PROFILING, STRUCTURE ELUCIDATION AND INTERROGATION OF BIOSYNTHESIS OF SPECIALIZED METABOLITES OF THE SOLANACEAE USING LIQUID CHROMATOGRAPHY/ MASS SPECTROMETRY AND STABLE ISOTOPE LABELING

By

Xiaoxiao Liu

A DISSERTATION

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

Chemistry – Doctor of Philosophy

2017

ABSTRACT

PROFILING, STRUCTURE ELUCIDATION AND INTERROGATION OF BIOSYNTHESIS OF SPECIALIZED METABOLITES OF THE SOLANACEAE USING LIQUID CHROMATOGRAPHY/ MASS SPECTROMETRY AND STABLE ISOTOPE LABELING

By

Xiaoxiao Liu

Hundreds of thousands of specialized metabolites are estimated to be produced and accumulated within the plant kingdom, many of which are considered potentially valuable natural compounds that can be used by humans. One of the greatest challenges for specialized metabolism analysis is the tremendous chemical diversity of such phytochemicals observed within and across plant species. To better address the diversification of plant specialized metabolites produced by glandular trichomes in the family *Solanaceae*, as well as understand the genetic and biochemical mechanisms underlying it, several mass spectrometry-based analytical platforms coupled with chromatographic separations, NMR, and stable isotope labeling have been employed in this study.

In chapter 2, the diversity of acylsugar metabolites identified from three petunia species was explored as a complementary research to the previous analysis of acylsugar within *Solanaceae* family using LC-MS based strategies. 28 novel acylsucorses, including a class of malonyl ester acylsucroses, have been successfully purified through solid phase extraction followed by HPLC separation, and structures of them are proposed according to UHPLC-MS and NMR analysis. Interestingly, all acylsucroses characterized from petunia species have not previously been detected from studies of other *Solanum* species including petunia hybrids. This work extended our knowledge of the acylsugar diversity within *Solanaceae* family, and can be served as

references in studying the bioactive of genes and enzymes involved in their biosynthesis.

To fulfill the need for rapid characterization of plant chemical diversity and improve the efficacy of metabolite identification, a UHPLC-MS/MS coupled with stable isotopic labeling strategy was established and summarized in chapter 3. [^{13}C]-Branched chain amino acids (BCAAs) were introduced as isotopic tracers to label and distinguish isomeric plant metabolites with diverse structures, particularly in respect to the lengths and branching type of aliphatic chains derived from metabolism of BCAAs. Metabolites tagged with stable isotopes were annotated easily according to their corresponded mass shift from unlabeled species using UHPLC-MS with both multiplexed CID and MS/MS analysis. This study provides a fast and simplified method to investigate the structural diversity of plant specialized metabolites without purification and NMR.

Another question addressed in the study is how do plants make the huge diversity of acylsugar metabolites. Chapter 4 described the *in vivo* investigation of acylsugar related acyl chain elongation mechanism using stable isotope labeling coupled with UHPLC-MS/MS methodology. Fragments of acyl chain generated from acylsugars with and without stable isotope labels derived from [^{13}C]-BCAA feeding of plants were analyzed using LC-MS/MS, and related aliphatic chain elongation pathways were proposed based on the resulting information. Evidence was observed to support that BCAA derived one-carbon elongation pathway existed in both wild tomato LA2560 and cultivated tomato, while two-carbon elongation was identified in *S. pennellii* LA0716 and LA2560.

ACKNOWLEDGEMENTS

Many people deserve thanks and appreciation in this thesis for their continuous support all those years throughout my entire PhD study. First and foremost, I would like to express my sincerest gratitude to my advisor Dr. A. Daniel Jones for his persistent understanding and support. I still can remember clearly the first day when I went to his office, with almost no experience about NMR and mass spectrometry but full of nervous, he kindly guided me to figure out what I really want to do and encouraged me to keep learning and be prepared to take challenges at the very first time. Thanks Dan for really caring about my thoughts and feelings, and offering valuable suggestions at any time when I need help, through the direction of research to my career and personal development. His patience and encouragement have assisted me to overcome those difficult times at graduate school. I feel so lucky to have such a knowledgeable and wonderful mentor for both scientific career and the rest of personal life.

Special thanks to my collaborators in trichome project, especially Dr. Rob Last's group and Dr. Cornelius Barry's group. It is an honor to work with all those brilliant scientists and learn plant biology from them. I also want to gratefully thank the support from RTSF Mass Spectrometry Core and Max T. Rogers NMR facility at Michigan State University. My thanks go especially to Dr. Dan Holmes for training me on the NMR instrument and helping me generate tons of NMR data to complete my research.

I also like to show my appreciation to members of my PhD committee, Dr. David Weliky, Dr. Dana Spence and Dr. Kevin D. Walker for their insightful suggestions and constructive

comments about my work. Additionally, I would like to extend my thanks to all group members at Dan Jones' research lab, particularly to Dr. Zhenzhen Wang, Dr. Chen Zhang, Dr. Sujana Pradhan, Dr. Afrand Kamali, Dr. Probodha Ekanayak and Steven Hurney. Thanks for their greatest companion and friendship all these years, both in- and off-campus. Without them, I cannot have so much joyful and memorable times in my graduate school life.

Last but not the least, my deepest thanks go to family and friends. I would like to sincerely thank my parents Dr. Fengling Liu and Chunyuan Yang for their unlimited love, concerns and encouragements. I really appreciate their companion every summer for traveling across such a long distance from China and bringing me all the happiness about staying with the family together. They are the strongest backup for me through the good and bad period of my entire life. Thanks to all my dearest friends for those fascinate breaks and long night talking. Without them, the graduate school will be boring and unbearable. Finally, thanks to many other people who provide me help and kindness at MSU. This work would not be possible without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One. Introduction

1.1 Importance and evolution of specialized metabolism in plants

1.1.1 Importance of plant specialized metabolites

By providing energy, oxygen, and organic matter derived from photosynthesis, plants are the basis of most land ecosystems on earth, and have been utilized as natural resources of foods, fuels, construction materials, medicines and pesticides by humans since prehistoric times. Within the last century, the dramatic rise in world population has led to a grave threat to limited natural resources, and brings great challenges to the sustainability and renewability of modern agriculture. Advances in biology and biotechnology, especially in genetic engineering, hold hopes for discovering novel solutions to end hunger and ensure environmental sustainability [1-3]. For example, several commercial crops, including tobacco, tomato, and soybean, have been genetically modified to increase production, tolerance to herbivores, and improve the quality of products [4-6]. Besides these achievements, a better understanding of diversity of biochemical compounds and related metabolic biosynthesis mechanisms is still essential to making the best use of metabolic engineering technology and producing valuable natural products derived from plants [7].

Hundreds of thousands of specialized metabolites are estimated to be synthesized and accumulated within the plant kingdom, many of which affect plant chemical defenses against both biotic and abiotic stress [8,9], attraction of pollinators [10], and competition and communication with other species [11-13]. To be distinguished from primary metabolites, which comprise a group of natural chemicals present among all species and are essential for growth and development, specialized metabolites used to be named as “secondary metabolites”, as compounds only observed in particular species and suspected to have no essential biological function but treated as waste by-products of primary metabolism [14, 15]. As studies focused on

plant secondary metabolites have accumulated during the past 50 years, recognition has grown that the biosynthesis of such natural compounds is restricted to specific plant lineages, and functions of these chemicals are often related to "species-specific" ecological situations. Such awareness led to the proposition of the new term "specialized metabolites" to describe these plant products [12,16,17]. Now, the total number of specialized metabolites produced within the plant kingdom has been estimated at approximately 200,000 [18], many of which are considered potentially valuable natural compounds that can be used as pharmaceuticals, fragrances, flavors, nutrients and natural pesticides by humans [19-21].

Plants are commonly considered as "green chemical factories" since specialized metabolites are assembled efficiently in environmentally friendly processes in specific plant tissues. A great portion of these bioactive compounds is known for their medicinal properties and many have been selected as or have inspired drug candidates by the modern pharmaceutical industry. In fact, more than half of the new small molecule chemical entities approved between 1981 and 2010 were derived from plants [22-24]. One of the most widely used medication, aspirin, originated from the discovery of salicin in bark of willow trees and has been utilized by humans for more than 2,000 years [25]. Beyond this, plant derived specialized metabolites have become indispensable to the development of anti-cancer drugs. For example, Topotecan and Irinotecan, used as cancer chemotherapeutic agents, are water-soluble analogues of camptothecin, which is a terpene indole alkaloid metabolite isolated from the bark and stem of *Camptotheca acuminata* in the 1960s [26]; Paclitaxel, purified from bark of pacific yew, was introduced as medicine to treat several types of cancer in 1993 [27]. Additionally, half of the Nobel Prize in Physiology or Medicine was awarded to Chinese scientist Youyou Tu in 2015, for her discovery of artemisinin (also called qinghaosu) from the plant *Artemisia annua,* which has been used as effective

medicine against malaria. This award serves as not only a milestone for research on traditional Chinese medicines, but also attracts more attention to the study of plant specialized metabolites, which are tremendous resources of natural chemicals with potential as drug candidates, as well as other valuable products.

1.1.2 History and growing interest in plant specialized metabolite research

Research on plant specialized metabolites has expanded significantly over the past two hundred years, transformed from specific compound identification to associated gene discoveries. In the late half of the 19th century, plant specialized metabolites (named as secondary metabolites) were considered as side-products produced through primary metabolite biosynthesis or metabolism, with functions not well demonstrated but known as chemicals that were not essential to the survival of plants [14,15]. The isolation of morphine from opium poppy in 1806 has been considered as the starting point of plant specialized metabolite research [21]. For the following 150 years, studies of such chemicals had progressed rapidly, while most efforts were mainly focused on isolation and structure elucidation of novel plant products. In the early 1950s, radioisotopic labeling techniques were introduced to reveal metabolic pathways of natural compounds, which turned the emphasis on specialized metabolites research toward aspects of their biosynthesis [28]. In the 1980s, the investigation of biological functions of specialized metabolites became a "hot-spot" in research, and this research proved that specialized metabolites play important roles in plant responses to various environmental stresses [29]. During the past 50 years, research on specialized metabolites have focused on the gene level, aiming to identify genes involved in the synthesis of plant products. At the same time, investigations into the evolutionary origins of plant specialized metabolites, including tracing

important genes and reconstructing the corresponding evolutionary mechanisms, would provide a new and exciting area that would lead answers about how plants have evolved to produce so many diverse metabolites [21].

1.1.3 Evolution of plant specialized metabolism

An enormous number of specialized metabolites, considered low molecular weight organic chemicals with complex structures, are produced by a vast array of plant species. Their diversity has excited the curiosity of researchers with regard to two major aspects: What advantages accrue to plants when they synthesize such large amounts of specialized metabolites? Also, what part of each plant's finite genome is devoted to specialized metabolism? It has been confirmed in many cases that most plant specialized metabolites play key roles as responses to environmental stress and communicating or competing with other species, and in consequence, the diversity of plant natural chemicals is necessary to help plants respond and thrive in a diverse set of complex and dynamic environments [8-13,30,31,32]. The reciprocal evolutionary relationship between butterflies and their food plants reported by Ehrlich and Raven suggested that the diversity of plant specialized metabolites was a result of stepwise evolutionary responses to the evolution of insect herbivores [31]. Furthermore, another evolutionary model proposed by Jones and Firn suggested that plant fitness is increased when they contain a very high diversity of specialized metabolites in order to reach a reasonable probability of making enough bioactive compounds to strengthen their defenses, while inactive chemicals are still retained as products of mutants that increase the probability of producing novel useful metabolites against other stresses [32]. These two models together help to explain the vast diversity of specialized metabolites accumulating within plants. Nevertheless, since consensus has not been reached regarding the costs endured by

plants when preserving such a large pool of specialized metabolites [33], it is challenging to speculate the benefits to plants when they produce more specialized metabolites in response to a particular environmental condition.

Compared with the size of most plant genomes, the number of specialized metabolites produced in many plant species far exceeds the number of genes that potentially code for biosynthetic enzymes [12]. Only 15-25% of plant genomes, which often contain 20,000 to 60,000 genes, encode enzymes annotated as having potential involvement in specialized metabolite biosynthesis, corresponding to a small fraction of all enzymes required for producing such a huge number of compounds in a given plant species [17,34,35]. It is believed that the biosynthesis of several specialized metabolites in plants is derived from primary metabolic mechanisms [13,16]. For instance, terpene metabolites are assembled using isopentenyl pyrophosphate and dimethylallyl pyrophosphate precursors, key components in isoprenoid biosynthetic pathways [16]. Alkaloid metabolites are synthesized using either products or intermediates of primary metabolism, including various amino acids and purine nucleosides [36]. Acylsugar metabolites are commonly derived from primary metabolites in plant trichomes, using glucose, sucrose and aliphatic acids as precursors for their biosynthesis [37]. Such related bioactive transitions of enzymes from primary metabolism to specialized metabolism have been attributed to gene duplication, random mutations that alter substrate specificity, and convergent evolution [12,16,17,30].

Examples of genes and encoded enzymes in plant specialized metabolism that have evolved, after multiple mutations, similar functions as primary metabolism genes, include large gene families discovered from different plant lineages, such as cytochrome P450s [38], terpene

synthases [39] and BAHD acyltransferases [40]. Novel genes arising from mutations of primary or specialized metabolic genes may encode enzymes involved in the synthesis of new chemicals through use of different substrates. Normally, the new group of substrates for such enzymes may have similar structures compared with their progenitors, suggesting that mutations that alter one or several amino acids, sequence truncation or substitutions could alter preferences of newly substrates [17, 41, 42]. Additionally, discoveries of specialized metabolites with similar functions may have evolved independently from distinct plant lineages, and may be explained using the concept of convergent evolution [12]. Function-directed convergent evolution has taken place in many plants, yielding either different or similar specialized metabolites that perform the same biological functions in different species [43,44]. Meanwhile, the same specialized metabolite synthesized using unrelated genes or enzymes has been observed from recent studies [45, 46]. For intense, an acyltransferase of the BAHD family was involved in making methyl anthranilate in grape (*Vitis labrusca)* using anthraniloyl-CoA as the substrate, while biosynthesis of the same compound in corn (*Zea mays*) was catalyzed by a methyl transferase of the SABATH family to transfer methyl group to carboxyl group of anthranilic acid [45]. Moreover, it is worth mentioning that evidence of convergent evolution between plants and other organisms to make the same compounds have been found, including biosynthesis of dopamine in banana, potato, and avocados, and melatonin has been identified in almost all plants tested [47,48], while more investigations need to be pursued in order to reveal the benefits and driving forces for plants to make these specialized compounds. Finally, limited by the size of genomes, plants continue to evolve to selectively produce useful specialized metabolites that can help plants adapt changes from changing circumstances, and eliminate old compounds and genes

that are no longer useful. Such natural selection gives rise to the lasting diversification of specialized metabolites within the plant kingdom [12].

1.2 Storage glandular trichomes as model tissues to study specialized metabolites in the family *Solanaceae*

Trichomes are aerial epidermal appendages found on aboveground plant tissues from approximately 30% of all vascular plant species [49,50]. These uni- and multi-cellular protuberances have long been considered as the first line of plant defense systems, function as physical and chemical barriers against herbivores, and protect plants against UV-B radiation, water loss, and extreme temperature stress [51,52,53]. Trichomes that have glands at their tips play a critical role in the synthesis, storage and secretion of large amounts of specialized metabolites, and are called storage/secreting glandular trichomes (SGTs). Cultivated tomato and its wild relatives in the family *Solanaceae* produce several different types of trichomes. Luckwill's taxonomic survey of tomato species documented four morphologically distinct SGTs: type I trichomes have a long multicellular stalk and a small gland at the tip; type IV trichomes consist of a relatively shorter stalk and a small glandular structure at the tip; type VI trichomes contain a four-celled glandular top on a short (~0.1 mm) stalk; and type VII trichomes have a one-celled stalk and a four-eight celled gland at the tip [50,54].

SGTs contribute to plant resistance through several ways, one of them involves biosynthesis of a variety of volatile and nonvolatile specialized metabolites toxic to insect and pathogens including terpenes, acylsugars, methyl ketones, flavonoids, and phenylpropenes in SGTs [37,55-58]. Terpenes are one of the largest groups of plant specialized metabolites produced by type VI glandular trichomes of *Solanum* species, which have the repellent effects on insects and protect

plants against pests [51,59]. Methyl ketones are a class of volatile compounds derived from fatty acids, with length of skeleton commonly ranging from 7-15 carbons, and function mainly as plant chemical defense against insects. In some accessions of wild tomato *S. habrochaites,* 2-tridecanone has been detected as the dominant component of type VI trichomes with a much higher concentration compared with that of cultivated tomato [60]. Flavonoids are derived from the phenylpropanoid pathway and play an important role in protecting plants against damage caused by solar UV-A and UV-B radiation. These classes of metabolites accumulate in glandular trichomes of various accessions of wild tomato *S. habrochaites* [61,62]. Phenylpropenes are well known for their ability to attract pollinators to plants, while their biosynthesis mechanisms are not fully understood [63]. Acylsugars are reported to be produced within type I/IV trichomes of various *Solanaceae* species and will be introduced in more detail in subsequent sections of this dissertation.

There is a significant chemical diversity of specialized metabolites observed not only within the different type of SGTs, but across different *Solanaceae* species [55-58]. Additionally, as aerial surface cells, glandular trichomes are accessible for cell type-specific sampling for analysis of trichome chemistry and assessment of gene expression. Therefore, SGTs have been selected as excellent experimental objects to study the complexity and chemical diversity of specialized metabolites, as well as metabolic biosynthesis mechanism and other underlying principles. Together with state-of-the-art analytical strategies and modern genetic tools, comprehensive investigation of glandular trichome specialized metabolism, including both chemical diversity characterization and identification of specific genes related to metabolic biosynthesis, has become a reality.

1.3 Introduction of acylsugars and related studies

1.3.1 Accumulation and biological functions of acylsugar metabolites in the plant family *Solanaceae*

Acylsugars (also named as sugar polyesters or resin glycosides), a class of natural compounds commonly accumulated within type I/IV glandular trichomes of Solanaceous plants, are the most abundant specialized metabolites detected in the leaf extracts of some plants in the *Solanaceae*, and have been reported as part of chemical defenses that protect plants against insect herbivores [37,64,65]. Acylsugars may act either in direct defense as sticky and toxic entrapping exudates [69] or in indirect defense as "Dangerous Lollipops" [70]. A recent study of *Nicotiana attenuata*, a wild tobacco species, demonstrated that acylsugar-enriched plants were selected as the first meal of Lepidopteran larvae, and did not directly affect the survival of herbivores, but indirectly helped predators to locate feeding larvae by producing several volatile chemicals (mainly branched chain aliphatic acids) with distinct odors from hydrolysis of acylsugars [70]. Due to their insect-resisting properties, acylsugar metabolites have been chosen as targets for breeding and genetic engineering of high-quality tomato species [71-73]. Besides, acylsugars are well known as commercially valuable compounds that have been utilized for decades as surfactants [74], emulsifiers [75], preservatives [76] and pharmaceutical excipients [77].

Previous liquid chromatography coupled with mass spectrometry (LC-MS)-based profiling of acylsugar metabolites from glandular trichomes showed significant diversity of acylsugar phenotypes within and across several plants from the *Solanaceae* [66-68]. Until recently, the number of known acylsugar metabolites was relatively small. Such metabolite profiling is essential for investigation of biosynthetic mechanism of acylsugars, and aids identification of novel genes and enzymes involved in acylsugar biosynthesis. Structure diversity of acylsugar

metabolites usually arises from various acyl chain lengths (C2-C12), different branching-types of aliphatic groups, sugar cores, the number of ester groups, and distinct positions of esterification. Acylsugars substituted with acyl chains longer than eight carbons have been identified with high insect resistance against psyllids, aphids, whiteflies and spider mites [70,78-80].

1.3.2 Structure diversity identification of acylsugar metabolites

With the development of modern analytical strategies, investigations of structural diversity of acylsugar metabolites in the *Solanaceae* have achieved remarkable progress. Acylsugars with either glucose or sucrose core have been identified as the two major classes present in the *Solanaceae*, and they are usually esterified with three to six aliphatic acids at specific positions on the sugar core. A universal nomenclature of acylsugar metabolites has been proposed according to our recent structural studies and explained in Figure 1.1 using tomato acylsugars G3:16(5,5,6) and S4:18(2,5,5,6) as examples [66-68]. The designation "G" refers to glucose core and "S" is used for sucrose core. The number following the sugar core designation indicates the

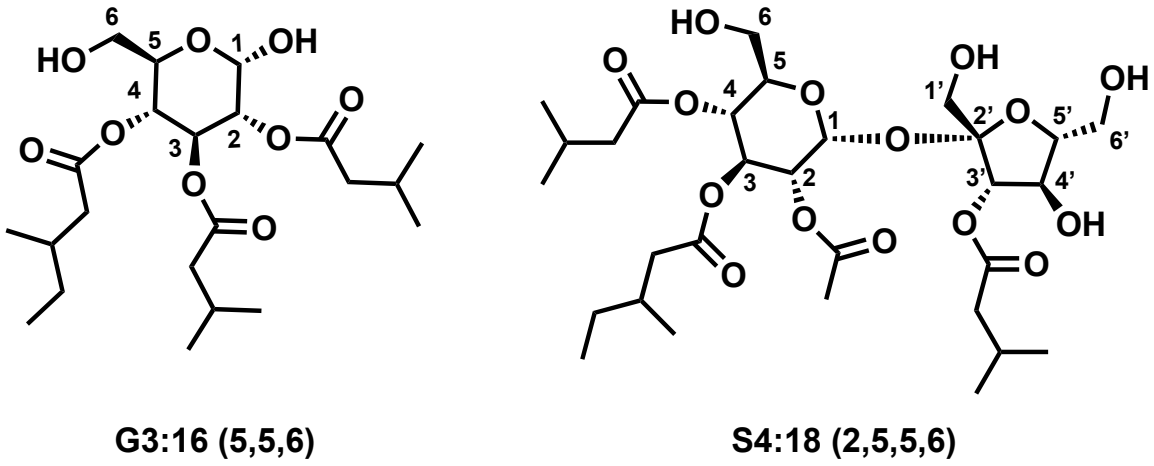


Figure 1.1. Nomenclature of acylsugar metabolites. *S. pennellii acylglucose G3:16(5,5,6) and cultivated tomato M82 acylsucrose S4:18(2,5,5,6) are used as examples. Substitution positions on sugar core are labeled with numbers in structures.*

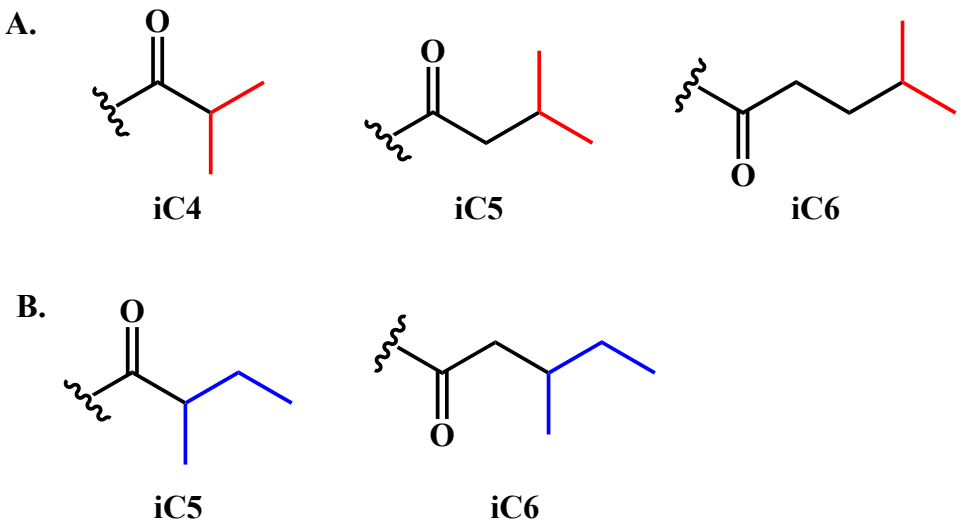


Figure 1.2. Common aliphatic acyl groups with different branching types. *A. Iso-branched iC4, iC5, iC6 acyl groups. Iso-branched (abbreviated as "i") terminals are labeled by red. B. Anteiso-branched aiC5, aiC6 acyl groups. Anteiso-branched (abbreviated as "ai") terminals are labeled by blue.*

total number of acyl groups attached to the sugar. In this case, there are tri-substituted acylglucose and tetra-substituted acylsucrose presented. The numeral following the colon indicates the total number of carbons in all acyl chains, which are 16 and 18 carbons separately for G3:16 and S4:18. The numbers in parentheses represent the lengths of each individual acyl group. For G3:16 (5,5,6), there are two C5 chains and one C6 group attached to a glucose core at the 2, 3, and 4 positions, while in S4:18, there is one acetyl (C2) chain, one C5 and one C6 group substituted at the 2, 3, and 4 positions on the sucrose pyranose ring, and another C5 chain substituted at the 3′ position of the furanose ring.

Structure diversity of acylsugar metabolites, in terms of sugar core types, acyl group composition, the various chain length and conservation of specific substitution positions, has been observed not only within each plant species, but also across different plants in the *Solanaceae* including tomato, potato, tobacco and petunia [66-68, 81-86]. For example, in tomato, tri- and tetra- acylsucroses were detected in cultivated tomato *Solanum lycopersicum*

M82, with one acyl group attached to the R3′ position of the furanose ring, while only tri-substituted acylsucroses and acylglucose were observed in wild tomato *S. pennellii,* with all three acyl groups attached to the pyranose ring [37, 66, 67]. Penta-acylsucroses were present in another wild tomato *Solanum habrochaites,* with two acyl groups substituted at furanose ring [67]. On the other hand, petunia species were observed to have both tetra- and penta-acylsucroses, with all four hydroxyl positions on the pyranose ring esterified by acyl chains [68]. Additionally, diversity of acyl chain length was determined across *Solanaceae* species, ranging from C2 to C12 observed in tomato and C4 to C8 detected in petunia, perhaps a result of different biosynthetic pathways active in these species [66-68]. Furthermore, malonyl acyl groups were identified only from acylsucroses from petunia plants, with the position of the malonyl group conserved at the R1′ position of the furanose ring (with two exceptions detected at R6′ in *P. exserta*) [68]. Furthermore, conservation of specific acyl groups at specific positions of the sugar core in *Solanaceae* acylsugars were identified, for example the R3′ position of tomato acylsucroses was only observed esterified as an acetyl ester, and the R6 position of petunia species was only esterified by aiC5 and iC4 groups [67,68], indicating the preference of substrate by acyltransferases that catalyze esterification at specific hydroxyl groups.

1.3.3 Biosynthesis of acylsugar metabolites result in vast chemical diversity

The great chemical diversity of acylsugar metabolites identified within the family *Solanaceae* is generated by a variety of biosynthetic and regulatory enzymes, which take part in complex acylsugar metabolism and control acylsugar abundance at various stages of development glandular trichomes. Several genes and enzymes expressed preferentially in trichomes are involved in catalyzing acylsugar synthesis in cultivated and wild tomato and petunia species

have been characterized in recent studies using *in vitro* enzyme assays and *in vivo* gene silencing [41,42, 87-91] (Figure 1.3). In brief, four acylsugar acyltransferases (ASATs) were identified that sequentially add specific acyl groups to R4, R3, R3′ and R3 hydroxyl positions on sucrose core in cultivated tomato, and are named as Sl-ASAT1-4 according to the order of acylation reactions [87]. Similar enzymes were characterized in the wild tomato *S. pennellii*, demonstrating a "flipped pathway" catalyzed by orthologs Sp-ASAT1, Sp-ASAT3 and Sp-ASAT2 in order, that attach acyl chains to R4, R2, R3 on the pyranose ring [42,90]. In addition, petunia acyltransferases from *P. axillaris* (Pax-ASATs) were discovered that participate in sequential esterifications at R2, R4, R3, and R6 positions on the pyranose ring [88]. Moreover, all characterized ASATs exhibit various degrees of substrate preference, using specific acyl CoA esters as substrates, contributing to acylsucrose chemical diversity arising from esterification by various acyl chains at specific substitution positions. In addition, two acylsugar acylhydrolases (ASH1 and ASH2) were recently described, which are involved in acylsugar turnover by catalyzing hydrolytic removal of acyl groups from R3 and R4 positions of certain tri- and tetra-acylsugars *in vitro* [89], allowing for remodeling of acylsucroses.

Various lengths of acyl groups, as well as their branching types, are other sources for structure diversification of acylsugar metabolites in Solanaceae plants. Different mechanisms elongating branched acyl groups generate chain length variation. It has been proposed that one-carbon elongation will form acyl groups shorter than C8 and longer acyl groups are synthesized through two-carbon elongation pathway as is the case in fatty acid biosynthesis [92,93]. Additionally, mutant versions of isopropylmalate synthase 3 (IPMS3) expressed in type I/IV trichomes in tomato species with a truncated C-terminus were characterized, which not only affect relative abundances of iC4 and iC5 acyl chain containing acylsugar metabolites, but may also lack one-

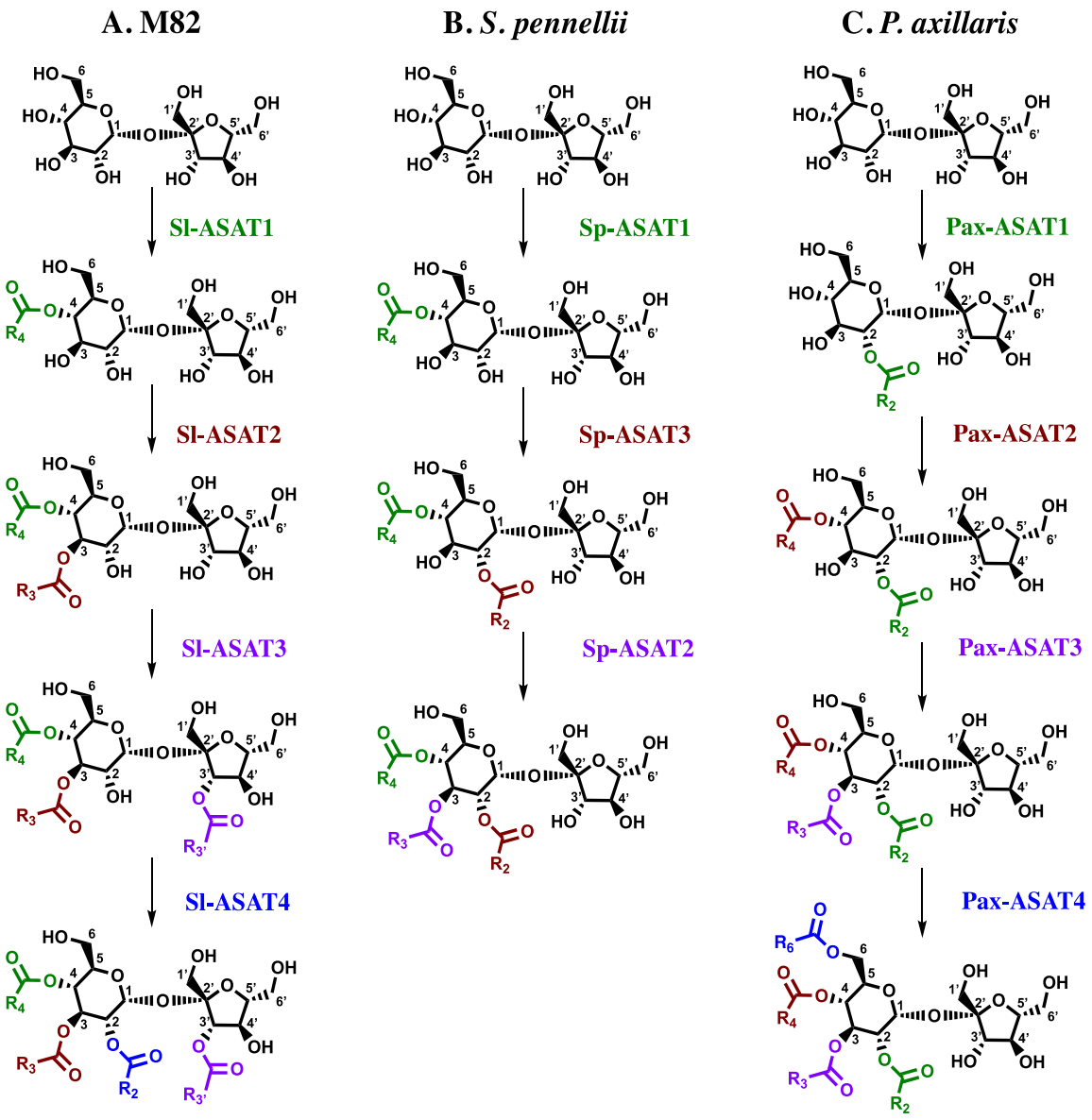


Figure 1.3. Comparison of acylsugar biosynthetic pathways of Solanaceae plants catalyzed by acylsugar acyltransferases (ASATs). *A. Biosynthetic pathway of tomato (M82) catalyzed by Sl-ASATs. B. Biosynthetic pathway of S. pennellii catalyzed by Sp-ASATs. C. Biosynthetic pathway of P. axillaris catalyzed by Pax-ASATs.*

carbon elongation pathway in *S. pennellii* LA0716 and other southern *S. pennellii* accessions [41]. More details about these acylsugar biosynthetic genes and enzymes that contribute to generating structure diversity are discussed in Chapter 4.

