske, W. F. Siems, J. E. Bruce, Mass spectrometry identifiable cross-linking strategy for studying protein-protein interactions, *Analytical Chemistry* **2005**, *77*, 311.

[79]   S. Kang, L. Mou, J. Lanman, S. Velu, W. J. Brouillette, P. E. Prevelige, Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry, *Rapid Communications in Mass Spectrometry* **2009**, *23*, 1719.

[80]   M. A. Nessen, G. Kramer, J. Back, J. M. Baskin, L. E. J. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, C. G. de Koster, Selective Enrichment of Azide-Containing Peptides from Complex Mixtures, *Journal of Proteome Research* **2009**, *8*, 3702.

[81]   Z. A. Chen, A. Jawhari, L. Fischer, C. Buchen, S. Tahir, T. Kamenski, M. Rasmussen, L. Lariviere, J.-C. Bukowski-Wills, M. Nilges, P. Cramer, J. Rappsilber, Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry, *EMBO J* **2010**, *29*, 717.

[82]   M. A. Lauber, J. P. Reilly, Structural Analysis of a Prokaryotic Ribosome Using a Novel Amidinating Cross-Linker and Mass Spectrometry, *Journal of Proteome Research* **2011**, *10*, 3604.

[83]   M. P. Washburn, D. Wolters, J. R. Yates, Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nature biotechnology* **2001**, *19*, 242.

[84]   Y. J. Lee, Probability-based shotgun cross-linking sites analysis, *Journal of The American Society for Mass Spectrometry* **2009**, *20*, 1896.

[85]   M. I. Rasmussen, J. C. Refsgaard, L. Peng, G. Houen, P. Højrup, CrossWork: Software-assisted identification of cross-linked peptides, *Journal of Proteomics* **2011**, *74*, 1871.

[86]   J. W. Back, L. de Jong, A. O. Muijsers, C. G. de Koster, Chemical Cross-linking and Mass Spectrometry for Protein Structural Modeling, *Journal of Molecular Biology* **2003**, *331*, 303.

[87]   S. Walsh, D. Diamond, Non-linear curve fitting using microsoft excel solver, *Talanta* **1995**, *42*, 561.

[88]   P. Quinn, W. Williams, The structural role of lipids in photosynthetic membranes, *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes* **1983**, *737*, 223.

[89]   M. Nishihara, K. Yokota, M. Kito, Lipid molecular species composition of thylakoid membranes, *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism* **1980**, *617*, 12.

[90]   P. G. Roughan, Phosphatidylglycerol and chilling sensitivity in plants, *Plant physiology* **1985**, *77*, 740.

[91]   A. Zavalin, J. Yang, K. Hayden, M. Vestal, R. M. Caprioli, Tissue protein imaging at 1 μm laser spot diameter for high spatial resolution and high imaging speed using transmission geometry MALDI TOF MS, *Analytical and Bioanalytical Chemistry* **2015**, *407*, 2337.

[92]   A. R. Korte, G. B. Yagnik, A. D. Feenstra, Y. J. Lee, in *Mass spectrometry imaging of small molecules*, Springer, **2015**, pp. 49.

[93]   G. Robichaud, K. P. Garrard, J. A. Barry, D. C. Muddiman, MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform, *Journal of The American Society for Mass Spectrometry* **2013**, *24*, 718.

[94]   D. C. Perdian, Y. J. Lee, Imaging MS Methodology for More Chemical Information in Less Data Acquisition Time Utilizing a Hybrid Linear Ion Trap–Orbitrap Mass Spectrometer, *Analytical Chemistry* **2010**, *82*, 9393.

[95]   J. Barber, K. Gounaris, in *Current topics in photosynthesis*, Springer, **1986**, pp. 237.

[96]   J. R. Kenrick, D. G. Bishop, The fatty acid composition of phosphatidylglycerol and sulfoquinovosyldiacylglycerol of higher plants in relation to chilling sensitivity, *Plant physiology* **1986**, *81*, 946.