1.4 Current challenges in specialized metabolism analysis

Chemical diversity of specialized metabolites within and across plant species presents one of the greatest challenges for specialized metabolism analysis. Multitudinous structures of these natural products provide a wide range of chemical properties, which not only makes isolation and identification of specific compounds more challenging, but also compromises effectiveness of sample preparation protocols, particularly with regard to finding suitable extraction methods for comprehensive analysis of metabolism [94]. Given the lack of commercially available standards for the majority of plant specialized metabolites, structural characterization and quantitative analysis of such diverse chemicals from plant extracts remains costly and time-consuming. Additionally, distinct phenotypes between species usually can be attributed to differences in genotype, and more complete definition of metabolic phenotypes can drive discoveries of novel genes and enzymes. As a result, development of powerful analytical platforms for fast and comprehensive assessment of the diversity of specialized metabolism is a high priority for improving our understanding of complex plant metabolic networks.

The wide range of specialized metabolite concentrations also challenges current analytical strategies applied to analyze plant metabolomes. Several reasons have been proposed to partially explain the huge dynamic range of concentrations of bioactive metabolites maintained within plants, including variations of molecular size, specific functions of particular chemicals and limitation of cellular space [94-96]. It is challenging to use a single analytical technique to analyze metabolites with various abundance levels, while some have attempted to overcome these problems. For instance, a stable isotopic labeling strategy using [^{15}N]KNO_3 as nitrogen source has been introduced to enhance the sensitivity of low abundant metabolites within

Arabidopsis cell for quantitative analysis [97], but this approach is limited to nitrogen-containing compounds.

Another significant challenge presented in specialized metabolism analysis is the dynamic behavior of specialized metabolites during growth and development of plants. Metabolite levels are dynamic, not only because related genes and enzymes exhibit temporal and spatial variation in expression [17], but also as a result of differential abundances of transporter proteins that move metabolites between subcellular compartments, cells and tissues [94]. Furthermore, as the turnover and degradation of metabolites can take place within hours, and sometimes even minutes in plants, specific techniques with the ability to either rapidly quench plant metabolism, or continuously monitor the flux of metabolism are required to complete our knowledge of plant specialized metabolism.

1.5 Current strategies for plant specialized metabolism investigation

1.5.1 Mass spectrometry-based strategies for metabolite profiling and accelerated metabolite annotation

Identification and measurement of specialized metabolites from crude matrices provides the basis of plant phenotype investigations, and is often the starting point for characterization of biochemical activities of novel genes and enzymes. Due to the enormous chemical diversity of metabolites, several hybrid analytical techniques, especially mass spectrometry-based analytical platforms coupled with chromatographic separations (*e.g.* GC-MS, LC-MS), are widely applied for deep profiling and fast annotation of natural products. GC-MS is a well-established analytical technique for profiling of volatile metabolites from complex plant extracts, providing high separation efficiency, and low level detection for annotation of metabolites over a wide range of

concentrations [98]. However, GC-MS analyses usually rely on mass spectrum databases for identification. Despite the growth of the NIST17 database which now includes spectra for more than 250,000 compounds, many specialized metabolites have yet to be discovered, and are not included in databases. Attribution of structure based on manual interpretation of mass spectra alone remains challenging, and is not applicable for analysis of entire metabolomes when unknown metabolites are present. Many specialized metabolites lack volatility needed to elute using gas chromatography, and LC-MS provides a widely used approach suitable for profiling of non-volatile specialized metabolites. Recently, the development of new column chemistries, especially for the introduction of HILIC separation and development of UHPLC techniques, expanded applications of LC-MS technique to a broader range of metabolites with different polarity [99,100]. In order to accelerate annotation of unknown metabolites within a single analysis, fragment ions with useful structural information are generated non-selectively using multiplexed collision-induced dissociation (CID), which usually involves the quasi-simultaneous acquisition of mass spectra with 3-5 different collision voltages applied ahead of the mass analyzer [66-68]. Since it is a non-selective, data-independent strategy, LC-MS with multiplexed CID technique has particular advantages for non-targeted deep profiling of specialized metabolites. Nevertheless, there is a substantial chance that signals of low abundance and coeluting metabolites may be obscured from signals from other high abundance compounds using this approach. To address this limitation, data-dependent and targeted MS/MS analyses have been applied to generate fragments of selected metabolite ions, which significantly reduces mis-assignments signals from the complex background. Moreover, mass spectrometry imaging (MSI) has been introduced as a complementary technique recently into plant metabolism analysis, with the ability to annotate metabolites with spatial localization information within

plant tissues. A recent study demonstrated metabolic profiling of single-cells with chemical spatial distribution information within glandular trichomes of tomato leaflets could be obtained using laser desorption/ionization mass spectrometry imaging [101].

1.5.2 NMR techniques for metabolite structure elucidation

Although multiplexed CID and MS/MS analyses provide some structural information based on molecular and fragment masses, and elemental formulas of each, mass spectrometry alone is not enough to completely elucidate the structures of novel compounds. Nuclear Magnetic Resonance (NMR) spectroscopy has long been considered as an indispensable tool for characterization of natural products. Comparing with other traditional procedure to determine the structure of natural compounds and their derivatives, NMR spectroscopy yields comprehensive information about atomic connections without *a priori* knowledge of skeleton and is considered a non-destructive technique [102]. One-dimensional NMR is commonly used as a first step toward identifying structures of specialized plant metabolites. For acylsugar metabolites, information about chemical shifts and coupling constants helps assignments of each hydrogen and carbon in the sugar core, and to distinguish anomers of acylglucoses [37]. Two-dimensional NMR techniques are especially helpful for establishing connectivities in complex structures. Correlation spectroscopy (COSY) is a technique that determines the correlation between hydrogens on adjacent carbon atoms. Its utility facilitates assignments of hydrogens. Heteronuclear single quantum coherence (HSQC) spectra have cross peaks representing correlation between hydrogens and the carbon atoms to which they are attached, and are especially helpful for distinguishing overlapping proton resonances by demonstrating connections with specific carbon atoms, which are usually more fully resolved Heteronuclear

multiple bond correlation (HMBC) spectra provide cross peaks for the interactions of protons and ^{13}C connected usually within 2-3 bonds, and this information aids final structure assignments. Additionally, other 2D NMR techniques including coupled-HSQC, Nuclear Overhouser Effect Spectroscopy (NOESY), and Total Correlation Spectroscopy (TOCSY) provide complementary information to assess the structural diversity of specialized metabolites. Nevertheless, a major weakness of NMR compared with mass spectrometry is poor relative sensitivity. Without narrow bore cryoprobe instruments, milligram amounts of purified metabolites are necessary for obtaining a complete set of spectra using NMR.

1.5.3 Stable isotope labeling strategies for investigation of metabolite biosynthesis

Isotope labeling strategies, using either radioisotope (such as ^{14}C, ^{3}H) or stable isotope (such as ^{13}C, ^{15}N) tracers, have been applied for metabolic network investigation for a long time [103]. With several order of magnitude improvements in sensitivity of mass spectrometers in recent years, precursors, intermediates, and final products labeled by stable isotope tracers can be distinguished from unlabeled species according to mass shifts due to the incorporation of isotope labels, making mass spectrometry an appropriate approach for studying mechanisms of metabolite biosynthesis. In addition, use of stable isotopes eliminates safety and containment requirements needed to handle radioisotope tracers. Furthermore, identification of radiolabeled metabolites often still requires mass spectrometry, and perhaps NMR, to establish locations of tracers within metabolic products. When coupled with LC-MS and LC-MS/MS analysis, tracing of stable isotope labeled metabolites successfully facilitates confirmation of specific metabolic pathways, identification of novel pathways, as well as measurement of fluxes through metabolic networks [104-106].

1.6 Summary of Research

The long-term goal of this project is to understand the biochemical and genetic mechanisms underlying the tremendous diversity of plant specialized metabolites produced by glandular trichomes in the family *Solanaceae*, and provide a fundamental basis for genetic engineering of plants and heterologous hosts to produce valuable compounds. Towards this goal, several chemical and biological issues need to be investigated, starting with comprehensive identification of specialized metabolites that accumulate in specific plant genotypes, defining similarities and differences within and across plant genotypes in terms of specialized metabolite production, and accelerating explorations into mechanisms that produce specialized metabolites evolve. To answer these questions, novel approaches with combinations of several modern analytical techniques including LC-MS, NMR, chromatographic isolation and stable isotope labeling have been employed for more comprehensive understanding of diversity in plant specialized metabolism.

Chapter 2 of this dissertation explores the diversity of acylsugar specialized metabolites produced within three petunia species which are members of the family *Solanaceae*. Acylsugars have been identified as the dominant class of petunia specialized metabolites based on deep LC/MS profiling of leaf dip extracts, which showed significant structural diversity within and between petunia species. More than 100 distinct molecular masses of acylsugar metabolites have been annotated using a modified 80-min LC-MS analysis. Of these, 28 have been successfully purified through solid phase extraction followed by semi-preparative scale HPLC separation, and structures are proposed according to NMR analysis. Malonyl ester acylsucroses have been first characterized as metabolites in all three petunia species, with one malonyl group esterified at either R1′ or R6′ positions of the furanose ring. Interestingly, all acylsucroses characterized from

petunia species have not previously been identified from previous studies of other *Solanum* species including petunia hybrids, which suggests new functions for genes and enzymes involved in their biosynthesis.

In recent years, advances in next-generation DNA sequencing have made whole genome and whole transcriptome analyses more rapid than structure elucidation for individual novel natural products. With increasing needs for rapid characterization of plant chemical diversity to keep pace with the genomic analysis, the importance of establishing relevant high-throughput analytical tools has drawn more attention. In Chapter 3, the newly developed LC-MS/MS analysis of metabolites from plants grown on ^{13}C-labeled amino acid precursors provided a simplified strategy to investigate the structural diversity of plant specialized metabolites without need for metabolite purification. A significant portion of acylsugar chemical diversity arising from the diversification of acyl groups, particularly in terms of different chain lengths and branching types. Branched chain amino acids (BCAAs, including Val, Leu and Ile) can be used as precursors to generate various acyl groups involved in acylsugar biosynthesis. Using uniformly labeled [^{13}C]-BCAAs as isotope tracers to feed plant seedlings, acyl chains with either iso- or anteiso- branching types could be labeled from labeled Val, Leu or Ile precursors separately, and the isotope incorporations were analyzed using LC/time-of-flight mass spectrometry (LC/TOF-MS) using both multiplexed CID for concentrated metabolites and MS/MS for low abundant compounds. The diversity of acyl groups was observed within *Solanaceae* plants including cultivated and wild tomato and petunia, with the presence of an unsaturated aiC5 chain from cultivated tomato acylsugar that has not been demonstrated during previous profiling of *Solanum* plants, perhaps due to its low concentration. Moreover, structural diversity of acylsugar isomeric metabolites was investigated using this approach. With the

knowledge of substrate preferences of four acylsugar acyltransferases detected in cultivated tomato, this hybrid strategy can be applied to propose detailed structural information of acylsugar metabolites without NMR.

Using similar stable isotope labeling coupled with the LC-MS/MS strategy described in Chapter 3, the Chapter 4 in this dissertation addresses the question of how do plants in the *Solanaceae* make such a huge diversity of acylsugar metabolites, particularly in terms of the variety of aliphatic chains. One of the direct incentives of this study was the discovery of several aiC6-containing acylsugar metabolites from wild tomato *S. pennellii* accession LA2560. The anteiso-branching of these C6 groups has been confirmed by NMR analysis. However, previously reported aliphatic group elongation mechanisms used by *S. pennellii* accessions do not explain the presence of aiC6 in acylsugars. On the other hand, as the recent study of ASATs in cultivated tomato documented the detailed acylsugar biosynthetic mechanism by sequential addition of acyl-CoA substrates onto specific positions of a sucrose core, a missing piece of the puzzle of acylsugar biosynthesis lies with the mechanisms that generate diverse acyl-CoA substrate pools across the *Solanaceae*. To answer these questions, acyl chain fragments of acylsugars generated with and without stable isotope labels derived from [^{13}C]-BCAA feeding of plants were analyzed using LC-MS/MS, and related aliphatic chain elongation pathways were proposed based on the resulting information.

REFERENCES

REFERENCES

1. Whitman, D. B., Genetically modified foods: harmful or helpful?. *CSA Discovery Guides* **2000**, 1-13.

2. Hoisington, D.; Khairallah, M.; Reeves, T.; Ribaut, J.-M.; Skovmand, B.; Taba, S.; Warburton, M., Plant genetic resources: What can they contribute toward increased crop productivity?. *Proceedings of the National Academy of Sciences of the USA* **1999**, *96* (11), 5937-5943.

3. Milo, R.; Last, R. L., Achieving diversity in the face of constraints: lessons from metabolism. *Science* **2012**, *336* (6089), 1663-1667.

4. Bomgardner, M. M., Replacing trans fat. *Chemical & Engineering News* **2012**, *90* (11), 30-32.

5. Kramer, M. G.; Redenbaugh, K., Commercialization of a tomato with an antisense polygalacturonase gene: The FLAVR SAVR™ tomato story. *Euphytica* **1994**, *79* (3), 293-297.

6. Vaeck, M.; Reynaerts, A.; Höfte, H.; Jansens, S.; De Beuckeleer, M.; Dean, C.; Zabeau, M.; Montagu, M. V.; Leemans, J., Transgenic plants protected from insect attack. *Nature* **1987**, *328* (6125), 33-37.

7. Schijlen, E.; de Vos, R. C. H.; Jonker, H.; Van Den Broeck, H.; Molthoff, J.; Van Tunen, A.; Martens, S.; Bovy, A., Pathway engineering for healthy phytochemicals leading to the production of novel flavonoids in tomato fruit. *Plant Biotechnology Journal* **2006**, *4* (4), 433-444.

8. Bolton, G. W., Plant phenolic compounds induce expression of the *Agrobacterium tumefaciens* loci needed for virulence. *Science* **1986**, *232*, 983-986.

9. Steppuhn, A.; Gase, K.; Krock, B.; Halitschke, R.; Baldwin, I. T., Nicotine's defensive function in nature. *PLoS Biology* **2004**, *2* (8), E217.

10. Papadopoulou, K.; Melton, R. E.; Leggett, M.; Daniels, M. J.; Osbourn, A. E., Compromised disease resistance in saponin-deficient plants. *Proceedings of the National Academy of Sciences of the USA* **1999**, *96* (22), 12923-12928.

11. Saito, K.; Matsuda, F., Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology* **2010**, *61*, 463-489.

12. Pichersky, E.; Lewinsohn, E., Convergent evolution in plant specialized metabolism. *Annual Review of Plant Biology* **2011**, *62*, 549-566.

13. Wink, M., Introduction: biochemistry, physiology and ecological functions of secondary metabolites. *Annual Plant Reviews Volume 40: Biochemistry of Plant Secondary Metabolism, Second Edition* **2010**, 1-19.

14. Fraenkel, G. S., The raison d'etre of secondary plant substances. *Science* **1959**, *129* (3361), 1466-1470.

15. Hartmann, T., The lost origin of chemical ecology in the late 19th century. *Proceedings of the National Academy of Sciences of the USA* **2008**, *105* (12), 4541-4546.

16. Pichersky, E.; Noel, J. P.; Dudareva, N., Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science* **2006**, *311* (5762), 808-811.

17. Pichersky, E.; Gang, D. R., Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends in Plant Science* **2000**, *5* (10), 439-445.

18. Dixon, R. A.; Strack, D., Phytochemistry meets genome analysis, and beyond. *Phytochemistry* **2003**, *62*, 815-816

19. Cragg, G. M.; Newman, D. J., Plants as a source of anti-cancer agents. *Journal of Ethnopharmacology* **2005**, *100* (1), 72-79.

20. Duke, S. O.; Canel, C.; Rimando, A. M.; Telle, M. R.; Duke, M. V.; Paul, R. N., Current and potential exploitation of plant glandular trichome productivity. *Advances in Botanical Research* **2000**, *31*, 121-151.

21. Hartmann, T., From waste products to ecochemicals: fifty years research of plant secondary metabolism. *Phytochemistry* **2007**, *68* (22), 2831-2846.

22. Atanasov, A. G.; Waltenberger, B.; Pferschy-Wenzig, E.-M.; Linder, T.; Wawrosch, C.; Uhrin, P.; Temml, V.; Wang, L.; Schwaiger, S.; Heiss, E. H., Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnology Advances* **2015**, *33* (8), 1582-1614.

23. Newman, D. J.; Cragg, G. M., Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products* **2012**, *75* (3), 311-335.

24. Kinghorn, A. D.; Pan, L.; Fletcher, J. N.; Chai, H., The relevance of higher plants in lead compound discovery programs. *Journal of Natural Products* **2011**, *74* (6), 1539-1555.

25. Jones A., Chemistry: an introduction for medical and health sciences. *John Wiley & Sons*, **2005,** 5-6

26. Wall, M. E.; Wani, M. C.; Cook, C. E.; Palmer, K. H.; McPhail, A. A.; Sim, G. A., Plant antitumor agents. I. The isolation and structure of camptothecin, a novel alkaloidal leukemia and tumor inhibitor from *Camptotheca acuminata*. *Journal of the American Chemical Society* **1966**, *88* (16), 3888-3890.

27. Mann, J., Natural products in cancer chemotherapy: past, present and future. *Nature Reviews Cancer* **2002**, *2* (2), 143-148.

28. Björkman, E., *Monotropa hypopitys* L.—an epiparasite on tree roots. *Physiologia Plantarum* **1960**, *13* (2), 308-327.

29. Nakabayashi, R.; Saito, K., Integrated metabolomics for abiotic stress responses in plants. *Current Opinion in Plant Biology* **2015**, *24*, 10-16.

30. Moghe, G. D.; Last, R. L., Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiology* **2015**, *169* (3), 1512-1523.

31. Ehrlich, P. R.; Raven, P. H., Butterflies and plants: a study in coevolution. *Evolution* **1964**, *18* (4), 586-608.

32. Jones, C. G.; Firn, R. D.; Malcolm, S. B., On the evolution of plant secondary chemical diversity. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **1991**, *333* (1267), 273-280.

33. Theis, N.; Lerdau, M., The evolution of function in plant secondary metabolites. *International Journal of Plant Sciences* **2003**, *164* (S3), S93-S102.

34. Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirkse, W.; Van Staveren, M.; Stiekema, W., Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana. *Nature* **1998**, *391* (6666), 485-488.

35. Somerville, C.; Somerville, S., Plant functional genomics. *Science* **1999**, *285* (5426), 380-383.

36. Facchini, P. J., Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation, and metabolic engineering applications. *Annual Review of Plant Biology* **2001**, *52* (1), 29-66.

37. Ghosh, B.; Jones, A. D., Profiling, characterization, and analysis of natural and synthetic acylsugars (sugar esters). *Analytical Methods* **2017**, *9* (6), 892-905.

38. Mizutani, M.; Ohta, D., Diversification of P450 genes during land plant evolution. *Annual Review of Plant Biology* **2010**, *61*, 291-315.

39. Falara, V.; Akhtar, T. A.; Nguyen, T. T. H.; Spyropoulou, E. A.; Bleeker, P. M.; Schauvinhold, I.; Matsuba, Y.; Bonini, M. E.; Schilmiller, A. L.; Last, R. L., The tomato terpene synthase gene family. *Plant Physiology* **2011**, *157* (2), 770-789.

40. D'Auria, J. C., Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **2006**, *9* (3), 331-340.

41. Ning, J.; Moghe, G. D.; Leong, B.; Kim, J.; Ofner, I.; Wang, Z.; Adams, C.; Jones, A. D.; Zamir, D.; Last, R. L., A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant Physiology* **2015**, *169* (3), 1821-1835.

42. Fan P.; Miller A. M.; Liu X.; Jones A. D.; Last R. L., Metabolic innovation through evolutionary pathway engineering in glandular hairs of tomato. *Submitted to Nature Communications* **2017**

43. Strack, D.; Vogt, T.; Schliemann, W., Recent advances in betalain research. *Phytochemistry* **2003**, *62* (3), 247-269.

44. Tanaka, Y.; Sasaki, N.; Ohmiya, A., Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *The Plant Journal* **2008**, *54* (4), 733-749.

45. Köllner, T. G.; Lenk, C.; Zhao, N.; Seidl-Adams, I.; Gershenzon, J.; Chen, F.; Degenhardt, J., Herbivore-induced SABATH methyltransferases of maize that methylate anthranilic acid using S-adenosyl-L-methionine. *Plant Physiology* **2010**, *153* (4), 1795-1807.

46. Köllner, T. G.; Schnee, C.; Gershenzon, J.; Degenhardt, J., The variability of sesquiterpenes emitted from two Zea mays cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *The Plant Cell* **2004**, *16* (5), 1115-1131.

47. Kulma, A.; Szopa, J., Catecholamines are active compounds in plants. *Plant Science* **2007**, *172* (3), 433-440.

48. Tan, D.-X.; Hardeland, R.; Manchester, L. C.; Korkmaz, A.; Ma, S.; Rosales- Corral, S.; Reiter, R. J., Functional roles of melatonin in plants, and perspectives in nutritional and agricultural science. *Journal of Experimental Botany* **2011**, *63* (2), 577-597.

49. Fahn, A., Structure and function of secretory cells. *Advances in Botanical Research* **2000**, *31*, 37-75.

50. Kang, J.-H.; Shi, F.; Jones, A. D.; Marks, M. D.; Howe, G. A., Distortion of trichome morphology by the hairless mutation of tomato affects leaf surface chemistry. *Journal of Experimental Botany* **2009**, *61* (4), 1053-1064.

51. Kennedy, G. G., Tomato, pests, parasitoids, and predators: tritrophic interactions involving the genus *Lycopersicon*. *Annual Review of Entomology* **2003**, *48* (1), 51-72.

52. Shepherd, R. W.; Bass, W. T.; Houtz, R. L.; Wagner, G. J., Phylloplanins of tobacco are defensive proteins deployed on aerial surfaces by short glandular trichomes. *The Plant Cell* **2005**, *17* (6), 1851-1861.

53. Wagner, G. J.; Wang, E.; Shepherd, R. W., New approaches for studying and exploiting an old protuberance, the plant trichome. *Annals of Botany* **2004**, *93* (1), 3-11.

54. Luckwill, L. C., The genus *Lycopersicon*. *Aberdeen Univ. Studies* **1943**, (120).

55. Gershenzon, J.; Dudareva, N., The function of terpene natural products in the natural world. *Nature Chemical Biology* **2007**, *3* (7), 408-414.

56. Fridman, E.; Wang, J.; Iijima, Y.; Froehlich, J. E.; Gang, D. R.; Ohlrogge, J.; Pichersky, E., Metabolic, genomic, and biochemical analyses of glandular trichomes from the wild tomato species *Lycopersicon hirsutum* identify a key enzyme in the biosynthesis of methylketones. *The Plant Cell* **2005**, *17* (4), 1252-1267.

57. Treutter, D., Significance of flavonoids in plant resistance: a review. *Environmental Chemistry Letters* **2006**, *4* (3), 147-157.

58. Gang, D. R.; Wang, J.; Dudareva, N.; Nam, K. H.; Simon, J. E.; Lewinsohn, E.; Pichersky, E., An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiology* **2001**, *125* (2), 539-555.

59. Bleeker, P. M.; Diergaarde, P. J.; Ament, K.; Schütz, S.; Johne, B.; Dijkink, J.; Hiemstra, H.; de Gelder, R.; de Both, M. T. J.; Sabelis, M. W., Tomato-produced 7-epizingiberene and R-curcumene act as repellents to whiteflies. *Phytochemistry* **2011**, *72* (1), 68-73.

60. Williams, W. G.; Kennedy, G. G.; Yamamoto, R. T.; Thacker, J. D.; Bordner, J., 2-Tridecanone: a naturally occurring insecticide from the wild tomato *Lycopersicon hirsutum f. glabratum*. *Science* **1980**, *207* (4433), 888-889.

61. Tattini, M.; Gravano, E.; Pinelli, P.; Mulinacci, N.; Romani, A., Flavonoids accumulate in leaves and glandular trichomes of *Phillyrea latifolia* exposed to excess solar radiation. *The New Phytologist* **2000**, *148* (1), 69-77.

62. Schmidt, A.; Li, C.; Shi, F.; Jones, A. D.; Pichersky, E., Polymethylated myricetin in trichomes of the wild tomato species *Solanum habrochaites* and characterization of trichome-specific 3′/5′-and 7/4′-myricetin O-methyltransferases. *Plant Physiology* **2011**, *155* (4), 1999-2009.

63. Tan, K.-H.; Nishida, R.; Toong, Y.-C., Floral synomone of a wild orchid, *Bulbophyllum cheiri*, lures Bactrocera fruit flies for pollination. *Journal of Chemical Ecology* **2002**, *28* (6), 1161-1172.

64. Arrendale, R. F.; Severson, R. F.; Sisson, V. A.; Costello, C. E.; Leary, J. A.; Himmelsbach, D. S.; Van Halbeek, H., Characterization of the sucrose ester fraction from *Nicotiana glutinosa*. *Journal of Agricultural and Food Chemistry* **1990**, *38* (1), 75-85.

65. Slocombe, S. P.; Schauvinhold, I.; McQuinn, R. P.; Besser, K.; Welsby, N. A.; Harper, A.; Aziz, N.; Li, Y.; Larson, T. R.; Giovannoni, J., Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiology* **2008**, *148* (4), 1830-1846.

66. Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Jones, A. D.; Last, R. L., Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. *The Plant Journal* **2010**, *62* (3), 391-403.

67. Ghosh, B.; Westbrook, T. C.; Jones, A. D., Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014**, *10* (3), 496-507.

68. Liu, X.; Enright, M.; Barry, C. S.; Jones, A. D., Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of Petunia species. *Metabolomics* **2017**, *13* (7), 85.

69. Goffreda, J. C.; Steffens, J. C.; Mutschler, M. A., Association of epicuticular sugars with aphid resistance in hybrids with wild tomato. *Journal of the American Society for Horticultural Science* **1990**, *115* (1), 161-165.

70. Weinhold, A.; Baldwin, I. T., Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. *Proceedings of the National Academy of Sciences of the USA* **2011**, *108* (19), 7855-7859.

71. Glas, J. J.; Schimmel, B. C. J.; Alba, J. M.; Escobar-Bravo, R.; Schuurink, R. C.; Kant, M. R., Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences* **2012**, *13* (12), 17077-17103.

72. Leckie, B. M.; De Jong, D. M.; Mutschler, M. A., Quantitative trait loci increasing acylsugars in tomato breeding lines and their impacts on silverleaf whiteflies. *Molecular Breeding* **2012**, *30* (4), 1621-1634.

73. Leckie, B. M.; Halitschke, R.; De Jong, D. M.; Smeda, J. R.; Kessler, A.; Mutschler, M. A., Quantitative trait loci regulating the fatty acid profile of acylsugars in tomato. *Molecular Breeding* **2014**, *34* (3), 1201-1213.

74. Dembitsky, V. M., Astonishing diversity of natural surfactants: 1. Glycosides of fatty acids and alcohols. *Lipids* **2004**, *39* (10), 933-953.

75. Garti, N.; Clement, V.; Fanun, M.; Leser, M. E., Some characteristics of sugar ester nonionic microemulsions in view of possible food applications. *Journal of Agricultural and Food Chemistry* **2000**, *48* (9), 3945-3956.

76. Baldwin, E. A.; Nisperos-Carriedo, M. O.; Baker, R. A., Edible coatings for lightly processed fruits and vegetables. *HortScience* **1995**, *30* (1), 35-38.

77. Lerk, P. C.; Sucker, H. H.; Eicke, H. F., Micellization and solubilization behavior of sucrose laurate, a new pharmaceutical excipient. *Pharmaceutical Development and Technology* **1996**, *1* (1), 27-36.

78. Chortyk, O. T.; Pomonis, J. G.; Johnson, A. W., Syntheses and characterizations of insecticidal sucrose esters. *Journal of Agricultural and Food Chemistry* **1996**, *44* (6), 1551-1557.

79. McKenzie, C. L.; Puterka, G. J., Effect of sucrose octanoate on survival of nymphal and adult *Diaphorina citri* (*Homoptera*: *Psyllidae*). *Journal of Economic Entomology* **2004**, *97* (3), 970-975.

80. Rodriguez, A. E.; Tingey, W. M.; Mutschler, M. A., Acylsugars of *Lycopersicon pennellii* deter settling and feeding of the green peach aphid (*Homoptera: Aphididae*). *Journal of Economic Entomology* **1993**, *86* (1), 34-39.

81. King, R. R.; Calhoun, L. A., 6 2, 3-Di-O-and 1, 2, 3-tri-O-acylated glucose esters from the glandular trichomes of *Datura metel*. *Phytochemistry* **1988**, *27* (12), 3761-3763.

82. Slocombe, S. P.; Schauvinhold, I.; McQuinn, R. P.; Besser, K.; Welsby, N. A.; Harper, A.; Aziz, N.; Li, Y.; Larson, T. R.; Giovannoni, J., Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiology* **2008**, *148* (4), 1830-1846.

83. King, R. R.; Singh, R. P.; Boucher, A., Variation in sucrose esters from the type B glandular trichomes of certain wild potato species. *American Journal of Potato Research* **1987**, *64* (10), 529-534.

84. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Sucrose esters associated with glandular trichomes of wild *Lycopersicon* species. *Phytochemistry* **1990**, *29* (7), 2115-2118.

85. King, R. R.; Singh, R. P.; Calhoun, L. A., Elucidation of structures for a unique class of 2, 3, 4, 3′-tetra-O-acylated sucrose esters from the type B glandular trichomes of *Solanum neocardenasii* Hawkes & Hjerting (PI 498129). *Carbohydrate Research* **1988**, *173* (2), 235-241.

86. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Characterization of 2, 3, 4, 3′-tetra-O-acylated sucrose esters associated with the glandular trichomes of *Lycopersicon typicum*. *Journal of Agricultural and Food Chemistry* **1993**, *41* (3), 469-473.

87. Fan, P.; Miller, A. M.; Schilmiller, A. L.; Liu, X.; Ofner, I.; Jones, A. D.; Zamir, D.; Last, R. L., In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. *Proceedings of the National Academy of Sciences of the USA* **2016**, *113* (2), E239-E248.