[97]   E. Páldi, G. Szalai, C. L. Marton, M. Pál, T. Janda, Role of some N-containing compounds in chilling tolerance of maize, *Acta Biol Szeged* **2002**, *46*, 99.

[98]   J. Mock, M. McNeill, Cold tolerance of maize inbred lines adapted to various latitudes in North America, *Crop Science* **1979**, *19*, 239.

[99]   W. Hou, H. Zhou, M. B. Khalil, D. Seebun, S. A. Bennett, D. Figeys, Lyso-form fragment ions facilitate the determination of stereospecificity of diacyl

glycerophospholipids, *Rapid Communications in Mass Spectrometry* **2011**, *25*, 205.

[100] E. Haverkate, L. Van Deenen, Isolation and chemical characterization of phosphatidyl glycerol from spinach leaves, *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism* **1965**, *106*, 78.

[101] G. Roughan, R. Slack, Glycerolipid synthesis in leaves, *Trends in Biochemical Sciences* **1984**, *9*, 383.

[102] S. A. Hogenhout, J. I. B. Bos, Effector proteins that modulate plant–insect interactions, *Current Opinion in Plant Biology* **2011**, *14*, 422.

[103] P. N. Dodds, J. P. Rathjen, Plant immunity: towards an integrated view of plant–pathogen interactions, *Nat Rev Genet* **2010**, *11*, 539.

[104] J. D. G. Jones, J. L. Dangl, The plant immune system, *Nature* **2006**, *444*, 323.

[105] J. Bohlmann, G. Meyer-Gauen, R. Croteau, Plant terpenoid synthases: molecular biology and phylogenetic analysis, *Proceedings of the National Academy of Sciences* **1998**, *95*, 4126.

[106] T. Sana, S. Fischer, G. Wohlgemuth, A. Katrekar, K.-h. Jung, P. Ronald, O. Fiehn, Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen Xanthomonas oryzae pv. oryzae, *Metabolomics* **2010**, *6*, 451.

[107] W. E. Riedell, E. A. Beckendorf, M. A. Catangui, Soybean Aphid Injury Effects on Shoot Nitrogen Components in Glycine max, *Crop Sci.* **2013**, 232.

[108] D. Sato, M. Sugimoto, H. Akashi, M. Tomita, T. Soga, Comparative metabolite profiling of foxglove aphids (Aulacorthum solani Kaltenbach) on leaves of resistant and susceptible soybean strains, *Molecular BioSystems* **2014**, *10*, 909.

[109] M. E. Studham, G. C. MacIntosh, Multiple Phytohormone Signals Control the Transcriptional Response to Soybean Aphid Infestation in Susceptible and Resistant Soybean Plants, *Molecular Plant-Microbe Interactions* **2013**, *26*, 116.

[110] M. Schad, R. Mungur, O. Fiehn, J. Kehr, Metabolic profiling of laser microdissected vascular bundles of Arabidopsis thaliana, *Plant methods* **2005**, *1*, 2.

[111] J. Fang, B. Schneider, Laser Microdissection: a Sample Preparation Technique for Plant Micrometabolic Profiling, *Phytochemical Analysis* **2013**.

[112] D. Eikel, M. Vavrek, S. Smith, C. Bason, S. Yeh, W. A. Korfmacher, J. D. Henion, Liquid extraction surface analysis mass spectrometry (LESA-MS) as a novel profiling tool for drug distribution and metabolism analysis: the terfenadine example, *Rapid Communications in Mass Spectrometry* **2011**, *25*, 3587.

[113] J. G. Swales, J. W. Tucker, N. Strittmatter, A. Nilsson, D. Cobice, M. R. Clench, C. L. Mackay, P. E. Andren, Z. Takats, P. J. H. Webborn, R. J. A. Goodwin, Mass Spectrometry Imaging of Cassette-Dosed Drugs for Higher Throughput Pharmacokinetic and Biodistribution Analysis, *Analytical Chemistry* **2014**, *86*, 8473.

[114] I. Sparkes, F. Brandizzi, Fluorescent protein-based technologies: shedding new light on the plant endomembrane system, *The Plant Journal* **2012**, *70*, 96.

[115] R. M. Caprioli, T. B. Farmer, J. Gile, Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS, *Analytical Chemistry* **1997**, *69*, 4751.

[116] H. Wei, K. Nolkrantz, D. H. Powell, J. H. Woods, M. C. Ko, R. T. Kennedy, Electrospray sample deposition for matrix-assisted laser desorption/ionization (MALDI) and atmospheric pressure MALDI mass spectrometry with attomole detection limits, *Rapid Communications in Mass Spectrometry* **2004**, *18*, 1193.

[117] N. Bjarnholt, B. Li, J. D'Alvise, C. Janfelt, Mass spectrometry imaging of plant metabolites - principles and possibilities, *Natural Product Reports* **2014**, *31*, 818.

[118] P. J. Horn, K. D. Chapman, Lipidomics in situ: Insights into plant lipid metabolism from high resolution spatial maps of metabolites, *Progress in Lipid Research* **2014**, *54*, 32.

[119] P. J. Horn, A. R. Korte, P. B. Neogi, E. Love, J. Fuchs, K. Strupat, L. Borisjuk, V. Shulaev, Y.-J. Lee, K. D. Chapman, Spatial Mapping of Lipids at Cellular Resolution in Embryos of Cotton, *The Plant Cell Online* **2012**, *24*, 622.

[120] A. R. Korte, Z. Song, B. J. Nikolau, Y. J. Lee, Mass spectrometric imaging as a high-spatial resolution tool for functional genomics: Tissue-specific gene expression of TT7 inferred from heterogeneous distribution of metabolites in Arabidopsis flowers, *Analytical Methods* **2012**, *4*, 474.

[121] J. Thunig, S. H. Hansen, C. Janfelt, Analysis of Secondary Plant Metabolites by Indirect Desorption Electrospray Ionization Imaging Mass Spectrometry, *Analytical Chemistry* **2011**, *83*, 3256.

[122] E. Cabral, D. Ifa, in *Mass Spectrometry Imaging of Small Molecules, Vol. 1203* (Ed.: L. He), Springer New York, **2015**, pp. 63.

[123] D. I. Campbell, C. R. Ferreira, L. S. Eberlin, R. G. Cooks, Improved spatial resolution in the imaging of biological tissue using desorption electrospray ionization, *Analytical and Bioanalytical Chemistry* **2012**, *404*, 389.

[124] V. Kertesz, G. J. Van Berkel, Improved imaging resolution in desorption electrospray ionization mass spectrometry, *Rapid Communications in Mass Spectrometry* **2008**, *22*, 2639.

[125] P. J. Roach, J. Laskin, A. Laskin, Nanospray desorption electrospray ionization: an ambient method for liquid-extraction surface sampling in mass spectrometry, *Analyst* **2010**, *135*, 2233.

[126] R. Shroff, F. Vergara, A. Muck, A. Svatoš, J. Gershenzon, Nonuniform distribution of glucosinolates in Arabidopsis thaliana leaves has important consequences for plant defense, *Proceedings of the National Academy of Sciences* **2008**, *105*, 6196.

[127] R. Shroff, K. Schramm, V. Jeschke, P. Nemes, A. Vertes, J. Gershenzon, A. Svatoš, Quantification of plant surface metabolites by MALDI mass spectrometry imaging: glucosinolates on Arabidopsis thaliana leaves, *The Plant Journal* **2015**.

[128] G. Hamm, V. Carré, A. Poutaraud, B. Maunit, G. Frache, D. Merdinoglu, J.-F. Muller, Determination and imaging of metabolites from Vitis vinifera leaves by laser desorption/ionisation time-of-flight mass spectrometry, *Rapid Communications in Mass Spectrometry* **2010**, *24*, 335.