88. Nadakuduti S. S.; Uebler J. B.; Liu X.; Jones A. D.; Barry C. S., Characterization of trichome-expressed BAHD acyltransferases in *Petunia axillaris* reveals distinct acylsugar assembly mechanisms within the Solanaceae. *Submitted to Plant Physiology* **2017**

89. Schilmiller, A. L.; Gilgallon, K.; Ghosh, B.; Jones, A. D.; Last, R. L., Acylsugar acylhydrolases: carboxylesterase-catalyzed hydrolysis of acylsugars in tomato trichomes. *Plant Physiology* **2016**, *170* (3), 1331-1344.

90. Schilmiller, A. L.; Moghe, G. D.; Fan, P.; Ghosh, B.; Ning, J.; Jones, A. D.; Last, R. L., Functionally divergent alleles and duplicated loci encoding an acyltransferase contribute to acylsugar metabolite diversity in Solanum trichomes. *The Plant Cell* **2015**, *27* (4), 1002-1017.

91. Schilmiller, A. L.; Charbonneau, A. L.; Last, R. L., Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proceedings of the National Academy of Sciences of the USA* **2012**, *109* (40), 16377-16382.

92. Kroumova, A. B.; Wagner, G. J., Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants. *Planta* **2003**, *216* (6), 1013-1021.

93. Kroumova, A. B.; Xie, Z.; Wagner, G. J., A pathway for the biosynthesis of straight and branched, odd-and even-length, medium-chain fatty acids in plants. *Proceedings of the National Academy of Sciences of the USA* **1994**, *91* (24), 11437-11441.

94. Fernie, A. R., The future of metabolic phytochemistry: larger numbers of metabolites, higher resolution, greater understanding. *Phytochemistry* **2007**, *68* (22), 2861-2880.

95. Fell, D.; Cornish-Bowden, A., Understanding the control of metabolism. *Portland press London*: **1997**; Vol. 2.

96. Sweetlove, L. J.; Fernie, A. R., Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. *New Phytologist* **2005**, *168* (1), 9-24.

97. Engelsberger, W. R.; Erban, A.; Kopka, J.; Schulze, W. X., Metabolic labeling of plant cell cultures with $K^{15}NO_3$ as a tool for quantitative analysis of proteins and metabolites. *Plant Methods* **2006**, *2* (1), 14.

98. Jorge, T. F.; Rodrigues, J. A.; Caldana, C.; Schmidt, R.; van Dongen, J. T.; Thomas-Oates, J.; António, C., Mass spectrometry- based plant metabolomics: Metabolite responses to abiotic stress. *Mass Spectrometry Reviews* **2016**, *35* (5), 620-649.

99. Tolstikov, V. V.; Fiehn, O., Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry. *Analytical Biochemistry* **2002**, *301* (2), 298-307.

100. Eugster P. J., Guillarme D., Rudaz S., et al. Ultra high pressure liquid chromatography for crude plant extract profiling. *Journal of AOAC International*, **2011**, *94*(1): 51-70.

101. Li, C.; Wang, Z.; Jones, A. D., Chemical imaging of trichome specialized metabolites using contact printing and laser desorption/ionization mass spectrometry. *Analytical and Bioanalytical Chemistry* **2014**, *406* (1), 171-182.

102. Bingol, K.; Brüschweiler, R., Knowns and unknowns in metabolomics identified by multidimensional NMR and hybrid MS/NMR methods. *Current Opinion in Biotechnology* **2017**, *43*, 17-24.

103. Lehmann, W. D., A timeline of stable isotopes and mass spectrometry in the life sciences. *Mass Spectrometry Reviews* **2017**, *36* (1), 58-85.

104. Wang, Z.; Jones, A. D., Profiling of stable isotope enrichment in specialized metabolites using liquid chromatography and multiplexed nonselective collision-induced dissociation. *Analytical Chemistry* **2014**, *86* (21), 10600-10607.

105. Creek, D. J.; Chokkathukalam, A.; Jankevics, A.; Burgess, K. E. V.; Breitling, R.; Barrett, M. P., Stable isotope-assisted metabolomics for network-wide metabolic pathway elucidation. *Analytical Chemistry* **2012**, *84* (20), 8442-8447.

106. Meissen, J. K.; Pirman, D. A.; Wan, M.; Miller, E.; Jatkar, A.; Miller, R.; Steenwyk, R. C.; Blatnik, M., Phenotyping hepatocellular metabolism using uniformly labeled carbon-13 molecular probes and LC-HRMS stable isotope tracing. *Analytical Biochemistry* **2016**, *508*, 129-137.

Chapter Two.

Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of Petunia species

This Chapter presents an extended version of research documented in the following journal article: X. Liu, M. Enright, C. S. Barry and A. D. Jones, Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of Petunia species. *Metabolomics* 13: 85 (2017); DOI: 10.1007/s11306-017-1224-9

2.1 Introduction

The plant kingdom synthesizes hundreds of thousands of diverse specialized metabolites that play key roles in chemical defenses against herbivores and resilience to environmental stress [1,2]. Some find commercial value as pharmaceuticals, fragrances, food additives and natural pesticides [3,4]. Additionally, since biosynthetic pathways of specialized metabolites have been modified by mutation and natural selection, plants function as efficient "chemical factories" to produce bioactive chemicals [5]. In the family *Solanaceae*, which includes numerous food and ornamental crops including tomato, potato, petunia, tobacco, eggplant, and pepper, hair-like epidermal appendages known as glandular trichomes (GTs) [6] play critical roles in synthesis, storage and secretion of bioactive specialized metabolites [7]. GTs can be selected as targets for plant metabolic engineering to increase resistance to insects [8]. Among trichome-enriched specialized metabolites, acylsugars have drawn broad attention owing to their importance in plant resistance against numerous insects [9,10,11]. In Solanaceous plants, acylsugars have been documented as glucose or sucrose cores esterified by several C2-C12 aliphatic acids [12,13,14]. Previous work showed that diversity in acylsugars, which derives from variation in acyl chain lengths, aliphatic chain branching, number of ester groups and positions of esterification, affects insecticidal properties [15,16]. An improved understanding of the genetic controls of acylsugar biosynthesis is essential to manipulate plant metabolic engineering [17,18], and advances in this area would be accelerated by a more complete understanding of the structures of acylsugars produced in nature.

Methods for acylsugar characterization were described in a recent review [19]. Recent efforts to identify acylsugars produced by plants in the family *Solanaceae* have relied on mass

spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy of purified metabolites [12, 17, 20]. Ultra-High Performance Liquid Chromatography (UHPLC) and time-of-flight Mass Spectrometry (TOF-MS) serve as powerful tools for metabolomic profiling [21]. Individual specialized metabolites are resolved by LC separation, recognized as acylsugars from their mass spectra using relative mass defect (RMD) filtering [22] and annotated based on retention times and accurate molecular and fragment ion masses. In addition, MS-based metabolite profiling of multiple genotypes builds connections between phenotype and genotype that facilitate discovery of gene and enzyme functions involved in specialized metabolite biosynthesis [23]. In wild tomato, a BAHD acetyltransferase SlAT2 (recently renamed SlASAT4) specifically catalyzes acetylation on the 2-position of acylsucroses, and was identified based on acylsugar profiles of cultivated tomato and *S. lycopersicum* X *S. pennellii* introgression lines [24]. A recent and extensive investigation identified four BAHD acyltransferases that account for all acylation reactions in the biosynthesis of tomato acylsucroses [25]. These efforts relied on a detailed understanding of acylsugar structures that describes both the diversity of acyl groups and their attachment positions on the carbohydrate core.

Current understanding of *Petunia* acylsugar biosynthesis is based on limited reports largely focused on metabolites extracted from hybrids generated from crosses of *Petunia* species [26, 27] including pooled tissues from 114 *Petunia* hybrid cultivars [15]. One notable recent paper profiled acyl groups released from acylsugar-containing extracts from five *Petunia* hybrids, and found a predominance of malonate and straight-chain aliphatic acyl groups ranging from 2-8 carbon atoms [28]. We are aware of one report of structures of two acylsucroses isolated from *Petunia* species other than hybrids [29]. A small number (~9) of *Petunia* acylsucroses have been documented with exact structure information supported by NMR of purified or pooled

metabolites [26,29] with acyl substitutions detected on 2, 3, 4, 6-positions on the pyranose ring of sucrose by C5-C8 aliphatic acyl groups and 1´, 4´, and 6´-positions on the furanose ring by acetyl and malonyl chains. Unique triacylsucroses with pyranose acylation on 2, 3, 4-positions with either iC6 or iC7 groups were derived from *P. nyctaginiflora* (synonymous with *P. axillaris*; [29]. However, much characterization of *Petunia* acylsugar has been performed on incompletely purified mixtures. Acylglucoses from *P. hybrida* were also reported but positions of specific acyl groups were not established [15]. Deeper and more detailed structural information, particularly with regard to selectivity of acylations at specific positions, is necessary to establish roles of position- and substrate-selectivity of acyltransferase enzymes and how these evolved to control acylsugar accumulation.

In this report, we profiled acylsugar metabolites derived from three *Petunia* species, *P. integrifolia*, *P. axillaris*, and *P. exserta* using UHPLC/MS, and demonstrated acylsugar chemical diversity within species based on structures generated from NMR spectra of purified metabolites. The broader goal of this investigation is to support discoveries of biosynthetic genes involved in specialized metabolite accumulation.

2.2 Materials and methods

2.2.1 Plant material preparation and extraction

Plants (*P. integrifolia*, *P. exserta* and *P. axillaris*) were germinated from seeds and grown in a greenhouse. Additional plant growth metadata are provided in Appendix Table 1. For acylsugar profiling, leaflets of each were extracted with isopropanol:acetonitrile:water (3:3:2, *v/v/v*). For acylsugar profiling, two terminal leaflets of *P. integrifolia*, *P. exserta* and *P. axillaris* were

sampled from 8-week-old plants and extracted by dipping into 2 mL of isopropanol: acetonitrile: water (3:3:2, v/v/v) for 2 min. Three biological replicates were used for each species. Extracts were transferred to microcentrifuge tubes and evaporated to dryness under vacuum using a SpeedVac (no heat applied). Residues were redissolved using 500 µL acetonitrile: water (4:1, v/v), and the tubes were vortexed for 30 s and centrifuged for 10 min (25 °C, 10000x*g*). A 200-µL aliquot of each supernatant was transferred to an autosampler vial fitted with a low-volume glass insert. A 2.0-µL aliquot of a 500 µM stock solution of telmisartan was added to each vial to deliver 5 µM telmisartan as internal standard and lock mass reference (*m/z* 513.2296 in negative mode; *m/z* 515.2442 in positive mode). Extracts were stored at -20 °C for UHPLC/MS analyses.

2.2.2 UHPLC/MS analyses

Metabolite profiling was performed using a Waters LCT Premier^{TM} time-of-flight mass spectrometer coupled with Shimadzu LC-20AD pumps. Extracts were resolved using UHPLC separation on an Ascentis Express C18 UHPLC column (10 cm×2.1 mm; 2.7 µm; Supelco, USA) using acetonitrile as mobile phase B, and aqueous 10 mM ammonium formate, adjusted to pH 2.83 with formic acid, as mobile phase A. The flow rate was set at 0.3 mL/min, column temperature was 40 °C and the total chromatographic run time was 80 min. The linear gradient was: hold at 1% B for 0-1 min, linear gradient to 40% B at 5 min, linear gradients to 70% B at 70 min and 100% B at 75 min, a hold at 100%B until 78 min, and a return to 1%B at 78.01 min followed by a hold until 80 min. Injection of 10 µL was performed using a SIL-5000 autosampler. Mass spectra were acquired using electrospray ionization in both positive- and negative-ion modes with multiplexed CID using W ion optics (resolution ~ 8000). Data

acquisition was performed using 0.1 s for each function over *m/z* 50 to 1500 with dynamic range extension. Five parallel functions were applied by switching Aperture 1 voltage between 10, 25, 40, 55, and 80 V. Several source parameters were: capillary voltage -2.5 kV in negative mode and +3.0 kV in positive mode, sample cone voltage 10 V, desolvation temperature 350 °C, source temperature 100 °C, cone gas flow 40 L/h and desolvation gas flow 400 L/h. MassLynx and QuanLynx software (v. 4.1, Waters Inc.) were used to process mass spectra data.

2.2.3 Purification of acylsugars from *Petunia* species

Acidic malonate ester acylsugars in *P. integrifolia, P. exserta* and *P. axillaris* were separated from neutral acylsugars using Supelclean^{TM} LC-SAX SPE tubes with Cl^- counter ion (3 mL, bed wt. 500 mg). Subsequent purification of individual acylsugars was achieved using HPLC.

Metabolites extraction for purification

Fourteen-week old *P. integrifolia, P. exserta* and *P. axillaris* plants were extracted for acylsugar purification separately. About 200 mature leaflets were collected from each Petunia species and extracted in 1 L of methanol for 2 min. Solvent was evaporated from each extract using a rotary evaporator without application of heat to the water bath. The residue was redissolved in 5 mL acetonitrile: water (4:1, v/v) with ultrasonication for 10 min and transferred to a 15-mL polypropylene centrifuge tube. After centrifugation (10000x*g*, 10 min, 25 °C), the supernatant was stored in glass vials until SPE fractionation.

Acylsugar purification method from *P. integrifolia*

A solid phase extraction (SPE) cartridge (Supelclean^{TM} LC-SAX, Cl^- counter ion, 3 mL, bed wt. 500 mg) was conditioned with 3 mL each of MilliQ water, methanol and conditioning solvent

(0.1 mM aqueous ammonium formate, pH 6.5) in sequence. A 2-mL aliquot Petunia extract that had been diluted 10-fold with conditioning solvent (see below) was loaded onto the column. Neutral acylsugars were eluted using 3 mL washing solvent I: 40% methanol/60% 0.5 mM aqueous ammonium formate (adjusted to pH 6.0 with ammonium hydroxide). An additional wash of the column with 2 mL washing solvent II: 50% methanol/50% 0.5 mM aqueous ammonium formate (adjusted to pH 6.0 with ammonium hydroxide) was performed to remove residual neutral metabolites. Malonate acylsugars were then eluted using 3 mL 80% methanol/20% aqueous 0.5 mM ammonium formate (adjusted to pH 2.0 with formic acid). Finally, the SPE cartridge was rinsed with 4 mL of 100% methanol. Acylsugar-containing fractions were evaporated to dryness at room temperature under vacuum in a Thermo Savant SpeedVac. Dried residues were redissolved immediately before further purification using HPLC.

Further purification of individual acylsugars was performed using a Waters 2795 HPLC Separations Module and a Dionex Acclaim 120 C18 semi-preparative HPLC column (15 cm × 4.6 mm; 5 μm particles). Solvent flow rate for all purification methods was 1.5 mL/min and sample injected volume was 200 μL. Fractions were collected every minute by an LKB fraction collector. Purity of acylsugars in particular fractions was confirmed using HPLC/MS on a Waters LCT Premier coupled to a Shimadzu LC-20AD pump.

Components of neutral acylsugar fractions from *P. integrifolia,* fractionated by SPE as described above, were purified using a 30-min LC method that employed 0.15% aqueous formic acid as mobile phase A and acetonitrile as mobile phase B. The purification gradient started with 90% A/10% B, increased by linear gradient to 40% B at 1 min, followed by a linear gradient to 70% B

over 1-25 min, then a steep linear gradient to 100% B at 26 followed by a hold at 100% B until 28 min. Solvent was returned to 10% B at 29 min and held for an additional 1 min.

Individual *P. integrifolia* malonate acylsugars from the acidic SAX-retained SPE fraction were purified using an 80-min LC method with 0.15% aqueous formic acid as mobile phase A and acetonitrile as mobile phase B. The purification gradient started with 5% B, held 1 min, then quickly increased to 41% B at 2 min, 41-42% B at 2-5 min, linear gradient to 43% B at 8 min, 45-47% B at 9-12 min, 47-49% B at 12-24 min, 49-50% B at 24-25 min, 50-52% B at 25-40 min, 52-53% B at 40-41 min, 53-55% B during 41-56 min, 55-57% B at 56-57 min, 57-59% B at 57-67 min, 59-60% B at 67-68 min, 60-61% B at 68-75 min, held 100% B at 76-78 min, returned to 5% B at 79 min and held 1 min.

Acylsugar purification method from *P. exserta.*

For *P. exserta,* SPE washing solution I was 60% methanol/40% aqueous 0.5 mM ammonium formate (adjusted to pH 6.0 with ammonium hydroxide) and washing solution II was 70% methanol/30% aqueous 0.5 mM ammonium formate (adjusted to pH 6.0 with ammonium hydroxide).

Individual *P. exserta* neutral acylsugars, fractionated by SPE as described above were purified using a 40 min LC method with 0.15% formic acid, as mobile phase A and acetonitrile as mobile phase B. The purification gradient started with 10% B, held for 1 min, then quickly increased to 45% B at 2 min, linear gradient to 50% B during 2-10 min, 54% B at 11 min, 54-56% B during 11-20 min, 57% B at 21 min, 57-59% B during 21-30 min, 61% B at 31 min, 61-70% B during 31-36 min, held 100% B at 37-38 min, returning to 10% B at 39 min and held 1 min.

P. exserta malonate acylsugars were purified using a 50 min LC method with 0.15% formic acid, as mobile phase A and acetonitrile as mobile phase B. The purification gradient start with 10% B, held 1 min, then quickly increased to 50% B at 2 min, 50-52% B at 2-5 min, linear gradient to 54% B at 10 min, 54-56% B at 10-20 min, 57-59% B at 21-31 min, 61% B at 32 min, 61-63% B at 32-38 min, 65% B at 39 min, 65-68% B during 39-44 min, held 100% B at 45-48 min, return to 10% B at 49 min and held 1 min.

Acylsugar purification method from *P. axillaris*

For *P. axillaris,* SPE washing solution I was 50% methanol/50% aqueous 0.5 mM ammonium formate adjusted to pH 6.0 with ammonium hydroxide, and washing solution II was 60% methanol/40% aqueous 0.5 mM ammonium formate, adjusted to pH 6.0 with ammonium hydroxide.

P. axillaris neutral acylsugars were purified using a 50-min LC method with 0.15% formic acid, as mobile phase A and acetonitrile as mobile phase B. The purification gradient started with 10% B, held for 3 min, then quickly increased to 40% B at 5 min, linear gradient to 70% B during 5-45 min, held 100% B at 46-48 min, returned to 10% B at 49 min and held 1 min.

P. axillaris malonate acylsugars were purified using a 60-min LC method with 0.15% formic acid, as mobile phase A and acetonitrile as mobile phase B. The purification gradient started with 10% B, held 1 min, then quickly increased to 44% B at 5 min, 44-46% B at 5-13 min, 46-48% B during 13-18 min, linear gradient to 49% B at 25 min, 49-51% B at 25-30 min, held 51% B for 4 min, jump to 57% B at 35 min, held 57% B for 5 min, increased to 59% B at 41 min, held for 4 min, 59-62% B during 45-46 min, held 62% B to 55 min, increased to 100% B at 56 min, held 100% B at 56-58 min, return to 10% B at 59 min and held 1 min.

2.3.4 Structure elucidation of purified acylsugars

NMR spectra were obtained using either Agilent DirectDrive2 500 MHz (Agilent) or Avance 900 MHz (Bruker) NMR spectrometers at the Max T. Rogers NMR Facility at Michigan State University. $CDCl_3$ was used as NMR solvent and chemical shift reference (δ = 7.24 ppm for ^{1}H, δ = 77.0 ppm for ^{13}C) for individual acylsugars. Initial NMR data were obtained on an Agilent DirectDrive2 500 MHz NMR spectrometer. Purified acylsugar fractions were evaporated to dryness under vacuum at room temperature using a SpeedVac, redissolved using 600 μL $CDCl_3$, and transferred to NMR tubes. When metabolite quantities were too low to generate useful NMR spectra these low-abundance samples were concentrated by evaporating solvent under a flow of N_2 gas at room temperature, redissolved in about 250 μL $CDCl_3$, and transferred into Shigemi NMR tubes (matched for $CDCl_3$). NMR spectra were generated on a Bruker Avance 900 MHz NMR (detailed NMR procedures are listed below). Vnmrj 3.2 software was used for processing of spectra from the 500 MHz NMR instrument, and Topspin 3.2 software was applied for processing of spectra obtained from the 900 MHz NMR spectrometer

2.3. Results and discussion

Twenty-nine distinct acylsucroses were purified from *P. integrifolia*, *P. axillaris* and *P. exserta*, including malonate ester acylsucroses that have not, to our knowledge, been isolated from other *Solanaceae* species. All are described in Table 2.1 with UHPLC retention times, accurate molecular masses and structures proposed based on NMR spectroscopy. NMR 1D and 2D spectra, chemical shift assignments, PubChem CID numbers, and InChI keys are presented in Appendix I. Since these metabolites were identified based on NMR and mass spectrometric characterization without available authentic standards or synthetic confirmation, their structures

should be considered as putative, meeting level 2 criteria of the Metabolomics Standards Initiative guidelines [30].

We propose a modified nomenclature for *Petunia* acylsugars derived from that proposed for acylsugars from tomato and its wild relatives [12,20]. "S" refers to the sucrose core, and designation of the letter "m" in parentheses identifies compounds containing a malonate ester. The first number indicates the number of acyl groups attached to the sugar core, and the number following the colon represents the total number of carbons in acyl chains. Numbers in parentheses indicate the number of carbons in the individual acyl chains. Isomeric acylsucroses with the same molecular mass are ordered based on increasing chromatographic retention times, which are labeled as a number in brackets (*e.g.* S(m)5:25[2]). Branching of acyl groups is specified by using the abbreviations 'i' for iso and 'ai' for anteiso branching.

2.3.1 Acylsugar profiling using UHPLC/TOF-MS

UHPLC/TOF-MS analyses demonstrated numerous chromatographic peaks that were annotated based on retention times and accurate pseudomolecular and fragment ion masses. Most observed metabolites were assigned as acylsucroses. Thirty-seven distinct acylsucrose pseudomolecular masses were resolved from *P. integrifolia, P. exserta* and *P. axillaris* (Figure 2.1), consisting of 2 triesters, 16 tetraesters and 19 pentaesters of sucrose. More than 100 distinct acylsugar forms were resolved using an 80-min LC gradient. All neutral aliphatic acyl groups were substituted on the pyranose ring as judged from fragment ions generated using non-selective CID in positive-ion mode. To our knowledge, none of these *Petunia* acylsugars have been identified in tomato or its wild relatives.

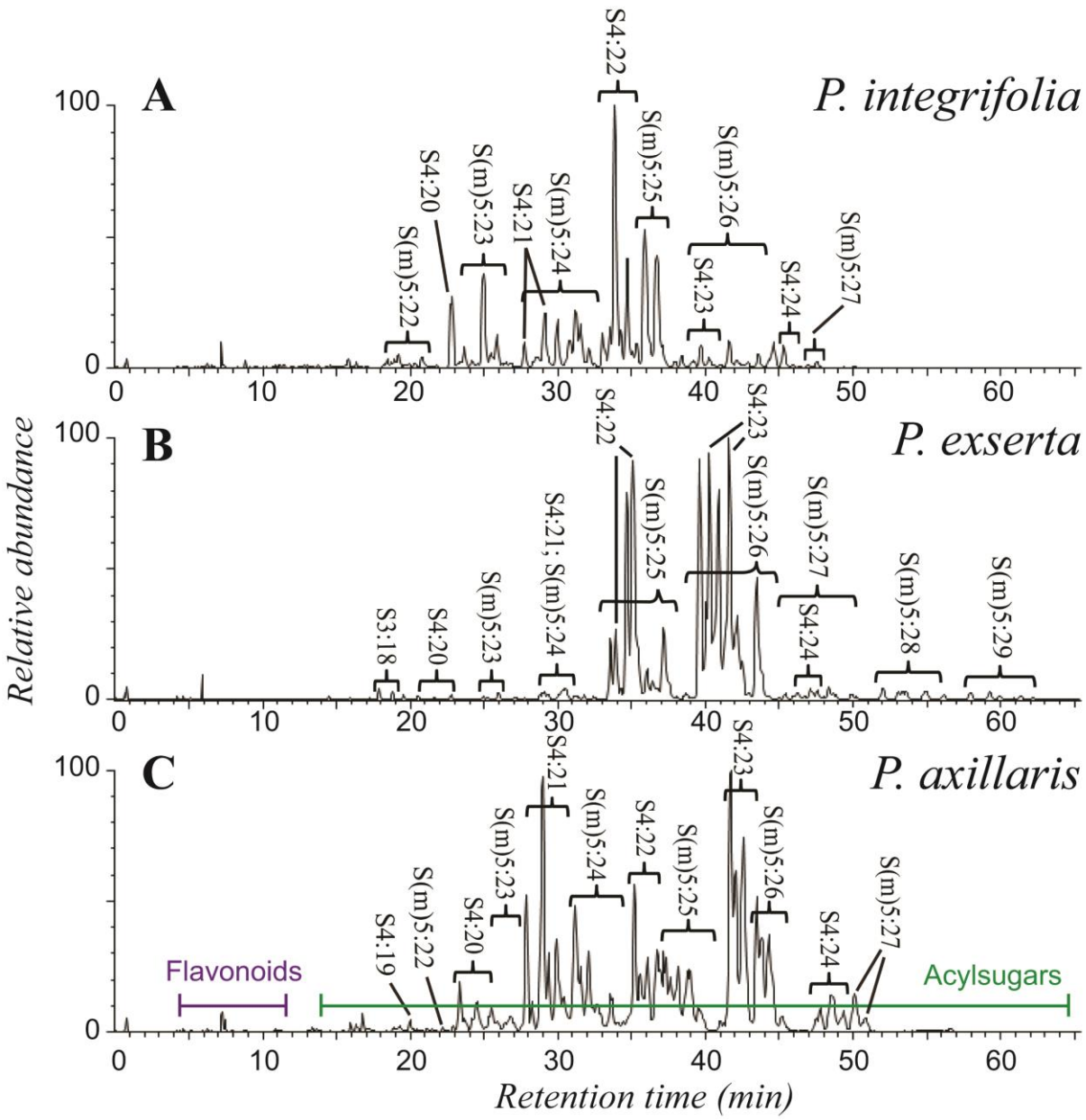


Figure 2.1. LC/MS Base Peak Intensity (BPI) chromatograms of leaflet extracts from three Petunia species*. Peaks for specialized metabolites are annotated. Data were generated using ESI in negative-ion mode with 80-min LC gradient. A. P. integrifolia; B. P. exserta; C. P. axillaris.*

Extracted ion chromatograms for carboxylate fragment ions (*m/z* 87, 101, 115…) generated at the highest collision potential provide information about the number of carbon atoms in the aliphatic acyl groups. All three genotypes yielded evidence for acyl groups ranging from C4 (*m/z*

87) to C9 (*m/z* 157), with relative amounts varying across genotypes. In contrast to a recent report of acyl groups in acylsugars of *Petunia* hybrids (Kroumova et al. 2016), our LC/MS data indicated low levels of C9 esters in acylsucroses from *P. axillaris* and *P. exserta* but not *P. integrifolia*. The isomeric diversity of acylsucroses was extensive based on extracted ion chromatograms for [M+formate]$^-$ ions (Figure 2.2), which suggest about 18 isomers of S4:21 across all three Petunia species. This finding is in accord with variation in positions of esterification by iso- and anteiso- groups.

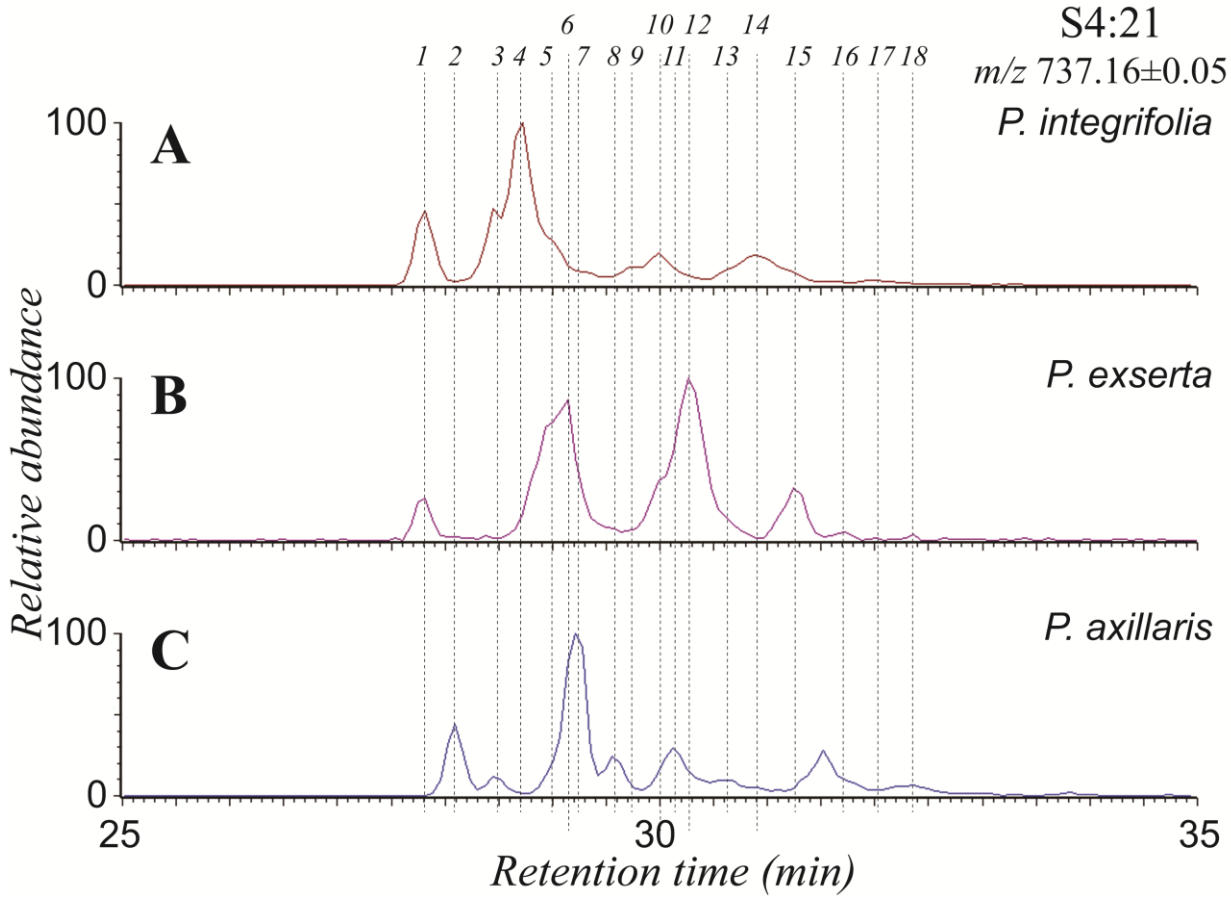


Figure 2.2. LC/MS extracted ion chromatograms for m/z 737.36±0.05, corresponding to neutral acylsucrose S4:21 from three wild petunia sepcies*, obtained in negative-ion mode for leaflet extracts from (A) P. integrifolia, (B) P. exserta, and (C) P. axillaris showing evidence for a total of 18 isomers, some incompletely resolved, across the three Petunia genotypes.*

Acylsucroses that possess a malonyl group on the furanose ring were abundant in all three *Petunia* species and could be distinguished by as ions with different nominal masses than the aliphatic ester series. For example, neutral acylsugars were detected in negative ion mode as $[M+formate]^-$ ions as a series of homologs differing in nominal mass by 14 Da (*m/z* 709, 723, 737, 751, 765; *e.g.* S4:23 appeared at *m/z* 765), whereas malonate esters formed a series of $[M-H]^-$ ions 2 Da lighter (*m/z* 763, 777, 791, 805…; *e.g.* S(m)5:23 at *m/z* 763). In addition, multiplexed CID spectra distinguished malonate acylsugars, which undergo facile decarboxylation to $[M-H-CO_2]^-$ at the low collision potential of 25 V (Figure 2.3), while neutral acylsugars did not yield significant fragment ions under these relatively gentle conditions.

CID spectra generated at multiple collision voltages provided important information to distinguish isomeric acylsugars. In negative-ion mass spectra, compositions of acyl chains and esterification positions in acylsugar were suggested by fragment ions formed at intermediate collision potentials (25, 40 and 55 V) corresponding to neutral losses of acyl groups as ketenes (*e.g.* 84 Da for C5, 112 Da for C7). Ions corresponding to deprotonated fatty acids (*e.g. m/z* 101 Da for C5 carboxylates and 129 Da for C7 carboxylates) were detected at the highest collision potential (80 V). In contrast, positive-ion mass spectra yielded fragment ions corresponding to cleavage of the glycosidic bond as the collision voltage increased. The masses of abundant key fragment ions facilitated assignment of acyl groups to either the pyranose or furanose rings. For neutral acylsucroses from tomato and its wild relatives, the most abundant fragment ion was usually formed with retention of charge on the furanose ring and its mass reflected the sum of acyl substitutions. However, for *Petunia* acylsucroses containing a malonate ester, the presence of malonate ester shifted the most abundant positively-charged fragment to the pyranose ring ion (*e.g. m/z* 527 for S(m)5:25; Figure 2.3A) and the malonylfructose fragment at *m/z* 249 was less

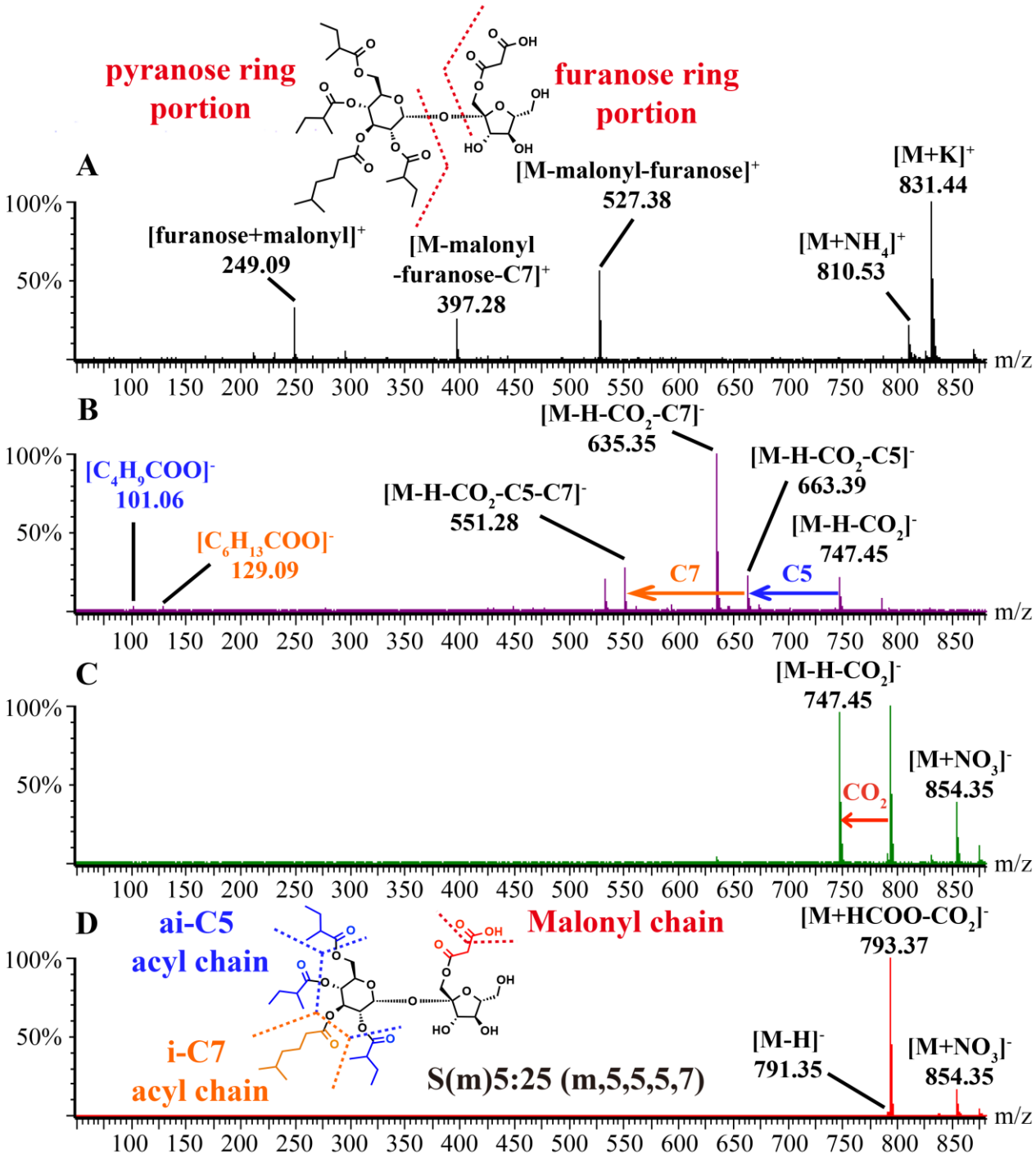


Figure 2.3. Multiplexed CID mass spectra of S(m)5:25 (RT ~ 36.16 min) from P. exserta. *A. positive-ion ESI spectrum (collision potential = 40 V) B. negative-ion ESI spectra (collision potentials = 80 V, 25 V, 10 V; top to bottom).*

abundant than typical when fructose groups are esterified by neutral acids. Figure 2.3 illustrates

how negative and positive mode mass spectra were applied to establish carbon number of acyl chains and their distribution on each ring of sucrose in malonate acylsugar S(m)5:25. Observation of $[M-H]^-$ (*m/z* 791) at lowest CID potential and $[M-H-CO_2]^-$ (*m/z* 747) formed at 25 V suggested malonyl ester decarboxylation. When CID potential was increased to 80 V, ions corresponding to additional neutral aliphatic ketene losses ($[M-H-CO_2-C5]^-$, *m/z* 663) and $[M-H-CO_2-C5-C7]^-$, *m/z* 551) were observed, as were carboxylate ions for C5 (*m/z* 101) and C7 (*m/z* 129) aliphatic acids. In positive-ion mass spectra, $[malonylfuranosyl]^+$ (*m/z* 249) was observed at a collision potential of 40 V, consistent with substitution of only the malonyl ester on the furanose ring. More specific assignments of acyl group attachments at specific positions on sucrose were made based on one-dimensional and two-dimensional NMR (1D/2D NMR) spectra of purified compounds as described below.

2.3.2 Isolation of malonate acylsugars using LC-SAX SPE

Fractionation on a strong anion exchange (SAX) solid phase extraction column (SPE)

Table 2.1. NMR elucidated structures of acylsugars purified from leaf extracts of P. integrifolia, P. exserta and P. axillaris.

						Acyl group at position						Acylsugar abundances		
Acylsugar[a]	**RT**[f] **(min)**	**Ion type**	**Measured *m/z***	**Calculated *m/z***	**Δm (ppm)**	**R2**	**R3**	**R4**	**R6**	**R1´**	**R6´**	***P. integrifolia***	***P. exserta***	***P. axillaris***
Neutral acylsucroses														
S4:19[d]	19.92	[M+formate]$^-$	709.3226	709.3288	-8.7	aiC5	iC5	aiC5	iC4	H	H	ND	ND	+
S4:20 [1][b]	23.28	[M+formate]$^-$	723.3418	723.3445	-3.7	aiC5	iC5	aiC5	aiC5	H	H	+	+	+
S4:20 [2][d]	24.47	[M+formate]$^-$	723.3459	723.3445	1.9	aiC5	aiC6	aiC5	iC4	H	H	+	+	+
S4:21 [1][b]	27.82	[M+formate]$^-$	737.3635	737.3601	4.6	iC6	aiC5	aiC5	aiC5	H	H	+	+	+
S4:21 [2][d]	28.94	[M+formate]$^-$	737.3620	737.3601	2.6	aiC5	aiC6	aiC5	aiC5	H	H	+	+	+
S4:21 [3][b,c,d]	29.66	[M+formate]$^-$	737.3646	737.3601	6.1	aiC5	iC6	aiC5	aiC5	H	H	++	++	++
S4:21 [4][d]	34.14	[M+formate]$^-$	737.3604	737.3601	0.4	aiC5	iC6	aiC5	iC5	H	H	++	+	++
S4:22 [1][b,c,d]	35.09	[M+formate]$^-$	751.3795	751.3758	4.9	aiC5	iC7	aiC5	aiC5	H	H	+++	+++	++
S4:22 [2][b]	35.18	[M+formate]$^-$	751.3785	751.3758	3.6	aiC5	iC7	iC5	aiC5	H	H	++	++	ND
S4:22 [3][c,d]	36.70	[M+formate]$^-$	751.3802	751.3758	5.9	aiC5	iC8	aiC5	iC4	H	H	++	++	+
S4:23[b,c,d]	41.31	[M+formate]$^-$	765.4017	765.3914	13.5	aiC5	iC8	aiC5	aiC5	H	H	+	++	++
S4:24 [1][d]	47.74	[M+formate]$^-$	779.4101	779.4071	3.8	iC6	iC8	aiC5	aiC5	H	H	ND	+	+
S4:24 [2][d]	48.51	[M+formate]$^-$	779.4064	779.4071	-0.9	iC8	aiC6	aiC5	aiC5	H	H	++	++	+
Malonate esters														
S(m)5:23 [1][b,c,d]	25.79	[M-H]$^-$	763.3383	763.3394	-1.4	aiC5	iC5	aiC5	aiC5	malonyl	H	++	+	++
S(m)5:23 [2][d]	26.77	[M-H]$^-$	763.3421	763.3394	3.5	aiC5	iC6	aiC5	iC4	malonyl	H	ND	ND	++
S(m)5:24 [1][c]	29.80	[M-H]$^-$	777.3506	777.3550	-5.7	aiC5	iC6	aiC5	aiC5	H	malonyl	ND	+	ND
S(m)5:24 [2][b]	29.99	[M-H]$^-$	777.3669	777.3550	15.3	iC4	iC7	aiC5	aiC5	malonyl	H	++	+	+
S(m)5:24 [3][b]	30.13	[M-H]$^-$	777.3668	777.3550	15.2	aiC5	iC7	iC4	aiC5	malonyl	H	++	+	+
S(m)5:24 [4][b]	30.50	[M-H]$^-$	777.3651	777.3550	13.0	iC5	iC7	iC4	aiC5	malonyl	H	+	+	++
S(m)5:24 [5][d]	30.60	[M-H]$^-$	777.3565	777.3550	1.9	aiC5	iC6	aiC5	aiC5	malonyl	H	+	+	++
S(m)5:24 [6][d]	31.11	[M-H]$^-$	777.3608	777.3550	7.5	aiC5	aiC6	aiC5	aiC5	malonyl	H	+	+	++
S(m)5:25 [1][c]	34.32	[M-H]$^-$	791.3682	791.3707	-3.2	aiC5	iC7	aiC5	aiC5	H	malonyl	ND	+++	ND
S(m)5:25 [2][b]	36.37	[M-H]$^-$	791.3758	791.3707	6.4	aiC5	iC5	iC7	aiC5	malonyl	H	+++	+	ND
S(m)5:25 [3][b,d]	37.12	[M-H]$^-$	791.3731	791.3707	3.0	aiC5	iC7	aiC5	aiC5	malonyl	H	+++	++	++
S(m)5:25 [4][d]	38.79	[M-H]$^-$	791.3732	791.3707	3.2	aiC5	iC8	aiC5	iC4	malonyl	H	+	+	++
S(m)5:26 [1][c]	39.28	[M-H]$^-$	805.3885	805.3863	2.7	aiC5	iC8	aiC5	aiC5	H	malonyl	ND	++	ND
S(m)5:26 [2][c]	40.54	[M-H]$^-$	805.3915	805.3863	6.5	aiC5	iC8	aiC5	aiC5	malonyl	H	ND	++	ND
S(m)5:27[c]	46.67	[M-H]$^-$	819.3998	819.4020	-2.7	aiC5	iC7	aiC5	iC7	malonyl	H	+	+	ND

a. The nomenclature of acylsugars and abbreviations used for branched-type of acyl chains are discussed in the text. Numbers in brackets indicate the relative elute order of isomeric acylsugars with same molecular mass.
iC4 = $(CH_3)_2CHCO$, iC5 = $(CH_3)_2CHCH_2CO$, aiC5 = $CH_3CH_2(CH_3)CHCO$, iC6 = $(CH_3)_2CHCH_2CH_2CO$, aiC6 = $CH_3CH_2(CH_3)CHCH_2CO$, iC7 = $(CH_3)_2CHCH_2CH_2CH_2CO$, iC8 = $(CH_3)_2CHCH_2CH_2CH_2CH_2CO$

b. Acylsugar metabolites purified from *P. integrifolia*

c. Acylsugar metabolites purified from *P. exserta*

d. Acylsugar metabolites purified from *P. axillaris*

e. Acylsugar abundance of acylsugars in each species were calculated as relative abundance (X), similar as that of *Solanum* acylsugars (Ghosh et al. 2014). X=(peak area*100/mg dry weight of leaf), symbol +, ++, +++ and ND indicate X>1000, 1000>X>100, 100>X>1 and X<1, respectively. Acylsugars in different species were matched based on their molecular mass, fragmentation behavior and retention time. ND=not detected.
f. RT: Retention Time

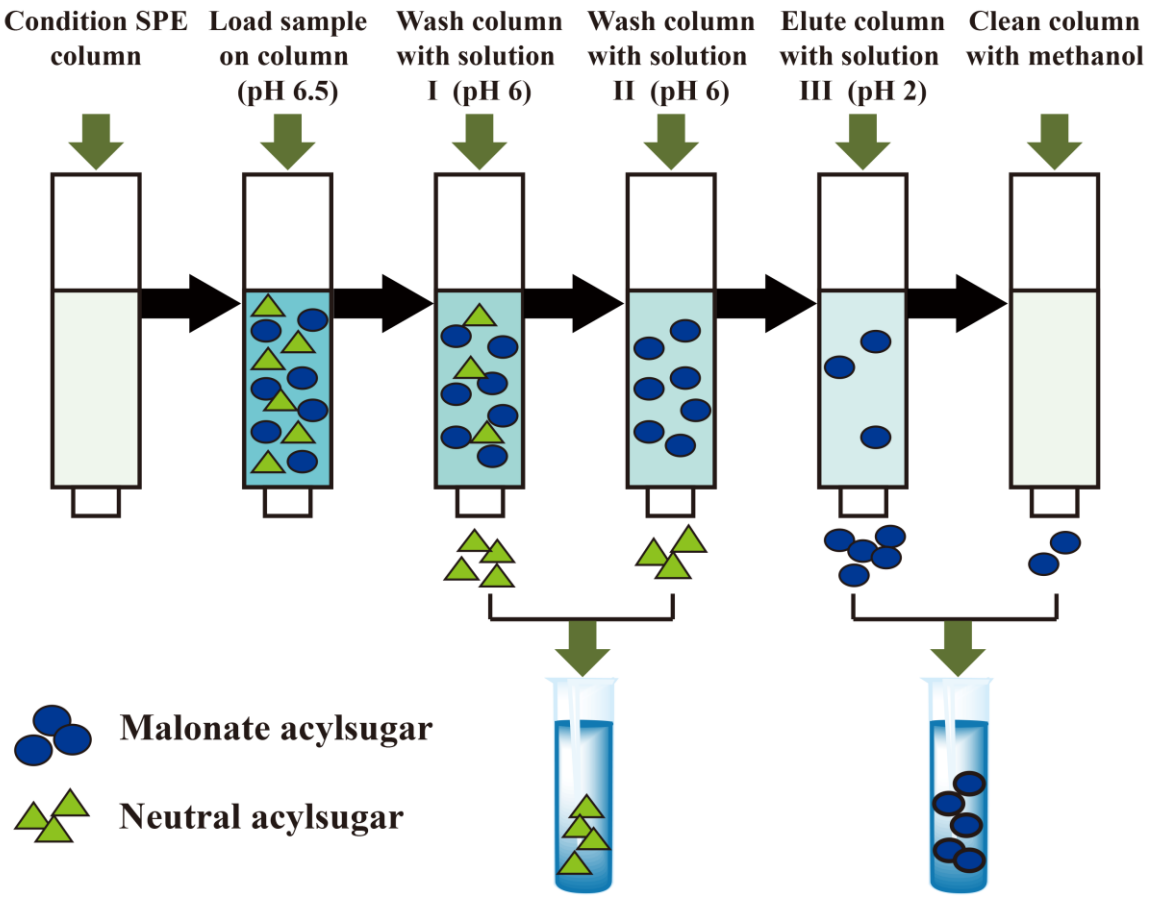


Figure 2.4. Workflow for malonate acylsugar fractionation from plant extracts using LC-SAX SPE column.

separated acidic malonate ester acylsugars from neutral metabolites and facilitated subsequent purification (Figure 2.4). Several malonate ester acylsugars that have similar UHPLC retention times (Table 2.1) as neutral acylsugars were challenging to isolate using only reversed phase HPLC owing to similar retention times. A solid phase anion exchange column retained malonate ester acylsugars, which are negatively-charged at pH 6, and neutral acylsugars were eluted using organic solvent followed by elution of acidic acylsugars using acidified methanol/water. Details were described in Materials and methods, section 2.2.3.

2.3.3 Acylsugar structure elucidation using NMR

To fully understand *Petunia* acylsugar chemical diversity with a broader goal to accelerate discoveries of biosynthetic genes involved in acylsugar metabolites accumulation, 14 neutral acylsucroses and 15 malonate acylsucroses were purified from three *Petunia* species. Their

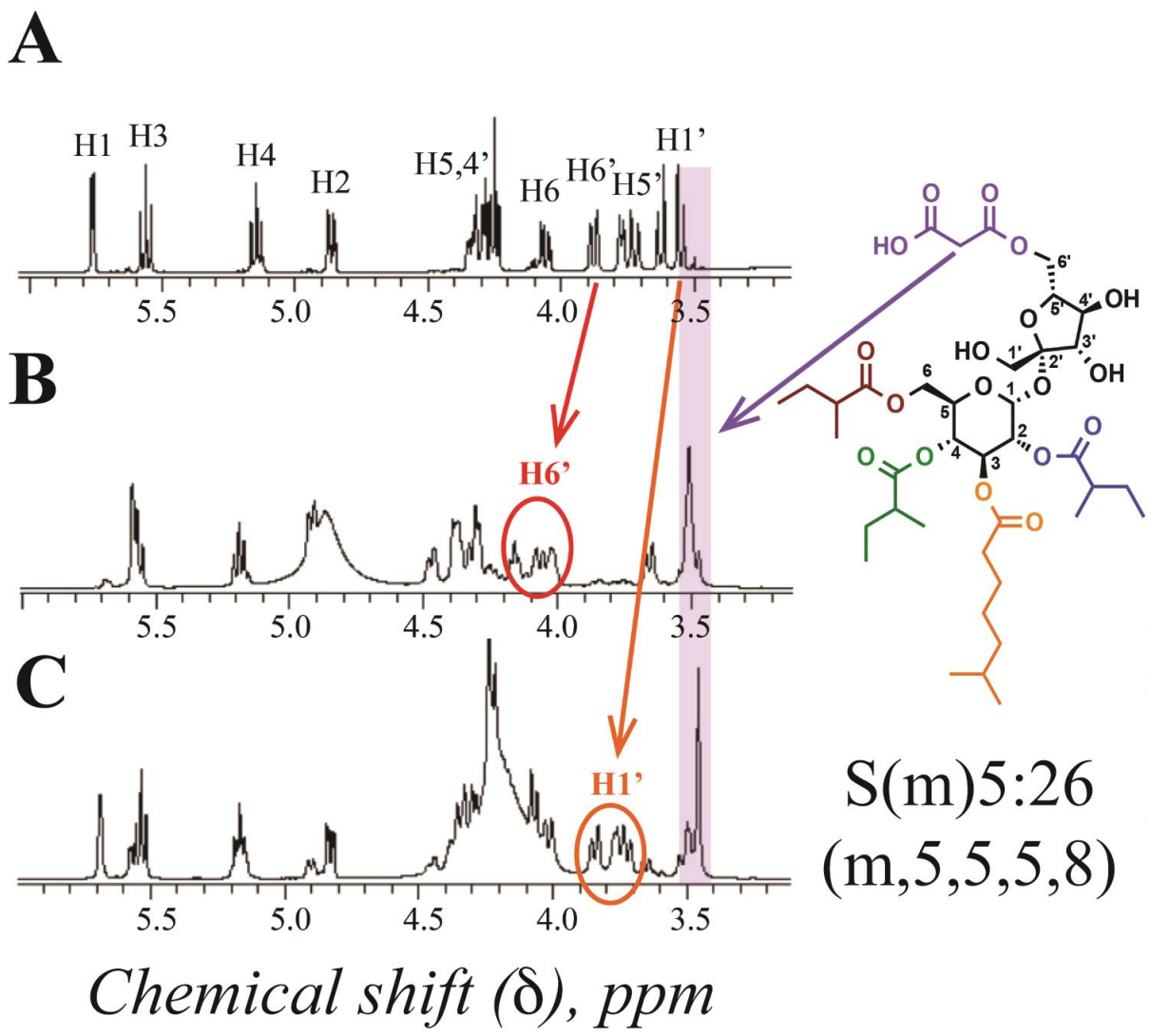


Figure 2.5. NMR spectra for P. exserta acylsugars. *^{1}H NMR spectra indicate that there are 2 substitution positions for malonate groups in P. exserta acylsugars in (A) S4:23 (5,5,5,8), (B) S(m)5:26(m,5,5,5,8)[1], and (C) S(m)5:26(m,5,5,5,8)[2].*

proposed structures were suggested based on 1D (^{1}H and ^{13}C) and 2D (COSY, HSQC and HMBC) NMR data. There are three distinct chemical shift regions for proton resonances in ^{1}H spectra of all purified acylsugars: 1) 0.7~1.8 ppm is the region for aliphatic H from acyl groups; 2) 2.0~2.5 ppm is the region for α-H of carbonyl; 3) 4.0~5.8 ppm is the region for protons on the pyranose ring. H attached to substituted C in sucrose ring fell in the 4.3~5.8 ppm range, and 3.6~4.3 ppm is the region for protons on furanose ring carbons that are not acylated. Acylation causes a downfield shift in resonances for protons attached at the esterified carbon positions and this information was used to identify substituted positions in simple cases (Figure 2.5A, B, and C). ^{1}H and COSY NMR spectra were applied to assign protons on sucrose carbons based on ^{1}H-^{1}H coupling, starting with the most downfield peaks (δ 5.6-5.8 ppm) that correspond to the 1-position proton on the pyranose ring. ^{13}C and HSQC spectra were used to not only assign carbon atoms on sucrose core and for acyl chains but also to distinguish ring protons when ^{1}H peaks overlapped in the region of 3.8-4.3 ppm since chemical shifts of their direct attached C differ between esterified and nonesterified positions (Table 2.2). Substitution positions for each acyl chain on the sucrose core were recognized using HMBC spectra, which showed correlation between H and C nuclei that are 2-4 bonds away from each other. Carbonyl carbons (δ 170~180 ppm) generated cross peaks with both pyranose ring protons and α-position protons on acyl chains to demonstrate specific esterification positions for individual acyl groups.

Table 2.2. Summary of acylsugar ^{13}C chemical shifts of sucrose core carbon atoms. Positions substituted with acyl groups are marked as bold.

Acylsugar	R1	R2	R3	R4	R5	R6	R1'	R2'	R3'	R4'	R5'	R6'
Neutral acylsucroses (13)												
S4:19	89.07	**70.72**	**68.99**	**67.64**	68.99	**61.73**	65.05	105.01	78.60	73.80	81.80	60.25
S4:20 [1]	89.10	**70.88**	**69.20**	**67.80**	69.06	**61.49**	65.00	104.98	78.59	73.83	81.95	60.51
S4:20 [2]	89.00	**70.67**	**68.94**	**67.48**	68.81	**61.37**	64.69	102.42	78.37	73.59	81.56	60.17
S4:21 [1]	88.97	**70.83**	**68.78**	**64.62**	69.03	**61.57**	64.92	104.53	78.55	73.92	81.44	60.16
S4:21 [2]	89.21	**70.71**	**69.45**	**67.77**	68.89	**61.46**	64.68	104.85	78.56	73.65	81.92	60.62
S4:21 [3]	89.34	**70.74**	**69.60**	**67.68**	68.81	**61.44**	64.39	104.79	78.00	73.80	81.91	60.87
S4:21 [4]	89.11	**70.77**	**69.31**	**67.72**	69.05	**61.74**	64.80	105.09	78.61	73.56	81.81	60.41
S4:22 [1]	89.02	**70.69**	**69.89**	**67.69**	69.01	**61.58**	64.49	105.06	78.72	73.80	81.71	60.14
S4:22 [2]	88.90	**70.73**	**68.87**	**67.75**	69.50	**61.60**	65.23	102.57	78.74	73.81	81.73	60.29
S4:22 [3]	89.11	**70.68**	**68.98**	**67.61**	69.07	**61.41**	64.97	105.04	78.82	73.81	81.82	60.23
S4:23	89.12	**70.68**	**69.25**	**67.60**	68.81	**61.34**	64.75	105.01	78.47	73.64	81.77	60.46
S4:24 [1]	89.00	**70.99**	**69.45**	**68.03**	67.65	**61.60**	64.95	102.61	78.45	73.95	81.54	60.19
S4:24 [2]	89.74	**71.68**	**69.96**	**68.67**	68.24	**62.37**	65.65	105.86	79.42	74.43	82.42	61.08
Malonate acylsugar (15)												
S(m)5:23 [1]	88.93	**70.87**	**69.12**	**67.64**	68.75	**61.65**	**64.97**	104.98	78.51	73.72	81.65	60.18
S(m)5:23 [2]	89.22	**70.73**	**69.34**	**67.59**	68.76	**61.32**	**64.34**	103.19	77.13	73.06	81.08	60.62
S(m)5:24 [1]	89.14	**70.67**	**69.71**	**67.68**	68.32	**61.61**	63.19	104.8	77.03	75.11	78.34	**64.57**
S(m)5:24 [2]	89.09	**70.64**	**69.24**	**67.54**	68.86	**60.39**	**63.18**	104.08	78.47	73.81	81.73	60.39
S(m)5:24 [3]	89.90	**70.73**	**69.80**	**67.56**	68.86	**60.39**	**65.01**	104.80	78.56	73.90	81.73	60.39
S(m)5:24 [4]	89.40	**70.82**	**69.71**	**67.93**	69.05	**61.23**	**63.28**	103.04	78.15	73.81	81.54	60.76
S(m)5:24 [5]	89.54	**70.44**	**69.58**	**67.59**	68.73	**61.32**	**63.89**	103.30	76.85	73.00	81.28	60.89
S(m)5:24 [6]	89.27	**70.77**	**69.47**	**67.64**	68.79	**61.35**	**64.21**	103.47	77.02	73.02	81.14	60.66
S(m)5:25 [1]	89.52	**70.42**	**69.93**	**67.81**	68.46	**61.64**	63.11	104.65	76.85	75.27	78.43	**65.02**
S(m)5:25 [2]	88.98	**70.67**	**69.09**	**67.68**	68.91	**60.44**	**61.34**	104.90	78.51	73.84	71.76	60.44
S(m)5:25 [3]	89.56	**70.90**	**69.84**	**67.95**	69.13	**61.58**	**64.41**	103.19	77.28	73.38	81.14	61.11
S(m)5:25 [4]	89.16	**70.62**	**69.36**	**67.76**	68.90	**61.35**	**64.44**	103.65	77.26	72.91	81.04	60.43
S(m)5:26 [1]	89.44	**70.32**	**69.94**	**67.93**	68.43	**61.64**	63.03	104.63	76.73	75.48	78.49	**65.04**
S(m)5:26 [2]	89.35	**70.64**	**69.72**	**67.74**	68.67	**61.35**	**64.02**	103.11	76.92	73.20	81.10	60.65
S(m)5:27	88.98	**70.91**	**69.58**	**67.75**	69.06	**61.32**	**64.28**	104.80	76.97	73.67	80.89	60.45

^{1}H and HMBC spectra provided clear evidence for malonyl ester groups. Compared with neutral acylsugar ^{1}H spectra, a unique singlet of the malonyl α-proton was present around δ3.5 ppm. Furthermore, two malonate carbonyl carbon resonances at 166-171 ppm correlated with the α-proton resonance at 3.5 ppm, producing two corresponding cross-peaks in HMBC spectra. For example, S(m)5:26(m,5,5,5,8) showed malonate α-proton resonance at 3.5 ppm correlated with carbonyl carbon resonances at 167.1 and 170.2 ppm, and the former correlated with 6´-position

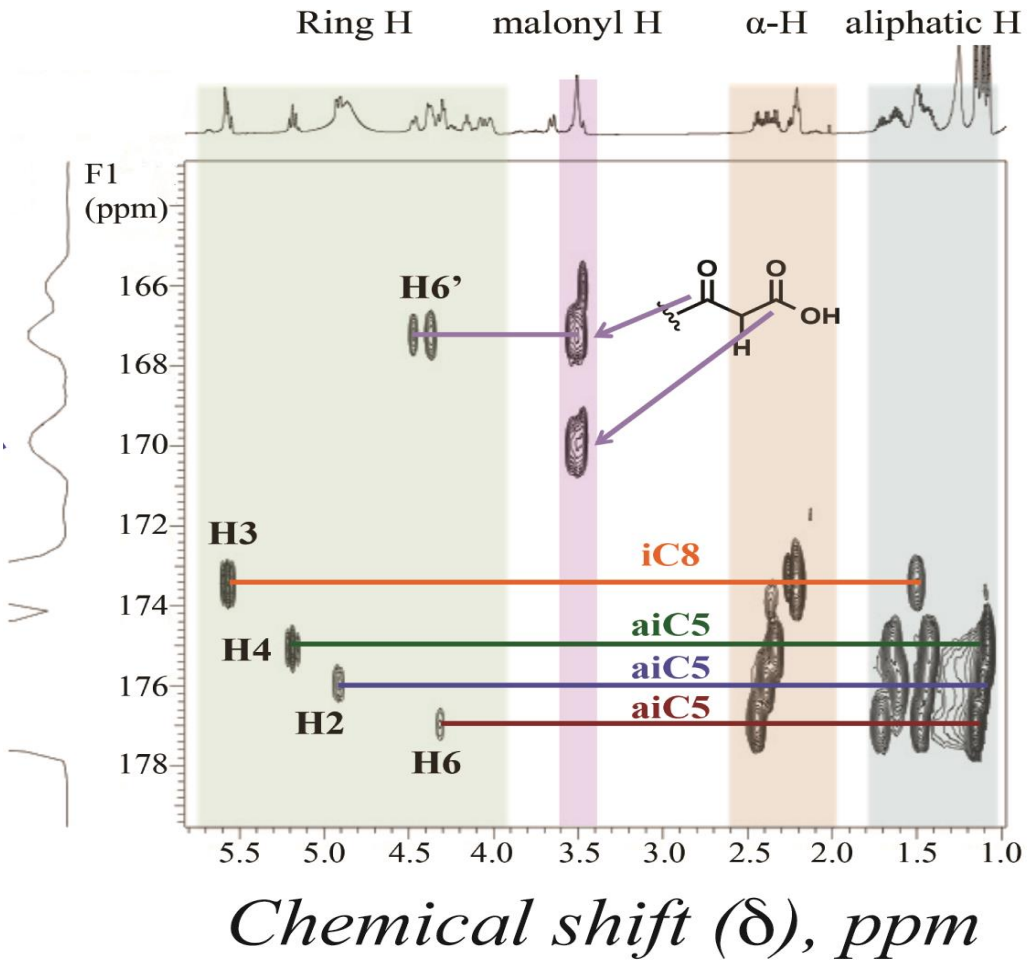


Figure 2.6. HMBC spectrum of S(m)5:26(m,5,5,5,8)[1] from P. exserta provided evidence for presence of the malonate group and its attachment on the 6' position.

protons (δ 4.33, 4.48 ppm) that demonstrated attachment of the malonyl chain at the 6´ position (Figure 2.6). Because of the facile decarboxylation of the malonyl group, the presence of a malonyl group on the furanose ring was confirmed by mass spectra. A new cross peak in HMBC spectra indicative of an acetyl group, instead of a malonyl group, substituted on furanose ring appeared for old extracts (stored > 2 weeks at -20 °C after purification; Figure 2.7). This finding may explain why esterification by acetyl groups has been reported in previous petunia acylsugar analyses. Table 2.1 lists all structures of purified acylsugars from *P. integrifolia, P. exserta* and *P. axillaris* elucidated from 1D and 2D NMR spectra, with relative abundances evaluated using chromatographic peak areas normalized to dry weight of leaflets.