[129] Y. Li, J. Zou, M. Li, D. D. Bilgin, L. O. Vodkin, G. L. Hartman, S. J. Clough, Soybean defense responses to the soybean aphid, *New Phytologist* **2008**, *179*, 185.

[130] M. V. Chiozza, M. E. O'Neal, G. C. MacIntosh, Constitutive and Induced Differential Accumulation of Amino Acid in Leaves of Susceptible and Resistant Soybean Plants in Response to the Soybean Aphid (Hemiptera: Aphididae), *Environmental Entomology* **2010**, *39*, 856.

[131] J. Diaz-montano, J. C. Reese, W. T. Schapaugh, L. R. Campbell, Chlorophyll Loss Caused by Soybean Aphid (Hemiptera: Aphididae) Feeding on Soybean, *Journal of Economic Entomology* **2007**, *100*, 1657.

[132] A. Evidente, V. Venturi, M. Masi, G. Degrassi, A. Cimmino, L. Maddau, A. Andolfi, In Vitro Antibacterial Activity of Sphaeropsidins and Chemical Derivatives toward Xanthomonas oryzae pv. oryzae, the Causal Agent of Rice Bacterial Blight, *Journal of Natural Products* **2011**, *74*, 2520.

[133] I. Ahuja, R. Kissen, A. M. Bones, Phytoalexins in defense against pathogens, *Trends in Plant Science* **2012**, *17*, 73.

[134] E. A. Schmelz, A. Huffaker, J. W. Sims, S. A. Christensen, X. Lu, K. Okada, R. J. Peters, Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins, *Plant J* **2014**, *79*, 659.

[135] T. Akatsuka, O. Kodama, H. Sekido, Y. Kono, S. Takeuchi, Novel phytoalexins (oryzalexins A, B, and C) isolated from rice blast leaves infected with *Pyricularia oryzae*. Part I: Isolation, characterization and biological activities of oryzalexins, *Agric. Biol. Chem.* **1985**, *49*, 1689.

[136] S. Taira, Y. Sugiura, S. Moritake, S. Shimma, Y. Ichiyanagi, M. Setou, Nanoparticle-assisted laser desorption/ionization based mass imaging with cellular resolution, *Analytical Chemistry* **2008**, *80*, 4761.

[137] K. Shrivas, T. Hayasaka, Y. Sugiura, M. Setou, Method for Simultaneous Imaging of Endogenous Low Molecular Weight Metabolites in Mouse Brain Using TiO2 Nanoparticles in Nanoparticle-Assisted Laser Desorption/Ionization-Imaging Mass Spectrometry, *Analytical Chemistry* **2011**, *83*, 7283.

[138] M. Paz, J. Martinez, A. Kalvig, T. Fonger, K. Wang, Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation, *Plant Cell Rep* **2006**, *25*, 206.

[139] J. A. Hankin, R. M. Barkley, R. C. Murphy, Sublimation as a method of matrix application for mass spectrometric imaging, *Journal of The American Society for Mass Spectrometry* **2007**, *18*, 1646.

[140] S. M. Kosina, A. Castillo, S. R. Schnebly, R. L. Obendorf, Soybean seed coat cup unloading on plants with low-raffinose, low-stachyose seeds, *Seed Science Research* **2009**, *19*, 145.

[141] S. M. Kosina, S. R. Schnebly, R. L. Obendorf, Are Raffinose and Stachyose Unloaded from Soybean Seed Coats to Developing Embryos?, *Open Plant Science Journal* **2013**, *7*.

[142] D. Price, A. Karley, D. Ashford, H. Isaacs, M. Pownall, H. Wilkinson, J. Gatehouse, A. Douglas, Molecular characterisation of a candidate gut sucrase in the pea aphid,< i> Acyrthosiphon pisum</i>, *Insect biochemistry and molecular biology* **2007**, *37*, 307.

[143] T. Wilkinson, D. Ashford, J. Pritchard, A. Douglas, Honeydew sugars and osmoregulation in the pea aphid Acyrthosiphon pisum, *Journal of Experimental Biology* **1997**, *200*, 2137.