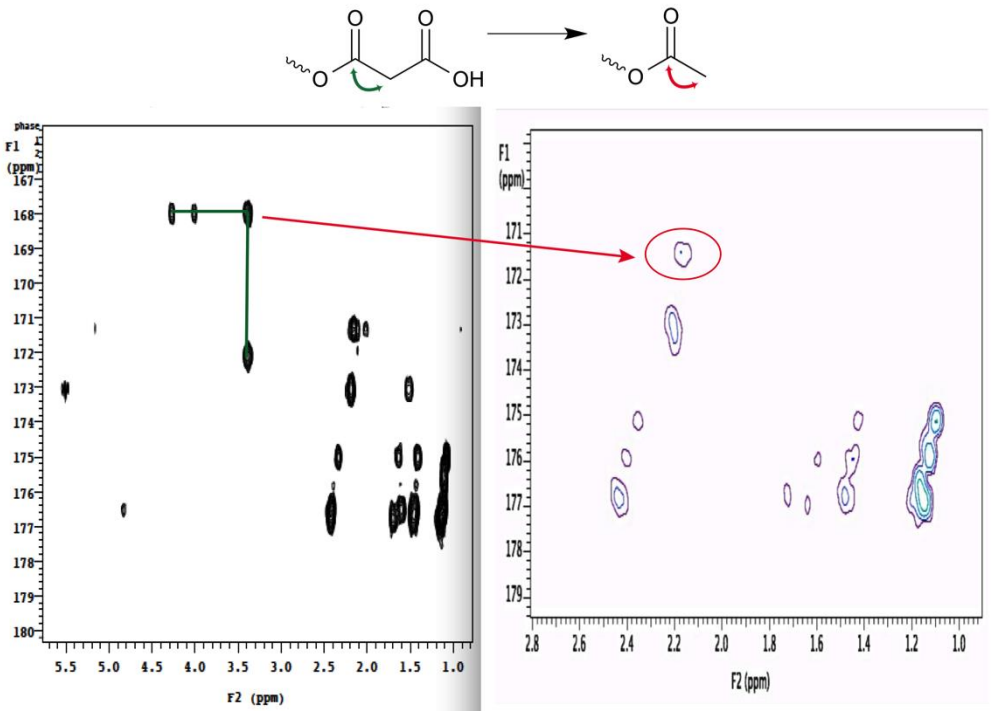


Figure 2.7. HMBC evidence for decarboxylation of an acylsugar malonate ester. *A new peak corresponding to C2 residue after loss CO_2*

2.3.4 Comparison of acylsugar profiles among 3 *Petunia* species and wild tomato.

Although the three species *P. integrifolia, P. exserta* and *P. axillaris* all belong to the genus *Petunia*, acylsugar profiles revealed substantial differences, not only in qualitative aspects (*e.g.* presence of acylsugar isomers differing in substitution pattern or acyl chain branching), but quantitative levels of both neutral and malonate acylsugars among the three species (see Table 2.1 and Figure 2.1). UHPLC/MS chromatograms of three *Petunia* plants exhibited significant differentiations in acylsugar profiles based on chromatographic retention times, fragment ion patterns and relative abundances of specific acylsugars based on integrated extracted ion chromatograms (Figure 2.1). Malonate ester acylsugars were more abundant relative to neutral acylsugars in *P. exserta* and *P. axillaris*, while in *P. integrifolia,* the levels of the two types were similar. Such diversity demonstrated within *Petunia* species may reflect differences in selectivity of acyltransferases as well as different levels of acyl-CoA pools.

Additionally, other specialized metabolites were detected in extracts of the three *Petunia* species using UHPLC/MS. Several lipids were found in three *Petunia* species; flavonoids quercetin monoglucoside, diglucoside and 3-methylquercetin were detected in *P. exserta* and a dimethylquercetin in *P. axillaris*. Neither terpenes nor alkaloids were detected based on MarkerLynx screen.

2.3.5 Diversity and conservation in position selective acylation

NMR structures of *Petunia* acylsugars summarized in Table 2.1 define sites of acyl group attachment and provide evidence of chemical diversity among three *Petunia* species. All hydroxyl groups on the pyranose ring of *Petunia* acylsucroses are substituted by acyl groups, while malonyl chains were detected only on the furanose ring. Acyl group diversity was greatest

on the 3-position of the pyranose ring, with all acyl groups except iC4 detected at this site. Moreover, in 26 of 29 purified *Petunia* metabolites, this 3-position was preferably substituted by the longest acyl group among a specific acylsucrose. In contrast, the 4-position in *P. exserta*, *P. axillaris* and 6-position in *P. integrifolia* is conserved by esterification by aiC5 groups, and aiC5 chains were also attached at the 2-position in all three species. Position-selective acylation by aiC5 and longer acyl groups (C6-C8) was similar to an earlier report of acylsucroses from *P. hybrida* [31]. Iso-C4 acylation at position 6 was observed in both *P. exserta* and *P. axillaris*, whereas the same substitution happened at the 2- and 4-positions in *P. integrifolia*. Only 3 purified acylsugars in *P. axillaris* contained aiC6 ester groups, all of them at the 3-position. The similar position-selectivity in attachment of acyl groups of *P. exserta* and *P. axillaris* is consistent with reported *Petunia* genus evolutionary analyses that these two species are more closely related than *P. integrifolia* [32].

To our knowledge, acylsugar related malonyltransferases have not been reported in *Petunia* glandular trichomes and we anticipate that identification of malonate acylsugars will aid efforts to discover these metabolic functions. Esterification by malonate esters at the 1´-position of the furanose ring was predominant in *P. integrifolia* and *P. axillaris*, while either 1´- or 6´-position in *P. exserta* was esterified by malonyl chain. Previous reports described similar findings for *P. hybrida* malonate acylsugars [15, 27], while we suggest that their observations of acetate groups at the 6´-position may come from malonyl decarboxylation during metabolite isolation.

Despite limited prior understanding of *Petunia* acylsugar biosynthetic mechanisms, some hypotheses can be generated by comparing with previous studies of tomato and closely related wild species [17,24,25]. The highly conserved acyl groups at specific sugar positions might be

attributed to the order of acylation catalyzed by position-selective acyltransferases or selective acyl chain removal and remodeling.

2.3.6 Diversity in acyl groups and implications for elongation mechanisms

Acyl CoA precursors of acylsugars are often derived from branched-chain ketoacids produced by metabolism of branched-chain amino acids, often involving chain elongation. Although wild tomato species use two-carbon elongation as is common for fatty acid biosynthesis, elongation in *Petunia* has been proposed to use a one-carbon elongation known as the α-KAE pathway [33]. Structure elucidation of *Petunia* acylsugars also assists the investigation of aliphatic acid elongation mechanisms. Confirmed by NMR data in this study, no acetate esters were observed in *Petunia* acylsugars (except via decarboxylation of malonate esters). Excluding malonate, the lengths of acyl chains in *Petunia* acylsugars ranged from C4 to C8, which is consistent with reported acylsugar profiles for other plants including *Nicotiana* spp. that use one-carbon metabolic elongation [33, 34]. At least two C5 moieties were present in all isolated *Petunia* acylsucroses, most of which are aiC5. Aside from the observation of aiC6 esters in *P. axillaris*, all other acyl groups longer than C5 were iso-branched, which could result from elongation of either valine-derived iC4 or leucine-derived iC5 to yield products as long as iC8 via either one- or two-carbon elongation. Anteiso-branching was only observed for aiC5 and aiC6 derived from isoleucine, the latter consistent with one-carbon elongation. Further experiments need to be performed to verify the proposed aliphatic acid elongation pathways in *Petunia* plants. Compared with structures of neutral acylsugars, malonate acylsugars only have one more position on the fructose ring that is acylated by a malonyl group, while all hydroxyls on the pyranose ring were esterified by aliphatic chains. Therefore, purified neutral acylsucroses in this

study could be used as substrates for potential screening of malonyltransferases in *Petunia* species.

2.4. Conclusions

A more comprehensive determination of diversity and complexity in *Petunia* acylsucrose metabolites has been reported. Structures have been assigned to 28 acylsucroses, including malonate esters that have never been reported in metabolite profiles of other *Solanaceous* plants. In the current study, these structures point to acylation position selectivity, particularly in the case of aiC5 on the pyranose ring within *P. exserta* and *P. axillaris* but less so in *P. integrifolia*. This result is consistent with phylogenetic analyses that *P. exserta* and *P. axillaris* are more closely related than *P. integrifolia*. The selectivity (and promiscuity) of acylation at specific positions is consistent with action of multiple acyltransferase enzymes and these findings should serve as the foundation for investigations into the functions of *Petunia* acyltransferase genes.

APPENDIX

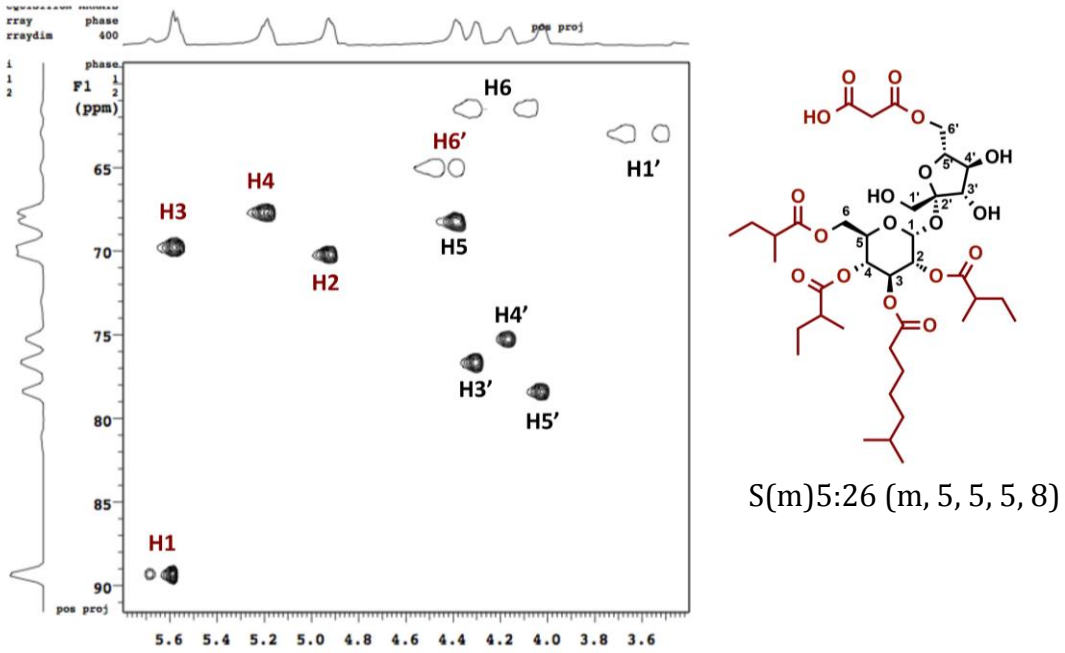


Figure S1. HSQC spectra for Petunia acylsugars led to assignment of Hs and Cs on sucrose ring.

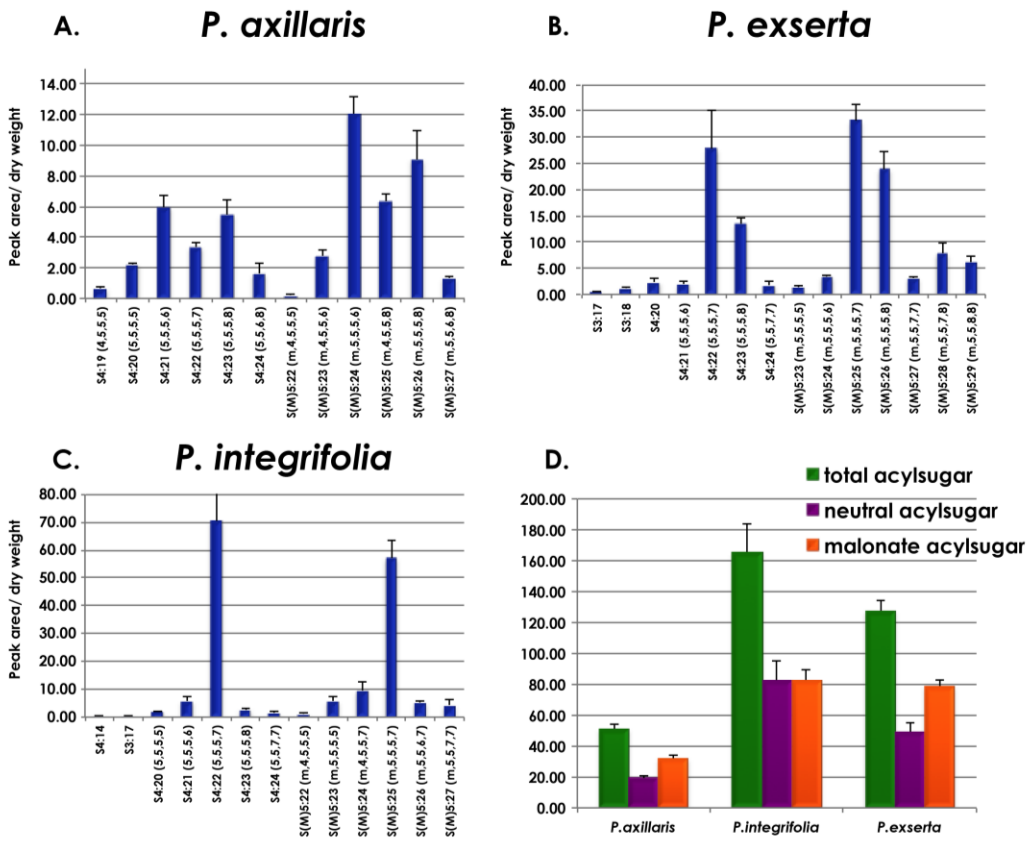


Figure S2. Quantitative analysis figures for neutral, malonate acylsugars in three Petunia species. *Three biological replicates for each species (N=3, mean ± SD). A. quantitative analysis result for P. axillaris; B. quantitative analysis result for P. exserta; C. quantitative analysis result for P. integrifolia; D. comparison of acylsugar abundance in three petunia species.*

Table S1. Plant growth metadata

Species	*Petunia axillaris* *Petunia integrifolia* *Petunia exserta*
Organ	Leaf
Organ specification	Leaflets. Leaflets were between 25%-75% of mature leaflet size.
Cell type	Glandular trichomes
Biosource amount	2 leaflets per species for metabolites profiling
Growth location	Plant greenhouse #20 , Michigan State University
Harvest date, time	December 3rd, 2013, around 17:00-19:00
Plant growth stage	8 weeks old, not flowering
Metabolism quenching method	Extract leaflets with isopropanol: acetonitrile: water (3:3:2, *v/v/v*) for 2 min
Harvest method	Leaflets were harvested using forceps and blade at the junction of stem and petiole.
Sample storage	Leaf tissues were extracted with isopropanol: acetonitrile: water (3:3:2, *v/v/v*) immediately after harvest. Extracts were transferred to 15 mL centrifuge tubes, then concentrated, redissolved, and transferred to 2 mL HPLC vials with 250 μL glass inserts. Samples were stored in -20 °C freezers.

Table S2. NMR data for purified acylsugars (chemical shift summary)

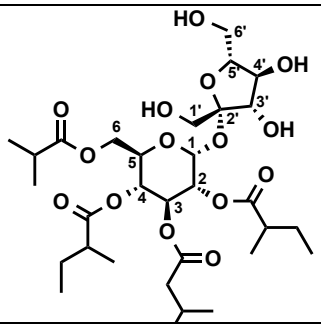
	**S4:19 (4,5,5,5)** **Purified from *P. axilaris*** **PubChem CID:** 91754396 **HRMS:** (ESI) *m/z* calculated for $C_{32}H_{53}O_{17}^-$ ([M+HCOO$^-$]): 709.3288, found: 709.3188 **Material recovered:** 0.5~1 mg	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.79 (d, *J*=3.9 Hz)	89.1[a]
2 (CH)	4.86 (dd, *J*=3.9, 10.0 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	176.8[c]
-2 (CH)	2.43[c,d] (m)	40.5[a]
-2' (CH_3)	1.16[c] (d, *J*=7.1 Hz)	16.1[a]
-3 (CH_2)	1.48 (m), 1.64[c,d] (m)	26.7[a]
-4 (CH_3)	0.89[b,d] (m)	11.2[c]
3 (CH)	5.57 (t, *J*=10.0 Hz)	69.0[a]
3-O-	–	–
-1 (CO)	–	172.6[c]
-2 (CH)	2.23[c,d] (m)	34.0[a]
-3 (CH)	1.54[c,d] (m)	24.4[a]
-4 (CH_3) x 2	0.95[a] (d, *J*=6.7 Hz)	22.2[c]
4 (CH)	5.11 (t, *J*=10.0 Hz)	67.6[a]
4-O-	–	–
-1 (CO)	–	175.0[c]
-2 (CH)	2.37[c,d] (m)	40.9[a]
-2' (CH_3)	1.11[c] (d, *J*=7.1 Hz)	16.6[a]
-3 (CH_2)	1.45, 1.67[c,d] (m)	26.7[a]
-4 (CH_3)	0.89[b,d] (m)	11.5[c]
5 (CH)	4.31[a,d] (m)	69.0[a]
6 (CH_2)	4.06 (m), 4.25[a,d] (m)	61.7[a]
6-O-	–	–
-1 (CO)	–	174.0[c]
-2 (CH)	2.41[c,d] (m)	33.7[a]
-3 (CH_3) x 2	1.16[c] (d, *J*=7.1 Hz)	18.9[a]
1' (CH_2)	3.54 (d, *J*=12.0 Hz), 3.63 (d, *J*=12.0 Hz)	65.0[a]
2' (C)	–	105.0[c]
3' (CH)	4.23[a] (d, *J*=8.2 Hz)	78.6[a]
4' (CH)	4.28[a] (t, *J*=8.2 Hz)	73.8[a]

Table S2. (cont'd)

5' (CH)	3.76[a,d] (m)	81.8[a]
6' (CH_2)	3.72 (dd, *J*=2.5, 12.8 Hz), 3.89[a] (dd, *J*=2.5, 12.8 Hz)	60.2[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

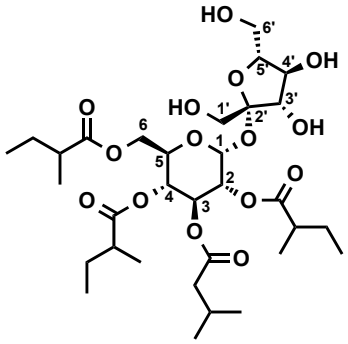
	**S4:20 [1] (5,5,5,5)** **Purified from *P. axillaris*** **PubChem CID**: 91754397 **HRMS:** (ESI) *m/z* calculated for $C_{33}H_{55}O_{17}^-$ ([M+HCOO$^-$]): 723.3445, found: 723.3367 **Material recovered:** 0.5~1 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.78 (d, *J*=3.8 Hz)	89.1[a]
2 (CH)	4.85 (dd, *J*=3.8, 10.0 Hz)	70.9[a]
2-O-	–	–
-1 (CO)	–	176.8[a]
-2 (CH)	2.44[d] (m)	40.78
-2' (CH_3)	1.15 (t, *J*=7.1 Hz)	16.3[a]
-3 (CH_2)	1.47 (m), 1.63[d] (m)	26.72
-4 (CH_3)	0.87[c,d] (m)	11.7[a]
3 (CH)	5.55 (t, *J*=10.0 Hz)	69.2[a]
3-O-	–	–
-1 (CO)	–	172.9[a]
-2 (CH)	2.23[d] (m)	33.9[a]
-3 (CH)	1.52[c,d] (m)	26.5[c]
-4 (CH_3) x 2	0.96[a] (d, *J*=6.6 Hz)	22.3[a]
4 (CH)	5.13 (t, *J*=10.0 Hz)	67.8[a]
4-O-	–	–
-1 (CO)	–	175.2[a]
-2 (CH)	2.37[d] (m)	40.91
-2' (CH_3)	1.11 (t, *J*=7.1 Hz)	16.51
-3 (CH_2)	1.45 (m), 1.67[d] (m)	26.38
-4 (CH_3)	0.89[c,d] (m)	13.8[a]
5 (CH)	4.32[a,d] (m)	69.1[a]
6 (CH_2)	4.07 (m), 4.29[a,d] (m)	61.5[a]
6-O-	–	–
-1 (CO)	–	176.4[a]

Table S2. (cont'd)

-2 (CH)	2.47 (m)	40.56
-2' (CH_3)	1.17 (t, *J*=7.1 Hz)	16.51
-3 (CH_2)	1.51 (m), 1.71[d] (m)	26.62
-4 (CH_3)	0.92[c,d] (m)	11.7[a]
1' (CH_2)	3.58 (d, *J*=12.0 Hz), 3.62 (d, *J*=12.0 Hz)	65.0[a]
2' (C)	–	104.98
3' (CH)	4.23[a] (d, *J*=8.2 Hz)	78.6[a]
4' (CH)	4.28[a] (t, *J*=8.2 Hz)	73.8[a]
5' (CH)	3.76[a,d] (m)	81.9[a]
6' (CH_2)	3.72 (dd, *J*=2.5, 12.8 Hz), 3.88[a] (dd, *J*=2.5, 12.8 Hz)	60.5[a]

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

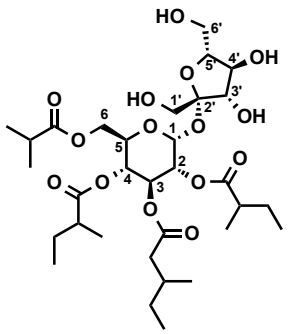
	**S4:20 [2] (4,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754398 **HRMS:** (ESI) *m/z* calculated for $C_{33}H_{55}O_{17}^-$ ([M+HCOO$^-$]): 723.3445, found: 723.3311 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.78 (d, *J*=3.8 Hz)	89.0[a]
2 (CH)	4.87 (dd, *J*=3.8, 10.0 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	176.9[c]
-2 (CH)	2.43[c,d] (m)	40.5[a]
-2' (CH_3)	1.17[c,d] (m)	15.9[a]
-3 (CH_2)	1.48 (m), 1.63[c,d] (m)	26.5[a]
-4 (CH_3)	0.92[c,d] (m)	11.4[a]
3 (CH)	5.55 (t, *J*=10.0 Hz)	68.9[a]
3-O-	–	–
-1 (CO)	–	172.9[c]
-2 (CH_2)	2.24[c] (d, *J*=7.2 Hz)	33.8[a]
-3 (CH)	1.54[c,d] (m)	24.3[a]
-3' (CH_3)	0.97[c,d] (m)	22.4[c]
-4 (CH_2)	1.31[c,d] (m)	22.2[a]
-5 (CH_3)	0.94[c,d] (m)	13.6[a]
4 (CH)	5.13 (t, *J*=10.0 Hz)	67.5[a]
4-O-	–	–
-1 (CO)	–	175.0[c]
-2 (CH)	2.36[c,d] (m)	40.7[a]
-2' (CH_3)	1.12[c,d] (m)	16.6[a]
-3 (CH_2)	1.45 (m), 1.67[c,d] (m)	26.5[a]
-4 (CH_3)	0.87[c,d] (m)	13.6[a]
5 (CH)	4.33[a,d] (m)	68.8[a]
6 (CH_2)	4.08 (m), 4.26[a,d] (m)	61.4[a]

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	173.5[c]
-2 (CH)	2.37[c,d] (m)	35.6[a]
-3 (CH_3) x 2	1.13[c] (d, *J*=6.7 Hz)	18.8[a]
1' (CH_2)	3.57 (d, *J*=11.9 Hz), 3.62 (d, *J*=11.9 Hz)	64.7[a]
2' (C)	–	102.42
3' (CH)	4.23[a] (d, *J*=8.0 Hz)	78.4[a]
4' (CH)	4.29[a] (t, *J*=8.0 Hz)	73.6[a]
5' (CH)	3.77[a,d] (m)	81.6[a]
6' (CH_2)	3.72 (dd, *J*=2.7, 12.9 Hz), 3.88[a] (dd, *J*=2.7, 12.9 Hz)	60.2[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

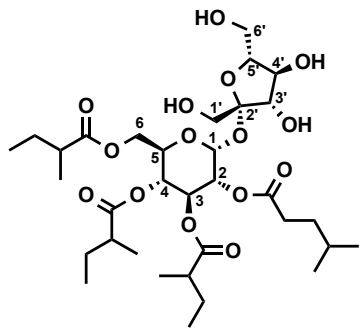
	**S4:21 [1] (5,5,5,6)** **Purified from *P. integrifolia*** **PubChem CID:** 91754399 **HRMS:** (ESI) *m/z* calculated for $C_{34}H_{57}O_{17}^-$ ([M+HCOO$^-$]): 737.3601, found: 737.3485 **Material recovered:** 0.5~0.8 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.82 (d, *J*=3.8 Hz)	89.0[a]
2 (CH)	4.84 (dd, *J*=3.8, 10.0 Hz)	70.8[a]
2-O-	–	–
-1 (CO)	–	172.6[c]
-2 (CH_2)	2.22[d] (m)	34.1[a]
-3 (CH_2)	1.52[c] (d, *J*=6.6 Hz)	22.7[a]
-4 (CH)	1.16[c,d] (m)	18.8[a]
-5 (CH_3) x 2	0.88[c,d] (m)	22.2[a]
3 (CH)	5.55 (t, *J*=10.0 Hz)	68.8[a]
3-O-	–	–
-1 (CO)	–	175.4[c]
-2 (CH)	2.56[d] (m)	40.7[a]
-2' (CH_3)	1.14[d] (m)	16.1[a]
-3 (CH_2)	1.45 (m), 1.61[c,d] (m)	26.9[c]
-4 (CH_3)	0.88 (d, *J*=7.5 Hz)	11.6[a]
4 (CH)	5.10 (t, *J*=10.0 Hz)	64.62
4-O-	–	–
-1 (CO)	–	176.9[c]
-2 (CH)	2.42[d] (m)	40.3[c]
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.6[a]
-3 (CH_2)	1.48 (m), 1.63[a,d] (m)	26.5[c]
-4 (CH_3)	0.88[d] (m)	14.0[a]
5 (CH)	4.30[a,d] (m)	69.0[a]
6 (CH_2)	4.09 (m), 4.24[d] (m)	61.6[a]
3-O-	–	–
-1 (CO)	–	176.3[c]

Table S2. (cont'd)

-2 (CH)	2.48[d] (m)	41.0[c]
-2' (CH_3)	1.11[c,d] (m)	16.4[c]
-3 (CH_2)	1.48 (m), 1.71[c,d] (m)	26.7[a]
-4 (CH_3)	0.92[d] (m)	12.0[a]
1' (CH_2)	3.59 (d, *J*=11.8 Hz), 3.66 (d, *J*=11.8 Hz)	64.9[a]
2' (C)	–	104.5[c]
3' (CH)	4.22 (d, *J*=9.2 Hz)	78.5[a]
4' (CH)	4.29[a,d] (m)	73.9[a]
5' (CH)	3.75[a,d] (m)	81.4[a]
6' (CH_2)	3.71 (m), 3.89[d] (m)	60.1[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

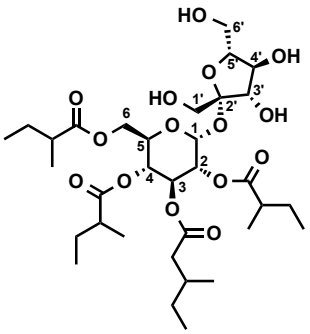
	**S4:21 [2] (5,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754400 **HRMS:** (ESI) *m/z* calculated for $C_{34}H_{57}O_{17}^-$ ([M+HCOO$^-$]): 737.3601, found: 737.3360 **Material recovered:** 2~3 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.76 (d, *J*=3.8 Hz)	89.21
2 (CH)	4.87 (dd, *J*=3.8, 10.0 Hz)	70.71
2-O-	–	–
-1 (CO)	–	176.80
-2 (CH)	2.41[d] (m)	40.54
-2' (CH_3)	1.14 (d, *J*=7.0 Hz)	15.97
-3 (CH_2)	1.48 (m), 1.63[c,d] (m)	26.75
-4 (CH_3)	0.86[d] (m)	11.39
3 (CH)	5.57 (t, *J*=10.0 Hz)	69.45
3-O-	–	–
-1 (CO)	–	173.28
-2 (CH_2)	2.23 (d, *J*=6.9 Hz)	32.07
-3 (CH)	1.55[d] (m)	24.22
-3' (CH_3)	0.87[d] (m)	22.14
-4 (CH_2)	1.42 (m), 1.45[c,d] (m)	33.38
-5 (CH_3)	0.96[d] (m)	11.56
4 (CH)	5.16 (t, *J*=10.0 Hz)	67.77
4-O-	–	–
-1 (CO)	–	175.04
-2 (CH)	2.36[d] (m)	40.91
-2' (CH_3)	1.10 (d, *J*=7.0 Hz)	16.42
-3 (CH_2)	1.45 (m), 1.67[c,d] (m)	26.39
-4 (CH_3)	0.88[d] (m)	13.84
5 (CH)	4.35[d] (m)	68.89
6 (CH_2)	4.04 (m), 4.31[d] (m)	61.46

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	176.52
-2 (CH)	2.45[d] (m)	40.81
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.37
-3 (CH_2)	1.48 (m), 1.72[c,d] (m)	26.61
-4 (CH_3)	0.91[d] (m)	11.46
1' (CH_2)	3.54 (d, *J*=11.9 Hz), 3.63 (d, *J*=11.9 Hz)	64.68
2' (C)	–	104.85
3' (CH)	4.26[a] (d, *J*=8.1 Hz)	78.56
4' (CH)	4.28[a] (t, *J*=8.1 Hz)	73.65
5' (CH)	3.78[d] (m)	81.9[a]
6' (CH_2)	3.72(dd, *J*=2.6, 13.1 Hz), 3.88 (dd, *J*=2.6, 13.1 Hz)	60.62
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