[144] A. Douglas, Phloem-sap feeding by animals: problems and solutions, *Journal of Experimental Botany* **2006**, *57*, 747.

[145] F. Wäckers, J. Lee, G. Heimpel, K. Winkler, R. Wagenaar, Hymenopteran parasitoids synthesize 'honeydew-specific'oligosaccharides, *Functional Ecology* **2006**, *20*, 790.

[146] P. W. Miles, Aphid saliva, *Biological Reviews* **1999**, *74*, 41.

[147] F. L. Goggin, Plant–aphid interactions: molecular and ecological perspectives, *Current Opinion in Plant Biology* **2007**, *10*, 399.

[148] H. Návarová, F. Bernsdorff, A.-C. Döring, J. Zeier, Pipecolic Acid, an Endogenous Mediator of Defense Amplification and Priming, Is a Critical Regulator of Inducible Plant Immunity, *The Plant Cell Online* **2012**, *24*, 5123.

[149] D. Vogel-Adghough, E. Stahl, H. Návarová, J. Zeier, Pipecolic acid enhances resistance to bacterial infection and primes salicylic acid and nicotine accumulation in tobacco, *Plant signaling & behavior* **2013**, *8*, e26366.

[150] J. Zhu, K.-C. Park, Methyl salicylate, a soybean aphid-induced plant volatile attractive to the predator Coccinella septempunctata, *Journal of chemical ecology* **2005**, *31*, 1733.

[151] J. Shah, J. Zeier, Long-distance communication and signal amplification in systemic acquired resistance, *Frontiers in plant science* **2013**, *4*.

[152] S. Murakami, R. Nakata, T. Aboshi, N. Yoshinaga, M. Teraishi, Y. Okumoto, A. Ishihara, H. Morisaka, A. Huffaker, E. A. Schmelz, Insect-Induced Daidzein, Formononetin and Their Conjugates in Soybean Leaves, *Metabolites* **2014**, *4*, 532.

[153] F. Meng, Y. Han, W. Teng, Y. Li, W. Li, QTL underlying the resistance to soybean aphid (Aphis glycines Matsumura) through isoflavone-mediated antibiosis in soybean cultivar 'Zhongdou 27', *Theoretical and applied genetics* **2011**, *123*, 1459.

[154] B. A. Vinatzer, G. M. Teitzel, M. W. Lee, J. Jelenska, S. Hotton, K. Fairfax, J. Jenrette, J. T. Greenberg, The type III effector repertoire of Pseudomonas syringae pv. syringae B728a and its role in survival and disease on host and non-host plants, *Molecular microbiology* **2006**, *62*, 26.

[155] B. C. Freeman, C. Chen, G. A. Beattie, Identification of the trehalose biosynthetic loci of Pseudomonas syringae and their contribution to fitness in the phyllosphere, *Environmental microbiology* **2010**, *12*, 1486.

[156] M. Kurz, A. Y. Burch, B. Seip, S. E. Lindow, H. Gross, Genome-driven investigation of compatible solute biosynthesis pathways of Pseudomonas syringae pv. syringae and their contribution to water stress tolerance, *Applied and environmental microbiology* **2010**, *76*, 5452.

[157] B. Sagot, M. Gaysinski, M. Mehiri, J.-M. Guigonis, D. Le Rudulier, G. Alloing, Osmotically induced synthesis of the dipeptide N-acetylglutaminylglutamine amide is mediated by a new pathway conserved among bacteria, *Proceedings of the National Academy of Sciences* **2010**, *107*, 12652.

[158] S. Li, X. Yu, G. A. Beattie, Glycine betaine catabolism contributes to Pseudomonas syringae tolerance to hyperosmotic stress by relieving betaine-mediated suppression of compatible solute synthesis, *Journal of bacteriology* **2013**, *195*, 2415.

[159] H. Robert, C. Le Marrec, C. Blanco, M. Jebbar, Glycine betaine, carnitine, and choline enhance salinity tolerance and prevent the accumulation of sodium to a level inhibiting growth of Tetragenococcus halophila, *Applied and environmental microbiology* **2000**, *66*, 509.

[160] C. Chen, G. A. Beattie, Characterization of the osmoprotectant transporter OpuC from Pseudomonas syringae and demonstration that cystathionine-β-synthase domains are required for its osmoregulatory function, *Journal of bacteriology* **2007**, *189*, 6901.

[161] C. Chen, G. A. Beattie, Pseudomonas syringae BetT is a low-affinity choline transporter that is responsible for superior osmoprotection by choline over glycine betaine, *Journal of bacteriology* **2008**, *190*, 2717.

[162] C. Chen, A. A. Malek, M. J. Wargo, D. A. Hogan, G. A. Beattie, The ATP-binding cassette transporter Cbc (choline/betaine/carnitine) recruits multiple substrate-binding proteins with strong specificity for distinct quaternary ammonium compounds, *Molecular microbiology* **2010**, *75*, 29.

[163] R. Panter, J. Mudd, Carnitine levels in some higher plants, *Febs Letters* **1969**, *5*, 169.

[164] C. Chen, S. Li, D. R. McKeever, G. A. Beattie, The widespread plant-colonizing bacterial species Pseudomonas syringae detects and exploits an extracellular pool of choline in hosts, *The Plant Journal* **2013**, *75*, 891.

[165] M. Jemal, High-throughput quantitative bioanalysis by LC/MS/MS, *Biomedical Chromatography* **2000**, *14*, 422.

[166] E. von Roepenack-Lahaye, T. Degenkolb, M. Zerjeski, M. Franz, U. Roth, L. Wessjohann, J. Schmidt, D. Scheel, S. Clemens, Profiling of Arabidopsis secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry, *Plant physiology* **2004**, *134*, 548.

[167] C. Böttcher, E. von Roepenack-Lahaye, J. Schmidt, C. Schmotz, S. Neumann, D. Scheel, S. Clemens, Metabolome analysis of biosynthetic mutants reveals a diversity of metabolic changes and allows identification of a large number of new compounds in Arabidopsis, *Plant physiology* **2008**, *147*, 2107.

[168] W. B. Dunn, Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes, *Physical biology* **2008**, *5*, 011001.

[169] A. Honda, K. Yamashita, T. Hara, T. Ikegami, T. Miyazaki, M. Shirai, G. Xu, M. Numazawa, Y. Matsuzaki, Highly sensitive quantification of key regulatory oxysterols in biological samples by LC-ESI-MS/MS, *Journal of lipid research* **2009**, *50*, 350.

[170] Y. Hsieh, Potential of HILIC-MS in quantitative bioanalysis of drugs and drug metabolites, *Journal of separation science* **2008**, *31*, 1481.

[171] W. Z. Shou, W. Naidong, Simple means to alleviate sensitivity loss by trifluoroacetic acid (TFA) mobile phases in the hydrophilic interaction chromatography–electrospray tandem mass spectrometric (HILIC–ESI/MS/MS) bioanalysis of basic compounds, *Journal of Chromatography B* **2005**, *825*, 186.

[172] K. Inoue, R. Obara, T. Hino, H. Oka, Development and application of an HILIC-MS/MS method for the quantitation of nucleotides in infant formula, *Journal of agricultural and food chemistry* **2010**, *58*, 9918.

[173] R. Oertel, V. Neumeister, W. Kirch, Hydrophilic interaction chromatography combined with tandem-mass spectrometry to determine six aminoglycosides in serum, *Journal of Chromatography A* **2004**, *1058*, 197.

[174] T. S. Moore, Phosphatidylcholine synthesis in castor bean endosperm, *Plant physiology* **1976**, *57*, 382.

[175] J. Bremer, Carnitine-metabolism and functions, *Physiol Rev* **1983**, *63*, 1420.