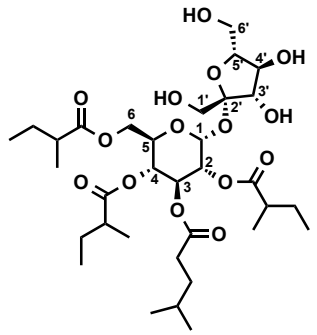
	**S4:21 [3] (5,5,5,6)** **Purified from *P. exserta*** **PubChem CID:** 91754401 **HRMS:** (ESI) *m/z* calculated for $C_{34}H_{57}O_{17}^-$ ([M+HCOO$^-$]): 737.3601, found: 737.3549 **Material recovered:** 0.2~0.4 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.71 (d, *J*=3.7 Hz)	89.34[a]
2 (CH)	4.86 (dd, *J*=3.7, 10.0 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	176.72[c]
-2 (CH)	2.41[d] (m)	40.66[a]
-2' (CH_3)	1.13 (d, *J*=7.0 Hz)	16.06[a]
-3 (CH_2)	1.48 (m), 1.62[c,d] (m)	26.83[c]
-4 (CH_3)	0.86[d] (m)	11.51[a]
3 (CH)	5.55 (t, *J*=10.0 Hz)	69.60[a]
3-O-	–	–
-1 (CO)	–	173.39[c]
-2 (CH_2)	2.22[d] (m)	34.13[c]
-3 (CH_2)	1.54[d] (m)	24.48[a]
-4 (CH)	1.28 (br)	31.29[a]
-5 (CH_3) x 2	0.95 (d, *J*=6.6 Hz)	22.36[a]
4 (CH)	5.14 (t, *J*=10.0 Hz)	67.68[a]
4-O-	–	–
-1 (CO)	–	175.20[c]
-2 (CH)	2.36[d] (m)	41.03[a]
-2' (CH_3)	1.09 (d, *J*=7.0 Hz)	16.54[a]
-3 (CH_2)	1.44 (m), 1.66[a,d] (m)	26.48[a]
-4 (CH_3)	0.87[d] (m)	13.95[c]
5 (CH)	4.34[d] (m)	68.81[a]
6 (CH_2)	4.06 (m), 4.28[d] (m)	61.44[a]
6-O-	–	–

Table S2. (cont'd)

-1 (CO)	–	176.72[c]
-2 (CH)	2.45[d] (m)	40.96
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.49[c]
-3 (CH_2)	1.48 (m), 1.73[c,d] (m)	26.71[a]
-4 (CH_3)	0.91[d] (m)	11.66[a]
1' (CH_2)	3.51 (d, *J*=12.0 Hz), 3.61 (d, *J*=12.0 Hz)	64.39[a]
2' (C)	–	104.79[c]
3' (CH)	4.26[a] (d, *J*=9.4 Hz)	78.00[a]
4' (CH)	4.26[a] (t, *J*=9.4 Hz)	73.80[a]
5' (CH)	3.79[d] (m)	81.9[a]
6' (CH_2)	3.72(dd, *J*=2.9, 13.1 Hz), 3.86 (dd, *J*=3.0, 13.1 Hz)	60.87[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

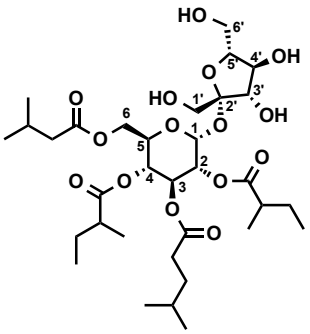
	**S4:21 [4] (5,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754402 **HRMS:** (ESI) *m/z* calculated for $C_{34}H_{57}O_{17}^-$ ([M+HCOO$^-$]): 737.3601, found: 737.3375 **Material recovered:** 1~1.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.76 (d, *J*=4.2 Hz)	89.11
2 (CH)	4.86 (dd, *J*=4.2, 10.0 Hz)	70.77
2-O-	–	–
-1 (CO)	–	177.0[c]
-2 (CH)	2.43[d] (m)	40.54
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.01
-3 (CH_2)	1.48 (m), 1.62[c,d] (m)	26.70
-4 (CH_3)	0.87[b,d] (m)	11.38
3 (CH)	5.55 (t, *J*=10.0 Hz)	69.31
2-O-	–	–
-1 (CO)	–	173.0[c]
-2 (CH_2)	2.23[d] (m)	34.04
-3 (CH_2)	1.54[d] (m)	24.28
-4 (CH)	1.29 (br, *J*=6.1 Hz)	29.13
-5 (CH_3) x 2	0.89 (d, *J*=6.6 Hz)	22.1[a]
4 (CH)	5.12 (t, *J*=10.0 Hz)	67.72
4-O-	–	–
-1 (CO)	–	175.1[c]
-2 (CH)	2.36[d] (m)	40.93
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.45
-3 (CH_2)	1.44 (m), 1.66[a,d] (m)	26.37
-4 (CH_3)	0.91[b,d] (m)	11.60
5 (CH)	4.32[a,d] (m)	69.05
6 (CH_2)	4.11 (m), 4.29[a,d] (m)	61.74
6-O-	–	–

Table S2. (cont'd)

-1 (CO)	–	174.2[c]
-2 (CH_2)	2.41[c,d] (m)	43.3[a]
-3 (CH)	1.51[c,d] (m)	24.66
-4 (CH_3) x 2	0.95 (d, *J*=6.6 Hz)	22.6[a]
1' (CH_2)	3.56 (d, *J*=12.0 Hz), 3.63 (d, *J*=12.0 Hz)	64.80
2' (C)	–	105.09
3' (CH)	4.25[a] (d, *J*=8.3 Hz)	78.61
4' (CH)	4.27[a] (t, *J*=8.1 Hz)	73.56
5' (CH)	3.75[a,d] (m)	81.81
6' (CH_2)	3.73 (dd, *J*=2.8, 12.6 Hz), 3.88[a] (dd, *J*=2.8, 12.6 Hz)	60.41
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

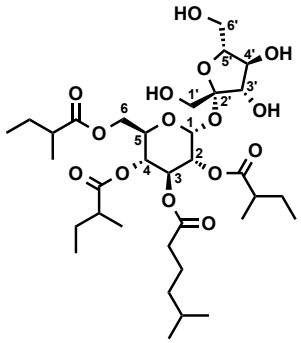
	**S4:22 [1] (5,5,5,7)** **Purified from *P. exserta*** **PubChem CID:** 91754403 **HRMS:** (ESI) *m/z* calculated for $C_{35}H_{59}O_{17}^-$ ([M+HCOO$^-$]): 751.3752, found: 751.3502 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.81 (d, *J*=4.0 Hz)	89.02
2 (CH)	4.84 (dd, *J*=4.0, 10.0 Hz)	70.69
2-O-	–	–
-1 (CO)	–	176.71
-2 (CH)	2.43[d] (m)	40.62
-2' (CH_3)	1.14 (d, *J*=7.1 Hz)	16.0[a]
-3 (CH_2)	1.48 (m), 1.65[c,d] (m)	26.7[c]
-4 (CH_3)	0.88[d] (m)	11.5[a]
3 (CH)	5.56 (t, *J*=10.0 Hz)	69.89
3-O-	–	–
-1 (CO)	–	173.14
-2 (CH_2)	2.22 (t, *J*=7.1 Hz)	34.2[a]
-3 (CH_2)	1.54[d] (m)	24.65
-4 (CH_2)	1.27 (br)	28.82
-5 (CH_2)	1.26 (br)	31.42
-6 (CH_3) x 2	0.88 (d, *J*=6.6 Hz)	22.49
4 (CH)	5.11 (t, *J*=10.0 Hz)	67.69
4-O-	–	–
-1 (CO)	–	175.03
-2 (CH)	2.37[d] (m)	40.93
-2' (CH_3)	1.11 (d, *J*=7.1 Hz)	16.4[a]
-3 (CH_2)	1.44 (m), 1.67[a,d] (m)	26.4[c]
-4 (CH_3)	0.90[d] (m)	11.2[a]
5 (CH)	4.31[d] (m)	69.01
6 (CH_2)	4.06 (m), 4.28[d] (m)	61.58

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	176.56
-2 (CH)	2.47[d] (m)	40.75
-2' (CH_3)	1.17 (d, *J*=7.1 Hz)	16.4[a]
-3 (CH_2)	1.51 (m), 1.72[c,d] (m)	26.6[c]
-4 (CH_3)	0.93[d] (m)	11.6[a]
1' (CH_2)	3.58 (d, *J*=11.8 Hz), 3.64 (d, *J*=11.8 Hz)	64.49
2' (C)	–	105.06
3' (CH)	4.21[a] (d, *J*=8.2 Hz)	78.72
4' (CH)	4.29[a] (t, *J*=8.2 Hz)	73.80
5' (CH)	3.75[d] (m)	81.71
6' (CH_2)	3.72 (dd, *J*=2.8, 12.9 Hz), 3.89 (dd, *J*=2.8, 12.9 Hz)	60.14
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

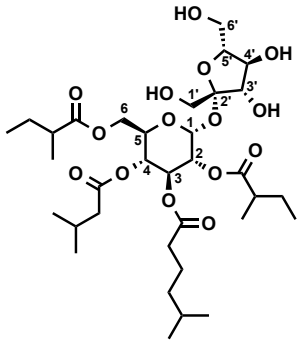
	**S4:22 [2] (5,5,5,7)** **Purified from *P. integrifolia*** **PubChem CID:** 91754404 **HRMS:** (ESI) *m/z* calculated for $C_{35}H_{59}O_{17}^-$ ([M+HCOO$^-$]): 751.3752, found: 751.3555 **Material recovered:** 0.2~0.4 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.81 (d, *J*=4.0 Hz)	88.9[a]
2 (CH)	4.84 (dd, *J*=4.0, 10.1 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	177.2[c]
-2 (CH)	2.41[d] (m)	41.0[c]
-2' (CH_3)	1.14 (d, *J*=7.0 Hz)	16.5[a]
-3 (CH_2)	1.48 (m), 1.64[c,d] (m)	26.8[c]
-4 (CH_3)	0.88[d] (m)	11.4[a]
3 (CH)	5.55 (t, *J*=10.1 Hz)	68.9[a]
3-O-	–	–
-1 (CO)	–	172.8[c]
-2 (CH_2)	2.23 (t, *J*=7.3 Hz)	34.1[a]
-3 (CH_2)	1.54[d] (m)	22.5[a]
-4 (CH_2)	1.28[c] (br)	28.8[c]
-5 (CH)	1.25[c] (br)	31.2[c]
-6 (CH_3) x 2	0.88 (d, *J*=6.7 Hz)	22.4[a]
4 (CH)	5.11 (t, *J*=10.1 Hz)	67.7[a]
4-O-	–	–
-1 (CO)	–	175.6[c]
-2 (CH_2)	2.54[c,d] (m)	40.8[a]
-3 (CH)	1.51[c,d] (m)	24.7[a]
-4 (CH_3) x 2	0.94 (d, *J*=6.7 Hz)	22.4[a]
5 (CH)	4.29[a,d] (m)	69.5[a]
6 (CH_2)	4.12 (m), 4.29[d] (m)	61.6[a]
6-O-	–	–

Table S2. (cont'd)

-1 (CO)	–	176.49[c]
-2 (CH)	2.47[d] (m)	40.7[c]
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.2[a]
-3 (CH_2)	1.50 (m), 1.71[c,d] (m)	26.6[c]
-4 (CH_3)	0.93[d] (m)	11.5[a]
1' (CH_2)	3.58 (d, *J*=10.6 Hz), 3.65 (d, *J*=10.6 Hz)	65.2[a]
2' (C)	–	102.57[c]
3' (CH)	4.20[a] (br)	78.7[a]
4' (CH)	4.28[a] (br)	73.8[a]
5' (CH)	3.74[a,d] (m)	81.7[a]
6' (CH_2)	3.73 (m), 3.91[d] (m)	60.3[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

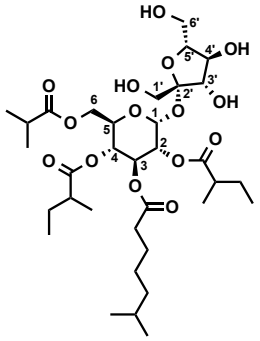
	**S4:22 [3] (4,5,5,8)** **Purified from *P. exserta*** **PubChem CID:** 91754405 **HRMS:** (ESI) *m/z* calculated for $C_{35}H_{59}O_{17}^-$ ([M+HCOO$^-$]): 751.3752, found: 751.3762 **Material recovered:** 0.2~0.4 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.80 (d, *J*=4.0 Hz)	89.1[a]
2 (CH)	4.87 (dd, *J*=4.0, 10.0 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	177.12
-2 (CH)	2.43[c,d] (m)	40.7[a]
-2' (CH_3)	1.2 (d, *J*=7.0 Hz)	16.2[a]
-3 (CH_2)	1.49 (m), 1.65[c,d] (m)	26.8[c]
-4 (CH_3)	0.89[c] (t, *J*=7.2 Hz)	11.6[a]
3 (CH)	5.54 (t, *J*=10.0 Hz)	69.0[a]
3-O-	–	–
-1 (CO)	–	172.9[c]
-2 (CH_2)	2.23 (t, *J*=7.6 Hz)	34.0[a]
-3 (CH_2)	1.55[d] (m)	24.6[c]
-4 (CH_2)	1.28 (br)[a]	29.0[a]
-5 (CH_2)	1.30 (br)[a]	31.5[a]
-6 (CH)	1.31[a] (br)	22.9[a]
-7 (CH_3) x 2	0.90 (d, *J*=6.6 Hz)	22.7[a]
4 (CH)	5.09 (t, *J*=10.0 Hz)	67.6[a]
4-O-	–	–
-1 (CO)	–	175.3[c]
-2 (CH)	2.37[c,d] (m)	41.0[a]
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.5[a]
-3 (CH_2)	1.45 (m), 1.69[c,d] (m)	26.3[c]
-4 (CH_3)	0.93[c] (t, *J*=7.2 Hz)	11.6[a]
5 (CH)	4.28[a,d] (m)	69.1[a]

Table S2. (cont'd)

6 (CH_2)	4.10 (m), 4.26[a,d] (m)	61.4[a]
6-O-	–	–
-1 (CO)	–	173.4[c]
-2 (CH)	2.38[c,d] (m)	40.8[a]
-3 (CH_3) x 2	1.17[c] (d, *J*=7.0 Hz)	18.9[a]
1' (CH_2)	3.59 (d, *J*=12.2 Hz), 3.64 (d, *J*=12.2 Hz)	65.0[a]
2' (C)	–	105.0[c]
3' (CH)	4.20[a] (d, *J*=8.1 Hz)	78.82
4' (CH)	4.28[a] (t, *J*=8.1 Hz)	73.81
5' (CH)	3.74[a,d] (m)	81.82
6' (CH_2)	3.72 (dd, *J*=2.5, 13.1 Hz), 3.89[a] (dd, *J*=2.5, 13.1 Hz)	60.23
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

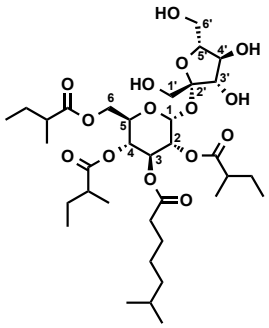
	**S4:23 (5,5,5,8)** **Purified from *P. exserta*** **PubChem CID:** 91754406 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{61}O_{17}^-$ ([M+HCOO$^-$]):765.3909, found: 765.3982 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.75 (d, *J*=3.8 Hz)	89.12
2 (CH)	4.87 (dd, *J*=3.8, 10.2 Hz)	70.68
2-O-	–	–
-1 (CO)	–	176.95
-2 (CH)	2.42[c,d] (m)	40.66
-2' (CH_3)	1.14 (d, *J*=7.01 Hz)	16.09
-3 (CH_2)	1.46 (m), 1.62[c,d] (m)	26.85
-4 (CH_3)	0.86 (t, *J*=7.4 Hz)	11.49
3 (CH)	5.56 (t, *J*=10.2 Hz)	69.25
3-O-	–	–
-1 (CO)	–	173.18
-2 (CH)	2.22 (t, *J*=7.7 Hz)	34.13
-3 (CH_2)	1.55[d] (m)	24.61
-4 (CH_2)	1.27 (br)[a]	29.23
-5 (CH_2)	1.26 (br)[a]	31.73
-6 (CH)	1.33 (br)[a]	29.00
-7 (CH_3) x 2	0.86 (d, *J*=6.6 Hz)	22.71
4 (CH)	5.14 (t, *J*=10.2 Hz)	67.60
4-O-	–	–
-1 (CO)	–	175.17
-2 (CH)	2.36[c,d] (m)	41.04
-2' (CH_3)	1.1 (d, *J*=7.0 Hz)	16.57
-3 (CH_2)	1.43 (m), 1.66[c,d] (m)	26.49
-4 (CH_3)	0.88 (t, *J*=7.4 Hz)	14.18
5 (CH)	4.31[a,d] (m)	68.81

Table S2. (cont'd)

6 (CH_2)	4.07 (m), 4.34[a,d] (m)	61.34
6-O-	–	–
-1 (CO)	–	176.61
-2 (CH)	2.46[c,d] (m)	40.92
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.51
-3 (CH_2)	1.49 (m), 1.71[c,d] (m)	26.74
-4 (CH_3)	0.92 (t, *J*=7.4 Hz)	11.68
1' (CH_2)	3.55 (d, *J*=11.9 Hz), 3.63 (d, *J*=11.9 Hz)	64.75
2' (C)	–	105.01
3' (CH)	4.24[a] (d, *J*=8.0 Hz)	78.47
4' (CH)	4.28[a] (t, *J*=8.0 Hz)	73.64
5' (CH)	3.78[d] (m)	81.77
6' (CH_2)	3.74 (dd, *J*=2.7, 12.8 Hz), 3.88 (dd, *J*=2.7, 12.8 Hz)	60.46
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

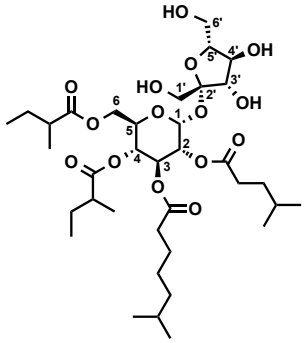
	**S4:24 [1] (5,5,6,8)** **Purified from *P. axillaris*** **PubChem CID:** 91754407 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{63}O_{17}^-$ ([M+HCOO$^-$]):779.4065, found: 779.3806 **Material recovered:** 0.5~1 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.80 (d, *J*=4.1 Hz)	89.0[a]
2 (CH)	4.87 (dd, *J*=4.1, 9.7 Hz)	71.0[a]
2-O-	–	–
-1 (CO)	–	173.5[c]
-2 (CH_2)	2.36[d] (m)	33.7[a]
-3 (CH_2)	1.54[d] (m)	24.8[c]
-4 (CH)	1.30 (br)	31.2[a]
-5 (CH_3) x 2	0.96 (d, *J*=6.6 Hz)	22.4[a]
3 (CH)	5.61 (t, *J*=9.7 Hz)	69.4[a]
3-O-	–	–
-1 (CO)	–	172.7[c]
-2 (CH_2)	2.22[d] (m)	34.1[a]
-3 (CH_2)	1.54[d] (m)	24.5[c]
-4 (CH_2)	1.28 (br)[a]	28.8[c]
-5 (CH_2)	1.32 (br)[a]	31.3[c]
-6 (CH)	1.30 (br)[a]	29.8[c]
-7 (CH_3) x 2	0.90 (d, *J*=6.6 Hz)	22.4[a]
4 (CH)	5.13 (t, *J*=9.7 Hz)	68.0[a]
4-O-	–	–
-1 (CO)	–	175.1[c]
-2 (CH)	2.34[c,d] (m)	41.0[a]
-2' (CH_3)	1.09 (d, *J*=7.0 Hz)	16.6[a]
-3 (CH_2)	1.42 (m), 1.65[c,d] (m)	26.4[a]
-4 (CH_3)	0.87[c] (t, *J*=7.4 Hz)	11.6[a]
5 (CH)	4.27[a] (br)	67.6[a]

Table S2. (cont'd)

6 (CH_2)	4.13 (m), 4.24[a,d] (m)	61.6[a]
6-O-	–	–
-1 (CO)	–	176.3[c]
-2 (CH)	2.43[c,d] (m)	40.9[a]
-2' (CH_3)	1.14 (d, *J*=7.0 Hz)	16.5[a]
-3 (CH_2)	1.47 (m), 1.71[c,d] (m)	26.7[c]
-4 (CH_3)	0.91[c] (t, *J*=7.4 Hz)	11.6[a]
1' (CH_2)	3.58 (d, *J*=13.6 Hz), 3.64 (d, *J*=13.6 Hz)	64.9[a]
2' (C)	–	102.61
3' (CH)	4.20[a] (d, *J*=9.9 Hz)	78.4[a]
4' (CH)	4.25[a] (br)	73.9[a]
5' (CH)	3.78[a,d] (m)	81.5[a]
6' (CH_2)	3.74 (dd, *J*=2.5, 13.6 Hz), 3.88[a] (dd, *J*=2.5, 13.6 Hz)	60.2[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

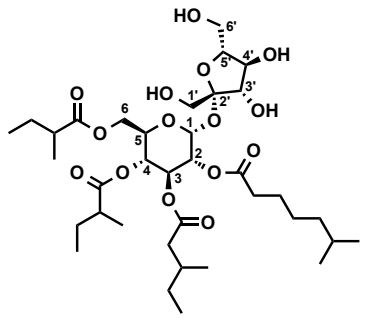
	**S4:24 [2] (5,5,6,8)** **Purified from *P. axillaris*** **PubChem CID:** 91754408 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{63}O_{17}^-$ ([M+HCOO$^-$]):779.4065, found: 779.3818 **Material recovered:** 0.8~1 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.79 (d, *J*=3.8 Hz)	89.7[a]
2 (CH)	4.88 (dd, *J*=3.8, 9.8 Hz)	71.7[a]
2-O-	–	–
-1 (CO)	–	173.5[c]
-2 (CH)	2.38 (t, *J*=7.9 Hz)	35.5[a]
-3 (CH_2)	1.55[d] (m)	24.3[a]
-4 (CH_2)	1.27 (br)[a]	30.3[a]
-5 (CH_2)	1.30 (br)[a]	33.1[a]
-6 (CH)	1.26 (br)[a]	30.3[a]
-7 (CH_3) x 2	0.86 (d, *J*=6.6 Hz)	23.2[a]
3 (CH)	5.62 (t, *J*=9.8 Hz)	70.0[a]
3-O-	–	–
-1 (CO)	–	172.8[c]
-2 (CH_2)	2.23[c] (d, *J*=6.8 Hz)	33.4[a]
-3 (CH)	1.54[c,d] (m)	24.3[a]
-3' (CH_3)	0.95[c,d] (m)	23.8[c]
-4 (CH_2)	1.43 (m), 1.47[c,d] (m)	34.8[a]
-5 (CH_3)	0.87[c,d] (m)	15.3[a]
4 (CH)	5.14 (t, *J*=9.8 Hz)	68.7[a]
4-O-	–	–
-1 (CO)	–	175.10
-2 (CH)	2.35[c,d] (m)	41.0[c]
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.5[c]
-3 (CH_2)	1.44 (m), 1.66[c,d] (m)	26.4[c]
-4 (CH_3)	0.87[c] (t, *J*=7.4 Hz)	11.5[c]

Table S2. (cont'd)

5 (CH)	4.28[a,d] (m)	68.2[a]
6 (CH_2)	4.08, 4.24[a,d] (m)	62.4[a]
6-O-	–	–
-1 (CO)	–	176.38
-2 (CH)	2.43[c,d] (m)	40.9[c]
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.5[c]
-3 (CH_2)	1.48 (m), 1.72[c,d] (m)	26.6[c]
-4 (CH_3)	0.89[c] (t, *J*=7.4 Hz)	11.5[c]
1' (CH_2)	3.57 (d, *J*=11.9 Hz), 3.63 (d, *J*=11.9 Hz)	65.6[a]
2' (C)	–	105.9[c]
3' (CH)	4.21[a] (d, *J*=8.1 Hz)	79.4[a]
4' (CH)	4.28[a] (t, *J*=8.1 Hz)	74.4[a]
5' (CH)	3.75[a,d] (m)	82.4[a]
6' (CH_2)	3.72 (dd, *J*=2.2, 13.1 Hz), 3.88[a] (dd, *J*=2.2, 13.1 Hz)	61.1[a]

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

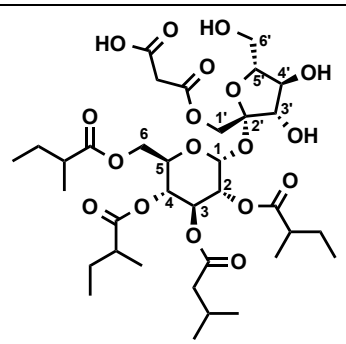
	**S(m) 5:23 [1] (m,5,5,5,5)** **Purified from *P. exserta*** **PubChem CID:** 91754409 **HRMS:** (ESI) *m/z* calculated for $C_{35}H_{55}O_{18}^-$ ([M-H$^-$]): 763.3388, found: 763.3298 **Material recovered:** 0.2~0.4 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.82 (d, *J*=3.7 Hz)	88.9[a]
2 (CH)	4.80 (dd, *J*=3.7, 9.5 Hz)	70.9[a]
2-O-	–	–
-1 (CO)	–	177.4[a]
-2 (CH)	2.43[c,d] (m)	40.7[a]
-2' (CH_3)	1.16[c] (d, *J*=7.0 Hz)	16.0[a]
-3 (CH_2)	1.47 (m), 1.65[c,d] (m)	26.6[a]
-4 (CH_3)	0.88[c,d] (m)	11.2[a]
3 (CH)	5.57 (t, *J*=9.5 Hz)	69.1[a]
3-O-	–	–
-1 (CO)	–	173.6[a]
-2 (CH)	2.23[c,d] (m)	34.0[a]
-3 (CH)	1.53[c,d] (m)	24.6[c]
-4 (CH_3) x 2	0.95[c] (d, *J*=7.3 Hz)	22.3[a]
4 (CH)	5.14 (t, *J*=9.5 Hz)	67.6[a]
4-O-	–	–
-1 (CO)	–	175.4[a]
-2 (CH)	2.37[d] (m)	40.9[a]
-2' (CH_3)	1.10 (d, *J*=7.0 Hz)	16.4[a]
-3 (CH_2)	1.45 (m), 1.68[d] (m)	26.4[a]
-4 (CH_3)	0.90[c,d] (m)	11.3[a]
5 (CH)	4.34[a,d] (m)	68.7[a]
6 (CH_2)	4.11 (m), 4.26[a,d] (m)	61.6[a]
6-O-	–	–
-1 (CO)	–	176.5[a]

Table S2. (cont'd)

-2 (CH)	2.47[c,d] (m)	40.7[a]
-2' (CH_3)	1.18[c] (d, *J*=7.0 Hz)	16.3[a]
-3 (CH_2)	1.51 (m), 1.72[c,d] (m)	26.5[a]
-4 (CH_3)	0.93[c,d] (m)	11.3[a]
1' (CH_2)	4.13 (m), 4.23[a,d] (m)	65.0[a]
1'-O-	–	–
-1 (CO)	–	166.4[c]
-2 (CH_2)	3.48 (s)	41.5[a]
-3 (CO)	–	170.5[c]
2' (C)	–	104.98[c]
3' (CH)	4.25[a] (br)	78.5[a]
4' (CH)	4.29[a] (br)	73.7[a]
5' (CH)	3.77[a,d] (m)	81.6[a]
6' (CH_2)	3.75 (m), 3.90[a,d] (m)	60.2[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

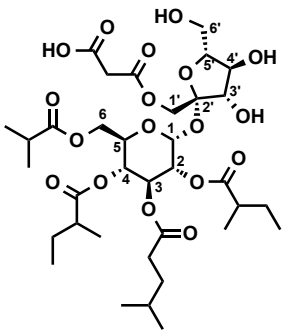
	**S(m) 5:23 [2] (m,4,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754410 **HRMS:** (ESI) *m/z* calculated for $C_{35}H_{55}O_{18}^-$ ([M-H$^-$]): 763.3388, found: 763.3331 **Material recovered:** 0.5~1 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.74 (d, *J*=3.8 Hz)	89.2[a]
2 (CH)	4.84 (dd, *J*=3.8, 9.9 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	177.2[c]
-2 (CH)	2.45[c,d] (m)	40.7[a]
-2' (CH_3)	1.20[c,d] (m)	15.8[c]
-3 (CH_2)	1.48 (m), 1.64[c,d] (m)	26.5[c]
-4 (CH_3)	0.95[c,d] (m)	11.2[a]
3 (CH)	5.55 (t, *J*=9.9 Hz)	69.3[a]
3-O-	–	–
-1 (CO)	–	173.0[c]
-2 (CH_2)	2.24[c] (d, *J*=7.0 Hz)	33.9[a]
-3 (CH_2)	1.54[c,d] (m)	24.7[a]
-4 (CH_2)	1.31[c,d] (m)	22.3[a]
-5 (CH_3) x 2	0.87[c,d] (m)	13.9[a]
4 (CH)	5.17 (t, *J*=9.9 Hz)	67.6[a]
4-O-	–	–
-1 (CO)	–	175.1[c]
-2 (CH)	2.36[c,d] (m)	41.0[a]
-2' (CH_3)	1.10[c,d] (m)	16.2[c]
-3 (CH_2)	1.45 (m), 1.66[c,d] (m)	26.2[c]
-4 (CH_3)	0.99[c,d] (m)	11.3[a]
5 (CH)	4.35[a,d] (m)	68.8[a]
6 (CH_2)	4.07 (m), 4.32[a,d] (m)	61.3[a]
6-O-	–	–

Table S2. (cont'd)

-1 (CO)	–	173.2[c]
-2 (CH)	2.39[c,d] (m)	35.9[a]
-3 (CH_3) x 2	1.13[c] (d, *J*=7.0 Hz)	19.0[a]
1' (CH_2)	4.14 (d, *J*=11.4 Hz), 4.24 (d, *J*=11.4 Hz)	64.3[a]
1'-O-	–	–
-1 (CO)	–	166.1[c]
-2 (CH_2)	3.48 (s)	41.7[a]
-3 (CO)	–	169.6[c]
2' (C)	–	103.19
3' (CH)	4.25[a] (d, *J*=8.7 Hz)	77.1[a]
4' (CH)	4.26[a] (br)	73.1[a]
5' (CH)	3.77[a] (br)	81.1[a]
6' (CH_2)	3.75 (d, *J*=13.0 Hz), 3.85[a] (d, *J*=13.0 Hz)	60.6[a]

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

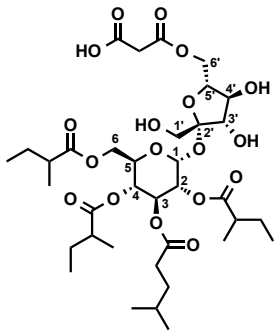
	**S(m) 5:24 [1] (m,5,5,5,6)** **Purified from *P. exserta*** **PubChem CID:** 91754411 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3358 **Material recovered:** 0.2~0.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.77 (d, *J*=4.0 Hz)	89.1[a]
2 (CH)	4.86 (dd, *J*=4.0, 9.5 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	176.0[c]
-2 (CH)	2.40[d] (m)	40.6[a]
-2' (CH_3)	1.13 (d, *J*=7.3 Hz)	16.1[a]
-3 (CH_2)	1.47 (m), 1.63[c,d] (m)	26.4[a]
-4 (CH_3)	0.88[c,d] (m)	11.2[c]
3 (CH)	5.57 (t, *J*=9.5 Hz)	69.7[a]
3-O-	–	–
-1 (CO)	–	173.4[c]
-2 (CH_2)	2.23[d] (m)	34.1[a]
-3 (CH_2)	1.54[d] (m)	24.6[a]
-4 (CH)	1.30[b] (br)	28.5[c]
-5 (CH_3) x 2	0.90[c] (d, *J*=7.2Hz)	22.3[c]
4 (CH)	5.18 (t, *J*=9.5 Hz)	67.7[a]
4-O-	–	–
-1 (CO)	–	175.3[c]
-2 (CH)	2.36[d] (m)	40.9[a]
-2' (CH_3)	1.10 (d, *J*=7.3 Hz)	16.5[a]
-3 (CH_2)	1.44 (m), 1.67[a,d] (m)	26.5[a]
-4 (CH_3)	0.89[c,d] (m)	11.3[c]
5 (CH)	4.38[a,d] (m)	68.3[a]
6 (CH_2)	4.13 (m), 4.25[a,d] (m)	61.6[a]
6-O-	–	–

Table S2. (cont'd)

-1 (CO)	–	177.1[c]
-2 (CH)	2.47[d] (m)	40.8[a]
-2' (CH_3)	1.20 (d, *J*=7.3 Hz)	16.3[a]
-3 (CH_2)	1.51 (m), 1.73[c,d] (m)	26.5[a]
-4 (CH_3)	0.92[c,d] (m)	11.2[c]
1' (CH_2)	3.55 (d, *J*=10.8 Hz), 3.67 (d, *J*=10.8 Hz)	63.2[a]
2' (C)	–	104.8[c]
3' (CH)	4.28[a] (d, *J*=7.1 Hz)	77.0[a]
4' (CH)	4.19[a] (t, *J*=7.1 Hz)	75.1[a]
5' (CH)	3.98[a] (br.)	78.3[a]
6' (CH_2)	4.43[a,d] (m)	64.6[a]
6'-O-	–	–
-1 (CO)	–	166.8[c]
-2 (CH_2)	3.51 (s)	41.4[a]
-3 (CO)	–	169.9[c]

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

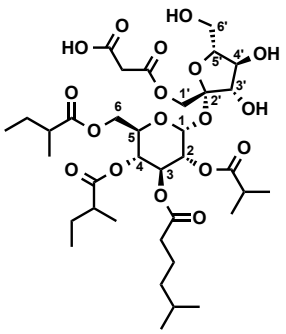
	**S(m) 5:24 [2] (m,4,5,5,7)** **Purified from *P. integrifolia*** **PubChem CID:** 91754412 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3198 **Material recovered:** 0.3~0.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.79 (d, *J*=3.9 Hz)	89.1[a]
2 (CH)	4.85 (dd, *J*=3.9, 10.1 Hz)	70.6[a]
2-O-	–	–
-1 (CO)	–	177.0[c]
-2 (CH)	2.58[c,d] (m)	33.8[a]
-3 (CH_3) x 2	1.20[c] (d, *J*=7.1 Hz)	18.7[a]
3 (CH)	5.56 (t, *J*=10.1 Hz)	69.2[a]
3-O-	–	–
-1 (CO)	–	173.1[c]
-2 (CH_2)	2.25[c] (t, *J*=7.9 Hz)	34.3[a]
-3 (CH_2)	1.56[c,d] (m)	22.5[a]
-4 (CH_2)	1.28[c] (br)	31.5[c]
-5 (CH_2)	1.32[c] (br)	29.6[c]
-6 (CH_3) x 2	0.89[c] (d, *J*=6.6 Hz)	22.3[a]
4 (CH)	5.15 (t, *J*=10.1 Hz)	67.5[a]
4-O-	–	–
-1 (CO)	–	175.2[c]
-2 (CH)	2.36[d] (m)	40.9[a]
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.5[a]
-3 (CH_2)	1.45 (m), 1.67[c,d] (m)	26.7[a]
-4 (CH_3)	0.93[c,d] (m)	11.3[c]
5 (CH)	4.34[d] (m)	68.9[a]
6 (CH_2)	4.15[a] (dd, *J*=7.1, 14.3 Hz)	60.4[a]
6-O-	–	–
-1 (CO)	–	176.5[c]

Table S2. (cont'd)

-2 (CH)	2.46[d] (m)	40.7[a]
-2' (CH_3)	1.14 (d, *J*=7.0 Hz)	16.3[a]
-3 (CH_2)	1.52 (m), 1.73[c,d] (m)	26.6[a]
-4 (CH_3)	0.88[c,d] (m)	11.2[c]
1' (CH_2)	4.05 (d, *J*=11.4 Hz), 4.34[a] (d, *J*=11.4 Hz)	63.2[a]
1'-O-	–	–
-1 (CO)	–	168.5[c]
-2 (CH_2)	3.51 (s)	43.9[a]
-3 (CO)	–	169.8[c]
2' (C)	–	104.1[c]
3' (CH)	4.23[a] (d, *J*=7.5 Hz)	78.5[a]
4' (CH)	4.27[a] (t, *J*=7.5 Hz)	73.8[a]
5' (CH)	3.78 (br.)	81.7[a]
6' (CH_2)	3.74[d] (m), 3.89 (d, *J*=3.5, 13.2 Hz)	60.4[a]

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

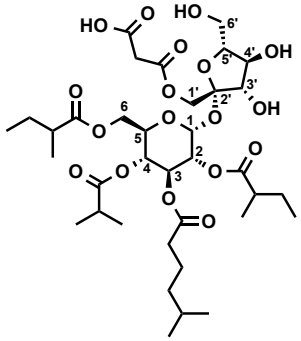
	**S(m) 5:24 [3] (m,4,5,5,7)** **Purified from *P. integrifolia*** **PubChem CID:** 91754413 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3181 **Material recovered:** 0.3~0.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.80 (d, *J*=3.6 Hz)	89.9[a]
2 (CH)	4.86 (dd, *J*=3.6, 10.0 Hz)	70.7[a]
2-O-	–	–
-1 (CO)	–	177.1[c]
-2 (CH)	2.43[d] (m)	40.5[a]
-2' (CH_3)	1.16 (d, *J*=6.9 Hz)	15.9[a]
-3 (CH_2)	1.47 (m), 1.63[c,d] (m)	26.7[a]
-4 (CH_3)	0.87[c,d] (m)	11.1[c]
3 (CH)	5.55 (t, *J*=10.0 Hz)	69.8[a]
3-O-	–	–
-1 (CO)	–	173.0[c]
-2 (CH_2)	2.22[c] (t, *J*=7.5 Hz)	34.2[a]
-3 (CH_2)	1.56[c,d] (m)	22.3[a]
-4 (CH_2)	1.26[c] (br)	31.5[c]
-5 (CH_2)	1.30[c] (br)	29.6[c]
-6 (CH_3) x 2	0.85[c] (d, *J*=6.6 Hz)	22.2[a]
4 (CH)	5.15 (t, *J*=10.0 Hz)	67.6[a]
4-O-	–	–
-1 (CO)	–	175.5[c]
-2 (CH)	2.54[c,d] (m)	33.8[a]
-3 (CH_3) x 2	1.14[c] (d, *J*=6.9 Hz)	18.9[a]
5 (CH)	4.34[a] (br.)	68.9[a]
6 (CH_2)	4.20[a,d] (m)	60.4[a]
6-O-	–	–
-1 (CO)	–	176.5[c]

Table S2. (cont'd)

-2 (CH)	2.47[d] (m)	40.9[a]
-2' (CH_3)	1.18 (d, *J*=6.9 Hz)	16.2[a]
-3 (CH_2)	1.52 (m), 1.73[c,d] (m)	26.4[a]
-4 (CH_3)	0.92[c,d] (m)	11.3[c]
1' (CH_2)	4.25 (d, *J*=12.0 Hz)	65.0[a]
1'-O-	–	–
-1 (CO)	–	166.4[c]
-2 (CH_2)	3.42 (s)	43.9[a]
-3 (CO)	–	168.7[c]
2' (C)	–	104.8[c]
3' (CH)	4.24[a] (d, *J*=8.4 Hz)	78.6[a]
4' (CH)	4.29[a,d] (m)	73.9[a]
5' (CH)	3.78[a,d] (m)	81.7[a]
6' (CH_2)	3.74 (dd, *J*=2.8, 12.3 Hz), 3.90 (dd, *J*=2.8, 12.3 Hz)	60.4[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

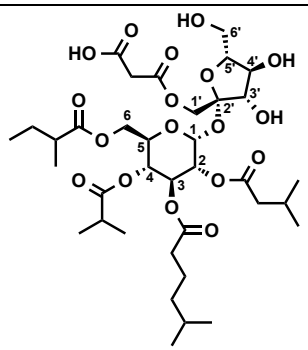
	**S(m) 5:24 [4] (m,4,5,5,7)** **Purified from *P. integrifolia*** **PubChem CID:** 91754414 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3184 **Material recovered:** 0.3~0.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.77 (d, *J*=3.7 Hz)	89.40
2 (CH)	4.85 (dd, *J*=3.7, 9.7 Hz)	70.82
2-O-	–	–
-1 (CO)	–	176.5[c]
-2 (CH_2)	2.43[c,d] (m)	40.0[a]
-3 (CH)	1.51[c,d] (m)	27.9[a]
-4 (CH_3) x 2	0.92[d] (m)	14.2[c]
3 (CH)	5.56 (t, *J*=9.7 Hz)	69.71
3-O-	–	–
-1 (CO)	–	173.1[c]
-2 (CH_2)	2.23[c,d] (m)	34.5[a]
-3 (CH_2)	1.54[c,d] (m)	22.3[a]
-4 (CH_2)	1.28[c] (br)	31.1[c]
-5 (CH)	1.31[c] (br)	29.4[c]
-6 (CH_3) x 2	0.89[c,d] (m)	22.1[c]
4 (CH)	5.15 (t, *J*=9.7 Hz)	67.93
4-O-	–	–
-1 (CO)	–	175.4[c]
-2 (CH)	2.54[c,d] (m)	34.0[a]
-3 (CH_3) x 2	1.15[c] (d, *J*=7.1 Hz)	18.5[c]
5 (CH)	4.37 (br.)	69.05
6 (CH_2)	4.13, 4.31[d] (m)	61.23
6-O-	–	–
-1 (CO)	–	176.9[c]
-2 (CH)	2.44[d] (m)	41.0[a]

Table S2. (cont'd)

-2' (CH_3)	1.20 (d, *J*=7.1 Hz)	16.5[a]
-3 (CH_2)	1.50, 1.72[c,d] (m)	26.8[a]
-4 (CH_3)	0.92[c,d] (m)	11.7[c]
1' (CH_2)	4.04 (d, *J*=11.4 Hz), 4.34 (d, *J*=11.4 Hz)	63.28
1'-O-	–	–
-1 (CO)	–	168.5[c]
-2 (CH_2)	3.42 (s)	44.3[a]
-3 (CO)	–	172.4[c]
2' (C)	–	103.04
3' (CH)	4.26[a,d] (m)	78.15
4' (CH)	4.27[a,d] (m)	73.81
5' (CH)	3.84 (br)	81.54
6' (CH_2)	3.75 (d, *J*=12.5 Hz), 3.89 (d, *J*=12.5 Hz)	60.76
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

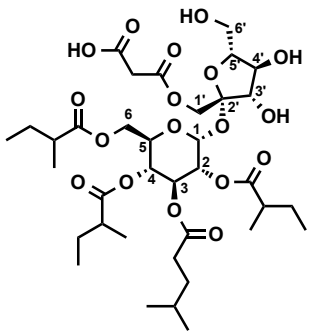
	**S(m) 5:24 [5] (m,5,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754415 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3134 **Material recovered:** 2~3 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.67 (d, *J*=3.9 Hz)	89.54
2 (CH)	4.84 (dd, *J*=3.9, 9.9 Hz)	70.44
2-O-	–	–
-1 (CO)	–	176.86
-2 (CH)	2.43[d] (m)	40.68
-2' (CH_3)	1.12 (d, *J*=7.0 Hz)	15.98
-3 (CH_2)	1.47 (m), 1.63[c,d] (m)	26.63
-4 (CH_3)	0.88[c,d] (m)	11.60
3 (CH)	5.52 (t, *J*=9.9 Hz)	69.58
3-O-	–	–
-1 (CO)	–	173.28
-2 (CH_2)	2.23[d] (m)	34.14
-3 (CH_2)	1.52[d] (m)	24.33
-4 (CH)	1.25 (br, *J*=6.6 Hz)	22.36
-5 (CH_3) x 2	0.93[c] (d, *J*=6.7 Hz)	22.30
4 (CH)	5.16 (t, *J*=9.9 Hz)	67.59
4-O-	–	–
-1 (CO)	–	175.16
-2 (CH)	2.34[d] (m)	40.95
-2' (CH_3)	1.07 (d, *J*=7.0 Hz)	16.56
-3 (CH_2)	1.45 (m), 1.61[a,d] (m)	26.80
-4 (CH_3)	0.89[c,d] (m)	11.60
5 (CH)	4.33[d] (m)	68.73
6 (CH_2)	4.03 (m), 4.27[d] (m)	61.32

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	177.06
-2 (CH)	2.43[d] (m)	40.68
-2' (CH_3)	1.14 (d, *J*=7.0 Hz)	16.42
-3 (CH_2)	1.45 (m), 1.69[c,d] (m)	26.48
-4 (CH_3)	0.92[c,d] (m)	11.42
1' (CH_2)	4.06 (d, *J*=11.4 Hz), 4.21 (d, *J*=11.4 Hz)	63.89
1'-O-	–	–
-1 (CO)	–	166.32
-2 (CH_2)	3.47 (s)	42.32
-3 (CO)	–	169.94
2' (C)	–	103.30
3' (CH)	4.23[a,d] (m)	76.85
4' (CH)	4.23[a,d] (m)	73.00
5' (CH)	3.78[d] (m)	81.3[a]
6' (CH_2)	3.72 (dd, *J*=2.8, 12.7 Hz), 3.84 (dd, *J*=2.8, 12.7 Hz)	60.89
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

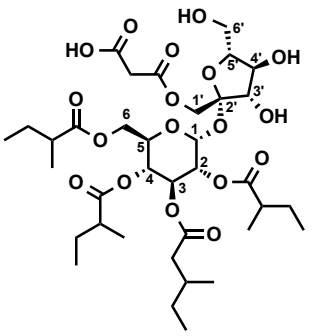
	**S(m) 5:24 [6] (m,5,5,5,6)** **Purified from *P. axillaris*** **PubChem CID:** 91754416 **HRMS:** (ESI) *m/z* calculated for $C_{36}H_{57}O_{18}^-$ ([M-H$^-$]): 777.3545, found: 777.3181 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.72 (d, *J*=3.7 Hz)	89.27
2 (CH)	4.84 (dd, *J*=3.7, 9.9 Hz)	70.77
2-O-	–	–
-1 (CO)	–	176.98
-2 (CH)	2.45[d] (m)	40.72
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.41
-3 (CH_2)	1.44 (m), 1.63[c,d] (m)	26.70
-4 (CH_3)	0.88[c,d] (m)	11.47
3 (CH)	5.54 (t, *J*=9.9 Hz)	69.47
3-O-	–	–
-1 (CO)	–	173.41
-2 (CH_2)	2.23 (d, *J*=7.1 Hz)	34.17
-3 (CH)	1.54[d] (m)	24.34
-3' (CH_3)	0.87[d] (m)	22.23
-4 (CH_2)	1.42[c,d] (m)	33.53
-5 (CH_3)	0.96[d] (m)	11.64
4 (CH)	5.17 (t, *J*=9.9 Hz)	67.64
4-O-	–	–
-1 (CO)	–	175.24
-2 (CH)	2.36[d] (m)	40.98
-2' (CH_3)	1.10 (d, *J*=7.0 Hz)	16.38
-3 (CH_2)	1.44 (m), 1.66[c,d] (m)	26.41
-4 (CH_3)	0.88[c,d] (m)	11.47
5 (CH)	4.35[d] (m)	68.79

Table S2. (cont'd)

6 (CH_2)	4.04 (m), 4.33[d] (m)	61.35
6-O-	–	–
-1 (CO)	–	176.92
-2 (CH)	2.45[d] (m)	40.72
-2' (CH_3)	1.16 (d, *J*=7.0 Hz)	16.41
-3 (CH_2)	1.49 (m), 1.72[c,d] (m)	26.54
-4 (CH_3)	0.92[c,d] (m)	11.64
1' (CH_2)	4.11 (d, *J*=11.4 Hz), 4.23 (d, *J*=11.4 Hz)	64.21
1'-O-	–	–
-1 (CO)	–	166.09
-2 (CH_2)	3.48 (s)	41.70
-3 (CO)	–	169.67
2' (C)	–	103.47
3' (CH)	4.26[a] (d, *J*=8.1 Hz)	77.02
4' (CH)	4.26[a] (t, *J*=8.1Hz)	73.02
5' (CH)	3.78[d] (m)	81.1[a]
6' (CH_2)	3.74 (dd, *J*=2.7, 12.5 Hz), 3.88 (dd, *J*=2.7, 12.5 Hz)	60.66
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

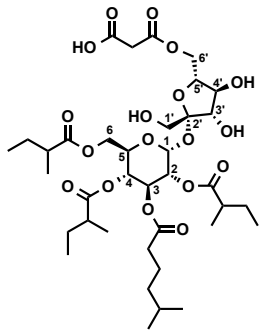
	**S(m) 5:25 [1] (m,5,5,5,7)** **Purified from *P. exserta*** **PubChem CID:** 91754417 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{59}O_{18}^-$ ([M-H$^-$]): 791.3701, found: 791.3377 **Material recovered:** 2~3 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.68 (d, *J*=3.9 Hz)	89.52
2 (CH)	4.91 (dd, *J*=3.9, 9.8 Hz)	70.42
2-O-	–	–
-1 (CO)	–	176.06
-2 (CH)	2.40[c,d] (m)	40.70
-2' (CH_3)	1.10 (d, *J*=7.1 Hz)	16.15
-3 (CH_2)	1.45 (m), 1.62[c,d] (m)	26.81
-4 (CH_3)	0.86[d] (m)	11.65
3 (CH)	5.54 (t, *J*=9.8 Hz)	69.93
3-O-	–	–
-1 (CO)	–	173.63
-2 (CH_2)	2.22 (t, *J*=7.4 Hz)	34.26
-3 (CH_2)	1.53[d] (m)	24.65
-4 (CH_2)	1.29[b] (br)	28.96
-5 (CH)	1.24[b] (br)	31.52
-6 (CH_3) x 2	0.88 (d, *J*=6.7 Hz)	22.5[a]
4 (CH)	5.18 (t, *J*=9.8 Hz)	67.81
4-O-	–	–
-1 (CO)	–	175.24
-2 (CH)	2.35[c,d] (m)	40.97
-2' (CH_3)	1.07 (d, *J*=7.1 Hz)	16.59
-3 (CH_2)	1.43 (m), 1.66[c,d] (m)	26.54
-4 (CH_3)	0.87[d] (m)	11.60
5 (CH)	4.36[d] (m)	68.46
6 (CH_2)	4.09 (m), 4.34[d] (m)	61.64

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	177.07
-2 (CH)	2.46[c,d] (m)	41.00
-2' (CH_3)	1.14 (d, *J*=7.1 Hz)	16.37
-3 (CH_2)	1.49 (m), 1.72[c,d] (m)	26.61
-4 (CH_3)	0.90[d] (m)	11.50
1' (CH_2)	3.54 (d, *J*=12.3 Hz), 3.68[a] (d, *J*=12.3 Hz)	63.11
2' (C)	–	104.65
3' (CH)	4.31[a] (d, *J*=8.6 Hz)	76.85
4' (CH)	4.18[a] (t, *J*=8.6 Hz)	75.27
5' (CH)	4.03[d] (m)	78.43
6' (CH_2)	4.40 (m), 4.47[a,d] (m)	65.02
6'-O-	–	–
-1 (CO)	–	167.61
-2 (CH_2)	3.49 (s)	41.40
-3 (CO)	–	170.35
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

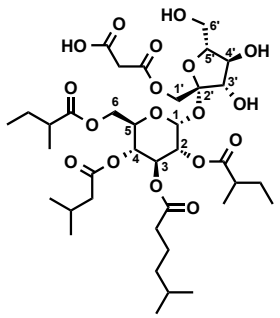
	**S(m) 5:25 [2] (m,5,5,5,7)** **Purified from *P. integrifolia*** **PubChem CID:** 91754418 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{59}O_{18}^-$ ([M-H$^-$]): 791.3701, found: 791.3093 **Material recovered:** 0.3~0.5 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.80 (d, *J*=3.7 Hz)	88.98
2 (CH)	4.84 (dd, *J*=3.7, 9.5 Hz)	70.67
2-O-	–	–
-1 (CO)	–	177.0[c]
-2 (CH)	2.43[c,d] (m)	40.5[a]
-2' (CH_3)	1.16[c] (d, *J*=7.1 Hz)	15.9[a]
-3 (CH_2)	1.49 (m), 1.69[c,d](m)	26.4[a]
-4 (CH_3)	0.86[a,d] (m)	11.3[a]
3 (CH)	5.57 (t, *J*=9.5 Hz)	69.09
3-O-	–	–
-1 (CO)	–	172.9[C]
-2 (CH_2)	2.24[c] (t, *J*=7.0 Hz)	34.2[a]
-3 (CH_2)	1.55[c,d] (m)	22.6[a]
-4 (CH_2)	1.31[b] (br, *J*=7.0 Hz)	29.5[c]
-5 (CH)	1.28[b] (br, *J*=7.0 Hz))	31.1[c]
-6 (CH_3) x 2	0.86 (d, *J*=7.0 Hz)	22.4[a]
4 (CH)	5.15 (t, *J*=9.5 Hz)	67.68
4-O-	–	–
-1 (CO)	–	171.4[c]
-2 (CH_2)	2.20 [d](m)	43.0[a]
-3 (CH)	2.08[c] (br)	25.3[a]
-4 (CH_3) x 2	0.95[c] (d, *J*=6.6 Hz)	14.2[c]
5 (CH)	4.34 (br)	68.91
6 (CH_2)	4.15[d] (m)	60.44
6-O-	–	–
-1 (CO)	–	176.6[c]

Table S2. (cont'd)

-2 (CH)	2.48[c,d] (m)	40.7[a]
-2' (CH_3)	1.18[c] (d, *J*=7.1 Hz)	16.2[a]
-3 (CH_2)	1.51 (m), 1.73[c,d] (m)	26.5[a]
-4 (CH_3)	0.92[a,d] (m)	11.4[a]
1' (CH_2)	4.10[a] (d, *J*=12.1 Hz)	61.34
1'-O-	–	–
-1 (CO)	–	165.6[C]
-2 (CH_2)	3.51 (s)	50.94
-3 (CO)	–	166.9[C]
2' (C)	–	104.90
3' (CH)	4.25[a] (d, *J*=8.1 Hz)	78.51
4' (CH)	4.22[a,d] (m)	73.84
5' (CH)	3.79[a,d] (m)	71.76
6' (CH_2)	3.74 (m), 3.89[a,d] (m)	60.44
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

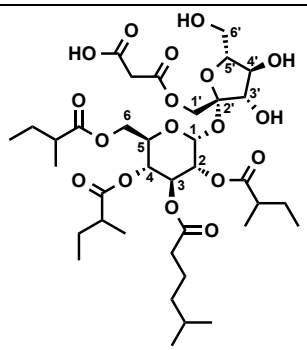
	**S(m) 5:25 [3] (m,5,5,5,7)** **Purified from *P. axillaris*** **PubChem CID:** 91754419 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{59}O_{18}^-$ ([M-H$^-$]): 791.3701, found: 791.3159 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.70 (d, *J*=4.0 Hz)	89.56
2 (CH)	4.84 (dd, *J*=4.0, 9.9 Hz)	70.90
2-O-	–	–
-1 (CO)	–	176.84
-2 (CH)	2.45[c,d] (m)	40.53
-2' (CH_3)	1.13 (d, *J*=7.1 Hz)	15.90
-3 (CH_2)	1.48 (m), 1.63[c,d] (m)	26.69
-4 (CH_3)	0.86[a,d] (m)	11.48
3 (CH)	5.54 (t, *J*=9.9 Hz)	69.84
3-O-	–	–
-1 (CO)	–	173.13
-2 (CH_2)	2.23 (t, *J*=7.0 Hz)	34.05
-3 (CH_2)	1.53[d] (m)	24.55
-4 (CH_2)	1.25[b] (br)	28.80
-5 (CH)	1.22[c] (br)	31.40
-6 (CH_3) x 2	0.88 (d, *J*=6.7 Hz)	22.42
4 (CH)	5.16 (t, *J*=9.9 Hz)	67.95
4-O-	–	–
-1 (CO)	–	175.07
-2 (CH)	2.34[c,d] (m)	40.91
-2' (CH_3)	1.08 (d, *J*=7.1 Hz)	16.45
-3 (CH_2)	1.44 (m), 1.66[c,d] (m)	26.38
-4 (CH_3)	0.90[a,d] (m)	11.32
5 (CH)	4.32[a,d] (m)	69.13
6 (CH_2)	4.10 (m), 4.31[a,d] (m)	61.58

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	176.84
-2 (CH)	2.45[c,d] (m)	40.53
-2' (CH_3)	1.15 (d, *J*=7.1 Hz)	16.30
-3 (CH_2)	1.48 (m), 1.71[c,d] (m)	26.53
-4 (CH_3)	0.87[a,d] (m)	11.50
1' (CH_2)	4.10 (d, *J*=11.5 Hz), 4.25[a] (d, *J*=11.5 Hz)	64.41
1'-O-	–	–
-1 (CO)	–	166.16
-2 (CH_2)	3.47 (s)	41.58
-3 (CO)	–	169.77
2' (C)	–	103.19
3' (CH)	4.28[a,d] (m)	77.28
4' (CH)	4.26[a,d] (m)	73.38
5' (CH)	3.77[d] (m)	81.14
6' (CH_2)	3.75 (m), 3.87[d] (m)	61.11
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

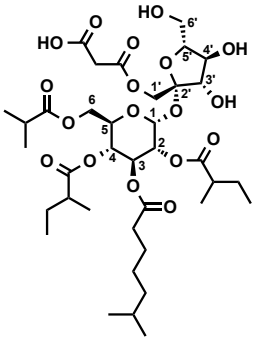
	**S(m) 5:25 [4] (m,4,5,5,8)** **Purified from *P. axillaris*** **PubChem CID:** 91754420 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{59}O_{18}^-$ ([M-H$^-$]): 791.3701, found: 791.3210 **Material recovered:** 0.5~1 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.74 (d, *J*=3.7 Hz)	89.2[a]
2 (CH)	4.85 (dd, *J*=3.7, 9.9 Hz)	70.6[a]
2-O-	–	–
-1 (CO)	–	177.0[c]
-2 (CH)	2.44[c,d] (m)	40.8[a]
-2' (CH_3)	1.17 (d, *J*=7.0 Hz)	15.9[c]
-3 (CH_2)	1.49, 1.64[c,d] (m)	26.7[c]
-4 (CH_3)	0.88[c,d] (m)	11.4[a]
3 (CH)	5.55 (t, *J*=9.9 Hz)	69.4[a]
3-O-	–	–
-1 (CO)	–	173.2[c]
-2 (CH_2)	2.23 (t, *J*=7.6 Hz)	34.1[a]
-3 (CH_2)	1.55[d] (m)	24.7[c]
-4 (CH_2)	1.28 (br)[a]	29.1[c]
-5 (CH_2)	1.26 (br)[a]	31.6[c]
-6 (CH)	1.30 (br)[a]	22.7[c]
-7 (CH_3) x 2	0.89[c] (d, *J*=6.9 Hz)	22.3[a]
4 (CH)	5.15 (t, *J*=9.9 Hz)	67.76
4-O-	–	–
-1 (CO)	–	175.2[c]
-2 (CH)	2.36[c,d] (m)	41.0[a]
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.4[c]
-3 (CH_2)	1.45 (m), 1.67[c,d] (m)	26.4[c]
-4 (CH_3)	0.92[c,d] (m)	11.6[c]
5 (CH)	4.34[a,d] (m)	68.9[a]

Table S2. (cont'd)

6 (CH_2)	4.11 (m), 4.28[a, d] (m)	61.3[a]
6-O-	–	–
-1 (CO)	–	173.9[c]
-2 (CH)	2.37[c,d] (m)	33.9[a]
-3 (CH_3) x 2	1.15[c] (d, *J*=6.9 Hz)	19.0[a]
1' (CH_2)	4.15 (d, *J*=11.4 Hz), 4.22[a] (d, *J*=11.4 Hz)	64.4[a]
1'-O-	–	–
-1 (CO)	–	165.9[c]
-2 (CH_2)	3.47 (s)	41.5[a]
-3 (CO)	–	169.4[c]
2' (C)	–	103.6[c]
3' (CH)	4.26[a] (br, *J*=8.9 Hz)	77.3[a]
4' (CH)	4.27[a] (br, *J*=8.9 Hz)	72.9[a]
5' (CH)	3.77[d] (m)	81.0[a]
6' (CH_2)	3.74 (m), 3.87[d] (m)	60.4[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

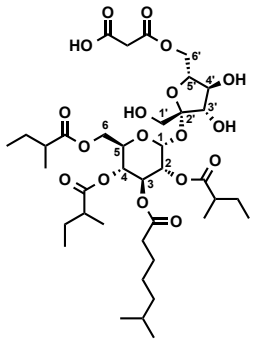
	**S5:26 [1] (m,5,5,5,8)** **Purified from *P. exserta*** **PubChem CID:** 91754421 **HRMS:** (ESI) *m/z* calculated for $C_{38}H_{61}O_{18}^-$ ([M-H$^-$]):805.3858, found: 805.3468 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.69 (d, *J*=3.7 Hz)	89.44
2 (CH)	4.93 (dd, *J*=3.7, 9.8 Hz)	70.32
2-O-	–	–
-1 (CO)	–	176.07
-2 (CH)	2.39[d] (m)	40.65
-2' (CH_3)	1.11 (d, *J*=6.8Hz)	16.01
-3 (CH_2)	1.45 (m), 1.61[c,d] (m)	26.67
-4 (CH_3)	0.85[d] (m)	11.38
3 (CH)	5.57 (t, *J*=9.8 Hz)	69.94
3-O-	–	–
-1 (CO)	–	173.62
-2 (CH)	2.22 (t, *J*=7.4 Hz)	34.22
-3 (CH_2)	1.50[d] (m)	24.78
-4 (CH_2)	1.26 (br)[a]	29.12
-5 (CH_2)	1.24 (br)[a]	28.85
-6 (CH)	1.29 (br)[a]	31.58
-7 (CH_3) x 2	0.95 (d, *J*=6.7 Hz)	22.50
4 (CH)	5.18 (t, *J*=9.8 Hz)	67.93
4-O-	–	–
-1 (CO)	–	175.20
-2 (CH)	2.34[c,d] (m)	40.99
-2' (CH_3)	1.08 (d, *J*=6.8 Hz)	16.43
-3 (CH_2)	1.42 (m), 1.65[c,d] (m)	26.41
-4 (CH_3)	0.86[d] (m)	14.23
5 (CH)	4.39[a] (br.m)	68.43

Table S2. (cont'd)

6 (CH_2)	4.07 (m), 4.32[a,d] (m)	61.64
6-O-	–	–
-1 (CO)	–	177.05
-2 (CH)	2.44[c,d] (m)	40.92
-2' (CH_3)	1.15 (d, *J*=6.8 Hz)	16.23
-3 (CH_2)	1.48 (m), 1.71[c,d] (m)	26.67
-4 (CH_3)	0.87[d] (m)	11.47
1' (CH_2)	3.52 (d, *J*=12.1Hz), 3.66[a] (d, *J*=12.1Hz)	63.03
2' (C)	–	104.63
3' (CH)	4.30 (d, *J*=8.5 Hz)	76.73
4' (CH)	4.17[a] (t, *J*=8.5 Hz)	75.48
5' (CH)	4.03 (br)	78.49
6' (CH_2)	4.38 (dd, *J*=4.2, 12.1 Hz), 4.47 (dd, *J*=4.2, 12.1 Hz)	65.04
6'-O-	–	–
-1 (CO)	–	167.18
-2 (CH_2)	3.52 (d, *J*=12.6 Hz)	41.71
-3 (CO)	–	169.93

[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.

Table S2. (cont'd)

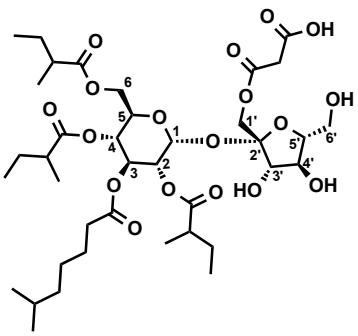
	**S5:26 [2] (m,5,5,5,8)** **Purified from *P. exserta*** **PubChem CID:** 91754422 **HRMS:** (ESI) *m/z* calculated for $C_{38}H_{61}O_{18}^-$ ([M-H$^-$]):805.3858, found: 805.3323 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.71 (d, *J*=3.8 Hz)	89.35
2 (CH)	4.85 (dd, *J*=3.8, 9.8 Hz)	70.64
2-O-	–	–
-1 (CO)	–	176.97
-2 (CH)	2.44[d] (m)	40.67
-2' (CH_3)	1.12 (d, *J*=6.8Hz)	16.3[b]
-3 (CH_2)	1.47 (m), 1.62[c] (m)	26.66
-4 (CH_3)	0.86[d] (m)	11.65
3 (CH)	5.56 (t, *J*=9.8 Hz)	69.72
3-O-	–	–
-1 (CO)	–	173.28
-2 (CH)	2.22[b] (t, *J*=7.6 Hz)	34.06
-3 (CH_2)	1.52[d] (m)	24.86
-4 (CH_2)	1.26 (br)[a]	29.22
-5 (CH_2)	1.28 (br)[a]	28.99
-6 (CH)	1.26 (br)[a]	31.73
-7 (CH_3) x 2	0.94 (d, *J*=6.8 Hz)	22.66
4 (CH)	5.19 (t, *J*=9.8 Hz)	67.74
4-O-	–	–
-1 (CO)	–	175.16
-2 (CH)	2.34[c,d] (m)	41.00
-2' (CH_3)	1.07 (d, *J*=6.8 Hz)	16.58
-3 (CH_2)	1.43 (m), 1.65[c,d] (m)	26.51
-4 (CH_3)	0.90[d] (m)	11.46

Table S2. (cont'd)

5 (CH)	4.37[a] (br)	68.67
6 (CH_2)	4.06 (m), 4.34[a,d] (m)	61.35
6-O-	–	–
-1 (CO)	–	176.99
-2 (CH)	2.44[c,d] (m)	40.67
-2' (CH_3)	1.14 (d, *J*=6.8 Hz)	16.33
-3 (CH_2)	1.47 (m), 1.71[c,d] (m)	26.83
-4 (CH_3)	0.87[d] (m)	11.61
1' (CH_2)	4.10 (d, *J*=11.3Hz), 4.25[a] (d, *J*=11.3Hz)	64.02
1'-O-	–	–
-1 (CO)	–	166.16
-2 (CH_2)	3.48 (s)	41.79
-3 (CO)	–	169.82
2' (C)	–	103.11
3' (CH)	4.26 (d, *J*=6.0 Hz)	76.92
4' (CH)	4.25[a] (t, *J*=6.0. Hz)	73.20
5' (CH)	3.77 (br)	81.10
6' (CH_2)	3.75 (dd, *J*=3.3, 12.8 Hz), 3.87 (dd, *J*=3.3, 12.8 Hz)	60.65
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

Table S2. (cont'd)

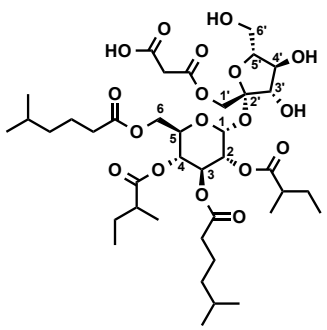
	**S(m) 5:27 (m,5,5,7,7)** **Purified from *P. exserta*** **PubChem CID:** 91754423 **HRMS:** (ESI) *m/z* calculated for $C_{37}H_{59}O_{18}^-$ ([M-H$^-$]): 819.4014, found: 819.3073 **Material recovered:** 0.2~0.4 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.75 (d, *J*=3.7 Hz)	89.0[a]
2 (CH)	4.91 (dd, *J*=3.7, 9.8 Hz)	70.9[a]
2-O-	–	–
-1 (CO)	–	176.0[c]
-2 (CH)	2.42[c,d] (m)	40.6[a]
-2' (CH_3)	1.15 (d, *J*=7.0 Hz)	16.1[a]
-3 (CH_2)	1.48 (m), 1.63[c,d] (m)	26.7[c]
-4 (CH_3)	0.93[c,d] (m)	11.2[c]
3 (CH)	5.58 (t, *J*=9.8 Hz)	69.6[a]
3-O-	–	–
-1 (CO)	–	172.7[c]
-2 (CH_2)	2.24[c,d] (m)	34.09
-3 (CH_2)	1.56[c,d] (m)	24.7[a]
-4 (CH_2)	1.29[b] (br)	28.6[c]
-5 (CH)	1.26[c] (br)	31.2[c]
-6 (CH_3) x 2	0.93 (d, *J*=7.2 Hz)	22.4[a]
4 (CH)	5.19 (t, *J*=9.8 Hz)	67.7[a]
4-O-	–	–
-1 (CO)	–	175.2[c]
-2 (CH)	2.36[c,d] (m)	41.0[a]
-2' (CH_3)	1.11 (d, *J*=7.0 Hz)	16.4[a]
-3 (CH_2)	1.45 (m), 1.67[c,d] (m)	26.5[c]
-4 (CH_3)	0.89[c,d] (m)	11.4[c]
5 (CH)	4.32[a,d] (m)	69.1[a]
6 (CH_2)	4.12 (m), 4.28[a,d] (m)	61.3[a]

Table S2. (cont'd)

6-O-	–	–
-1 (CO)	–	173.5[c]
-2 (CH_2)	2.23[c,d] (m)	34.09
-3 (CH_2)	1.53[c,d] (m)	24.7[a]
-4 (CH_2)	1.29[b] (br)	28.9[c]
-5 (CH)	1.26[c] (br)	31.3[c]
-6 (CH_3) x 2	0.89[c] (d, *J*=7.2 Hz)	22.5[a]
1' (CH_2)	4.09 (d, *J*=12.0 Hz), 4.43[a] (d, *J*=12.0 Hz)	64.3[a]
1'-O-	–	–
-1 (CO)	–	166.5[c]
-2 (CH_2)	3.51 (s)	41.8[a]
-3 (CO)	–	169.4[c]
2' (C)	–	104.8[c]
3' (CH)	4.29[a,d] (m)	77.0[a]
4' (CH)	4.29[a,d] (m)	73.7[a]
5' (CH)	3.79[a,d] (m)	80.9[a]
6' (CH_2)	3.75 (m), 3.89[a,d] (m)	60.4[a]
[a]Determined by HSQC. [b]Determined by COSY. [c]Determined by HMBC. [d]Multiplicity not certain due to overlap.		

REFERENCES

REFERENCES

1. Duke, S. O.; Canel, C.; Rimando, A. M.; Telle, M. R.; Duke, M. V.; Paul, R. N., Current and potential exploitation of plant glandular trichome productivity. *Advances in Botanical Research* **2000**, *31*, 121-151.

2. Wagner, G. J., Secreting glandular trichomes: more than just hairs. *Plant Physiology* **1991**, *96* (3), 675-679.

3. Balandrin, M. F.; Klocke, J. A.; Wurtele, E. S.; Bollinger, W. H., Natural plant chemicals: sources of industrial and medicinal materials. *Science* **1985**, *228* (4704), 1154-1159.

4. Cragg, G. M.; Newman, D. J., Plants as a source of anti-cancer agents. *Journal of Ethnopharmacology* **2005**, *100* (1), 72-79.

5. DellaPenna, D.; O'Connor, S. E., Plant gene clusters and opiates. *Science* **2012**, *336* (6089), 1648-1649.

6. Fahn, A., Structure and function of secretory cells. *Advances in Botanical Research* **2000**, *31*, 37-75.

7. Schilmiller, A. L.; Last, R. L.; Pichersky, E., Harnessing plant trichome biochemistry for the production of useful compounds. *The Plant Journal* **2008**, *54* (4), 702-711.

8. Glas, J. J.; Schimmel, B. C. J.; Alba, J. M.; Escobar-Bravo, R.; Schuurink, R. C.; Kant, M. R., Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences* **2012**, *13* (12), 17077-17103.

9. McKenzie, C. L.; Puterka, G. J., Effect of sucrose octanoate on survival of nymphal and adult *Diaphorina citri* (Homoptera: Psyllidae). *Journal of economic entomology* **2004**, *97* (3), 970-975.

10. Rodriguez, A. E.; Tingey, W. M.; Mutschler, M. A., Acylsugars of *Lycopersicon pennellii* deter settling and feeding of the green peach aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **1993**, *86* (1), 34-39.

11. Weinhold, A.; Baldwin, I. T., Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. *Proceedings of the National Academy of Sciences* **2011**, *108* (19), 7855-7859.

12. Ghosh, B.; Westbrook, T. C.; Jones, A. D., Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014**, *10* (3), 496-507.

13. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Sucrose esters associated with glandular trichomes of wild Lycopersicon species. *Phytochemistry* **1990**, *29* (7), 2115-2118.

14. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Characterization of 2, 3, 4, 3'-tetra-O-acylated sucrose esters associated with the glandular trichomes of *Lycopersicon typicum*. *Journal of Agricultural and Food Chemistry* **1993**, *41* (3), 469-473.

15. Chortyk, O. T.; Kays, S. J.; Teng, Q., Characterization of insecticidal sugar esters of Petunia. *Journal of Agricultural and Food Chemistry* **1997**, *45* (1), 270-275.

16. Maluf, W. R.; Maciel, G. M.; Gomes, L. A. A.; Cardoso, M. d. G.; Gonçalves, L. D.; da Silva, E. C.; Knapp, M., Broad-spectrum arthropod resistance in hybrids between high- and low-acylsugar tomato lines. *Crop Science* **2010**, *50* (2), 439-450.

17. Kim, J.; Kang, K.; Vigil, E. G.; Shi, F.; Jones, A. D.; Barry, C. S.; Last, R. L., Striking natural diversity in glandular trichome acylsugar composition is shaped by variation at the Acyltransferase 2 locus in the wild tomato *Solanum habrochaites*. *Plant Physiology* **2012**, *160* (4). 1854-1870.

18. Leckie, B. M.; Halitschke, R.; De Jong, D. M.; Smeda, J. R.; Kessler, A.; Mutschler, M. A., Quantitative trait loci regulating the fatty acid profile of acylsugars in tomato. *Molecular Breeding* **2014**, *34* (3), 1201-1213.

19. Ghosh, B.; Jones, A. D., Profiling, characterization, and analysis of natural and synthetic acylsugars (sugar esters). *Analytical Methods* **2017**, *9*: 892-905.

20. Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Daniel Jones, A.; Last, R. L., Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. *The Plant Journal* **2010**, *62* (3), 391-403.

21. Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry- based metabolomics. *Mass Spectrometry Reviews* **2007**, *26* (1), 51-78.

22. Ekanayaka, E. A. P.; Celiz, M. D.; Jones, A. D., Relative mass defect filtering of mass spectra: a path to discovery of plant specialized metabolites. *Plant Physiology* **2015**, *167* (4), 1221-1232.

23. Kliebenstein, D. J.; Osbourn, A., Making new molecules–evolution of pathways for novel metabolites in plants. *Current Opinion in Plant Biology* **2012**, *15* (4), 415-423.

24. Schilmiller, A. L.; Charbonneau, A. L.; Last, R. L., Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proceedings of the National Academy of Sciences of the USA* **2012**, *109* (40), 16377-16382.

25. Fan, P.; Miller, A. M.; Schilmiller, A. L.; Liu, X.; Ofner, I.; Jones, A. D.; Zamir, D.; Last, R. L., In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. *Proceedings of the National Academy of Sciences of the USA* **2016**, *113* (2), E239-E248.

26. Moser, D.; Klaiber, I.; Vogler, B.; Kraus, W., Molluscicidal and antibacterial compounds from *Petunia hybrida*. *Pest Management Science* **1999**, *55* (3), 336-339.

27. Ohya, I.; Shinozaki, Y.; Tobita, T.; Takahashi, H.; Matsuzaki, T., Sucrose esters from the surface lipids of *Petunia hybrida*. *Phytochemistry* **1996**, *41* (3), 787-789.

28. Kroumova, A. B. M.; Zaitlin, D.; Wagner, G. J., Natural variability in acyl moieties of sugar esters produced by certain tobacco and other *Solanaceae* species. *Phytochemistry* **2016**, *130*, 218-227.

29. Singh, A. P.; Singh, A. K.; Begum, A. S.; Sahai, M., Two acyl sucroses from *Petunia nyctaginiflora*. *Phytochemistry* **2003**, *63* (4), 485-489.

30. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L., Proposed minimum reporting standards for chemical analysis. *Metabolomics* **2007**, *3* (3), 211-221.

31. Kraus, W., Investigation of biologically active natural products using online LC-bioassay, LC- NMR, and LC- MS techniques. *Journal of Toxicology: Toxin Reviews* **2003**, *22* (4), 495-508.

32. Kulcheski, F. R.; Muschner, V. C.; Lorenz-Lemke, A. P.; Stehmann, J. R.; Bonatto, S. L.; Salzano, F. M.; Freitas, L. B., Molecular phylogenetic analysis of Petunia juss.(*Solanaceae*). *Genetica* **2006**, *126* (1), 3-14.

33. Kroumova, A. B.; Wagner, G. J., Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants. *Planta* **2003**, *216* (6), 1013-1021.

34. Forkner, R. E.; Hare, J. D., Genetic and environmental variation in acyl glucose ester production and glandular and nonglandular trichome densities in *Datura wrightii*. *Journal of Chemical Ecology* **2000**, *26* (12), 2801-2823.

Chapter Three.

Distinguishing isomeric acylsugar metabolites in the *Solanaceae* by *in vivo* labeling using [^{13}C]-labeled amino acid precursors and analysis using liquid chromatography-tandem mass spectrometry

3.1. Introduction

It has been estimated that the plant kingdom synthesizes about 200,000 specialized metabolites that play important roles in adaptation to ecological challenges from herbivores and dynamic environments [1-4]. Many plant-derived natural products and synthetic compounds they inspire find commercial value as pharmaceuticals, fragrances, food additives and pesticides [5,6]. More than 50% of approved clinical drugs over the past 30 years have natural product origins [7]. Compared with chemical synthesis, plant metabolic biosynthesis and engineering provide timesaving and environmentally friendly ways to generate bioactive chemicals [8]. Improvements in these processes have been limited by gaps in our knowledge regarding the identities of specialized metabolites and discoveries of gene functions involved in regulation of their biosynthesis.

Identification of plant specialized metabolites first requires recognition that they exist, and this is achieved through untargeted "discovery" metabolomic measurements performed on plant extracts. Such metabolomic investigations are expected bridge the gap between phenotype and genotype, and lead to discoveries of relevant genes and enzymes involved in biosynthesis of specific metabolites and the factors that regulate their abundance [9]. However, comprehensive identification and structural characterization of specialized metabolites from plant tissues remains one of the grand challenges in phytochemistry [10]. It has been estimated that around 15,000 metabolites may be present in a single plant species [11]. Chemical diversity of metabolites is much greater than that of nucleotides and proteins, which are produced from a limited range of building blocks. The diversity of chemical substructures in specialized metabolites has been ascribed to result from gene duplication, mutation and neofunctionalization of a redundant gene, and positive selection for metabolic phenotypes within and among species

[12]. Nevertheless, compared with progress achieved in genomics and proteomics, more efforts need to be devoted to increase structural and functional characterization of plant metabolomes [13], as entire genomes can be sequenced faster and at lower cost than elucidation of the structure of a single novel specialized metabolite. To overcome these challenges, wide ranges of powerful analytical technologies are applied to resolve and identify the huge structural diversity present in plant metabolomes.

Nuclear Magnetic Resonance (NMR) spectroscopy and Gas Chromatography or High Performance Liquid Chromatography coupled to Mass Spectrometry (GC-MS, HPLC-MS) have emerged as popular analytical tools for establishing structural diversity of specialized metabolites from plant tissues [10, 14-16]. GC-MS is a traditional analytical technique for the analysis of nonpolar, volatile and thermostable plant metabolites owing to its reproducibility of retention index and availability of mass spectrum databases [17,18]. These features make GC-MS a reliable approach for targeted metabolic analysis of known chemicals, though the lack of a truly comprehensive spectrum database limits its application for identification of novel compounds. NMR provides a robust, non-destructive and reproducible technique that yields atomic connectivities needed to establish structures of unknown compounds pre-purified from biological extracts [19]. Previous investigation of tomato specialized metabolites from the Jones lab demonstrates that the combination of 1D and 2D NMR spectroscopic analysis provides a powerful tool to determine structures and reveal structural diversity of a class of metabolites across genotypes [20]. With the improvement of instrument sensitivity and data processing methods, NMR has found success in identification of novel metabolites from mixtures [14,21,22]. However, in order to obtain good quality spectra, at least a few hundred micrograms of purified sample has been required for NMR analysis, which becomes a drawback of this

technique. HPLC-MS is one of the most commonly used approaches for metabolite analysis owing to its capability to detect a wide range of compounds, diverse in concentration and structure, with high sensitivity and resolution [15,16]. Metabolite identification is performed based on distinct retention times, accurate molecular masses, and relative mass defects of measured ions, whereas detailed structural information could be gathered from fragmentation patterns generated using multiplexed collision induced dissociation (CID) or multistage MS^n analysis [23,24]. Nevertheless, one limitation of HPLC-MS in plant metabolite profiling is that it cannot provide full structural characterization of unknown metabolites because fragment ion masses often fail to distinguish isomers. Therefore, chemical diversity arising from structural isomers differing in connectivity but with the same molecular and fragment ion masses may not be resolved by HPLC-MS.

Nowadays, plant metabolic analyses benefit from stable isotope-enabled methodology to aid metabolite identification, quantification, metabolic biosynthetic pathway determination, as well as measurements of metabolite turnover and flux [25,26]. Different from radiolabeling, stable isotope labeled compounds are widely analyzed using mass spectrometry, which can distinguish labeled substances by shift of masses from the naturally abundant isotopologs. A variety of stable isotope labeling approaches have been reported, performed using uniformly labeled isotopic tracers including [$^{13}C_6$]glucose, [$^{13}C_3$]lactate and [$^{13}C_5$]glutamine, demonstrated the fate of target precursor metabolites, along with illustration of correlated metabolic pathways [27,28]. Additionally, stable isotope labeled standards generated from ^{13}C-enriched precursors was used to improve quantitative analysis of all volatile metabolites of interest [29]. Identification of bioactive metabolites as well as central intermediates represents another application of stable isotope assisted metabolomics [28,30]. Though great efforts have been devoted to use of stable

isotope tracers to establish biosynthetic pathways, to our knowledge, their application of chemical diversity analysis, particularly for distinguishing isomeric specialized metabolites that could not be distinguished by mass spectra, is at an early stage.

Isomeric metabolites differing by branching in aliphatic substructures present unusual challenges for mass spectrometric identification because ions often fail to undergo fragmentation that yields ion masses that distinguishing branching isomers. To address this, UHPLC-MS based stable isotope labeling strategy for rapid distinguishing of isomeric plant metabolites has been reported here, using uniformly ^{13}C-Branched Chain Amino Acids (BCAA) valine (Val), leucine (Leu) and isoleucine (Ile) as isotopic tracers. Each is proposed as a biosynthetic precursor of branched aliphatic ester groups in tomato specialized metabolites. Acylated sugar metabolites (named as "acylsugar") in six plant species in the family *Solanaceae* were selected as targets and labeled *in vivo* with ^{13}C-BCAA precursors.

3.2. Materials and Methods

3.2.1 Chemicals

L-valine ($^{13}C_5$, 99%; ^{15}N, 99%), L-leucine ($^{13}C_6$, 99%; ^{15}N, 99%) and L-isoleucine ($^{13}C_6$, 99%; ^{15}N, 99%) were purchased from Cambridge Isotope Laboratories. L-valine, L-leucine, L-isoleucine and Bacto™ Agar were purchased from Sigma-Aldrich. HPLC grade acetonitrile, 2-propanol, methanol, formic acid (88% aqueous solution), ammonium formate (14.8 M, aqueous solution) were purchased from VMR Scientific. Half-strength Hoagland nutrient solution was obtained from Michigan State University Growth Chamber Facility.

3.2.2 ^{13}C amino acid feeding of plants

Tomato seeds (*Solanum lycopersicum* M82, *Solanum pennellii* accessions LA0716 and LA2560) were obtained from the UC Davis C. M. Rick Tomato Genetics Resource Center, and Petunia seeds (*P. exserta*, *P. axillaris*, *P. integrifolia*) were obtained from Dr. Cornelius Barry at the Michigan State University Department of Horticulture. All seeds were sterilized with 40 mL, 40% fresh-made bleach (16 mL bleach in 24 mL sterilized MilliQ water), by shaking for 5 min. Liquid was removed, and seeds were rinsed with 40 mL sterilized water, shaking for 5 min. The rinsing step was repeated at least 4 times to minimize residual bleach. Agar media (0.8%) was prepared by adding 4 g Bacto^{TM} agar into 500 mL ¼-strength Hoagland solution in a 1-L Erlenmeyer flask, preparing about 1.5 L media in total. Prepared 0.8% agar media was autoclaved using liquid cycle, with sterilization for 45 min and cooling down for 10 min. Under a Laminar flow bench, hot agar solutions were separated into 250 mL flasks (about 150 mL each), followed by adding about 4.5 mL of pre-filtered (using MILLEX GV filter unit, 0.22 μm) 0.1 mg/mL amino acid solutions: [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu, [$^{13}C_6$,^{15}N]Ile, L-Val, L-Leu, L-Ile and 4.5 mL of ¼-strength Hoagland solutions as control, to make the final concentration of amino acids to 3 μg/mL. This step was repeated to prepare enough amino acid-containing media for all seeds used in labeling experiments. Then, amino acid-containing agar medias were separated into labeled 50-mL polypropylene centrifuge tubes and allowed to cool for 1 h in the laminar flow bench.

Each of the bleached seeds was transferred separately onto 15 mL 0.8% agar media (containing labeled or unlabeled amino acids and Hoagland solution) in a 50-mL polypropylene centrifuge tube. The tubes were covered by 3M film tape (0.22 μm) to ensure bacteria could not penetrate, while air could pass though the film. All equipment and solutions involved were pre-sterilized or

autoclaved, and experiments were performed under a Laminar flow bench. Each tube was transferred into a plant growth chamber at Michigan State University, incubated with 24 hrs daylight, at 25 °C, for 21 days (27 days for *S. pennellii* species). Seeds germinated with 3 µg/mL [$^{13}C_5$,^{15}N]Val-, [$^{13}C_6$,^{15}N]Leu-, and [$^{13}C_6$,^{15}N]Ile-containing agar media, respectively, were labeled as stable isotope labeling groups, and seeds germinated using 3 µg/mL unlabeled L-Val, L-Leu, L-Ile and ¼X-Hoagland solution were assigned as unlabeled control groups. When possible, 3 replicates were applied for each growth condition.

3.2.3 Sample Preparation

After 21 days germination (27 days for *S. pennellii* accessions), all leaflets from each centrifuge tube were sampled and extracted by dipping into 1 mL of isopropanol: acetonitrile: water (3:3:2, *v/v/v*) for 2 min. Extracts were concentrated by drying under vacuum using a SpeedVac (no heating applied). Residues were redissolved into 100 µL acetonitrile: water (4:1, *v/v*). Extracts were vortexed for 30 s and centrifuged for 10 min (25 °C, 10000*g*), followed by transferring about 80 µL supernatant to glass autosampler insert in LC vials. 5 µM telmisartan was added as internal standard and to obtain accurate mass (lock mass at *m/z* 513.2296 in negative-ion mode; *m/z* 515.2442 in positive-ion mode). Extracts were stored at -20 °C until thawed just before UHPLC-MS analyses.

3.2.4 UHPLC-MS analysis with Multiplexed CID

Metabolite profiling of all extracts was performed using a Waters LCT Premier^{TM} time-of-flight mass spectrometer coupled to Shimadzu LC-20AD pumps. Extracted metabolites were resolved using UHPLC separation on an Ascentis Express fused core C18 UHPLC column (10 cm×2.1

mm; 2.7 μm; Supelco, USA) using aqueous 10 mM ammonium formate, adjusted to pH 2.73 with formic acid, as mobile phase A and acetonitrile as mobile phase B. The flow rate was set at 0.3 mL/min, column temperature was 40 °C and the chromatographic run time was 80 min. The UHPLC gradient was: hold at 1% B for 1 min, linear gradient to 80% B over 1-100 min, increase to 100% B during 100-101 min, hold at 100% B at 101-105 min and returned to 1% B over 105-106 min. The system was allowed to re-equilibrate at 1% B for 4 min. 10 μL of each extract was injected using a Shimadzu SIL-5000 autosampler. Mass spectra were acquired in using electrospray ionization in both positive- and negative-ion modes under control of MassLynx software (v.4.1, Waters Inc.), with multiplexed CID approach using W optics. Mass resolution was 8000 (full width half maximum) and data acquisition was performed within 0.1 s for each function over *m/z* 50 to 1500 using dynamic range extension. To generate molecular ion and fragment ions quasi-simultaneously, five parallel functions were applied by switching Aperture 1 voltage between 10, 25, 40, 55, 80 V. Several source parameters were: capillary voltage -2.5 kV in negative mode and +3.0 kV in positive mode, sample cone voltage 10 V, desolvation temperature 350 °C, source temperature 100 °C, cone gas flow 40 L/h and desolvation gas flow 400 L/h. MassLynx and Quanlynx software (v.4.1, Waters Inc.) were used to process mass spectra data. Resolved acylsugar metabolites were annotated based on accurate mass to charge ratio, retention time and fragment ion masses.

3.2.5 Data dependent MS/MS analysis

To trace stable enrichment of specific acylsugar metabolite substructures, data dependent MS/MS analyses were carried out using a Waters Xevo G2-XS QTOF mass spectrometer. In ESI negative mode, anticipated stable isotope precursor ions with *m/z* corresponding to 4 Da and 5

Da heavier than unenriched metabolites were selected as MS/MS precursor for each acylsugar labeled with [$^{13}C_5$,^{15}N]Val, and either [$^{13}C_6$,^{15}N]Leu, or [$^{13}C_6$,^{15}N]Ile, respectively. For example, for unenriched acylsugar S4:17 with [M+formate]$^-$ at *m/z* 681, in its relevant [$^{13}C_6$,^{15}N]Ile labeled sample, the corresponding ion with *m/z* 686 (5 Da heavier) was selected for MS/MS analysis. Mass range for the survey scan was *m/z* 50-1500. Collision potential ramp from 30-80 V was applied in MS/MS function to generated fragments and quadrupole resolution for precursor ion selection for MS/MS analysis was set to 1 *m/z*. LC separation was performed using the same conditions described above.

3.3 Results and discussion

Acylsugars, a group of specialized metabolites that accumulate in glandular trichomes of the plant family *Solanaceae*, are defined as molecules with a carbohydrate core (usually glucose or sucrose) esterified by several carboxylic acids (acyl groups) [31]. Common acyl groups include linear and branched chain aliphatic groups ranging from two to twelve carbon atoms. Plant acylsugars have been detected containing esters drawn from about 20 different acyl groups. Acylsugar nomenclature uses “S” to refer to the sucrose core, followed by a number indicating the number of acyl groups attached to the sugar core, while a second number after the colon represents the total number of carbons in acyl chains. Numbers in parentheses indicate the number of carbon atoms of each individual acyl chain [20,32]. Sucrose contains 8 hydroxyls that could be esterified, and the number of permutations of 20 acyl groups, from 1-8 at a time yields more than 5 billion possible acylsucroses structures from known structural components. Such huge acylsugar diversity exceeds the resolving capacity of liquid chromatography and mass spectrometry alone (Figure 3.1). The biosynthetic pathway of acylsugar-related CoA esters from

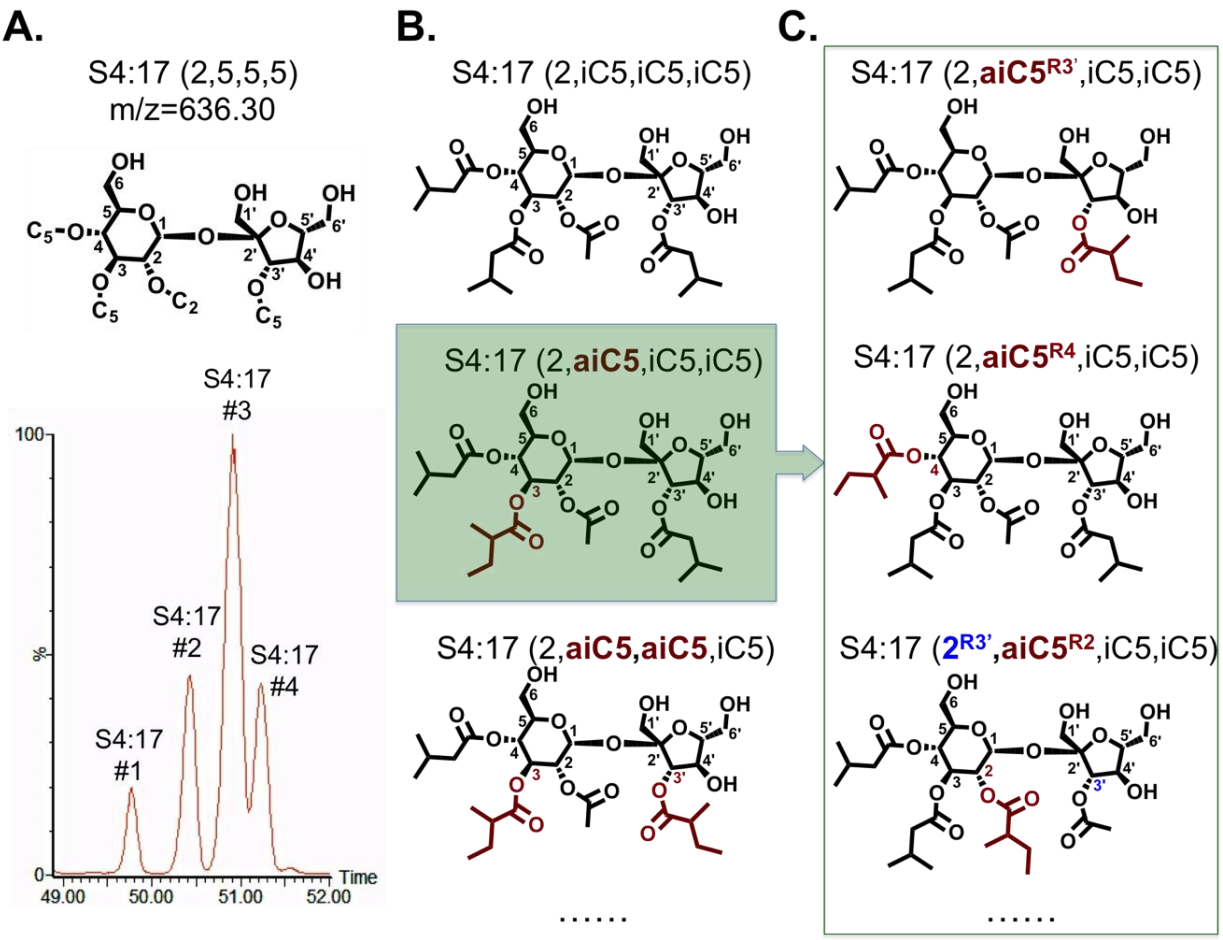


Figure 3.1. Structural diversity of acylsugar S4:17 isomers from tomato. *A. LC/MS chromatogram of four isomeric S4:17 metabolites resolved in a tomato leaf extract; B. Examples of S4:17 isomers with different branching types of C5 aliphatic group. iC5=$(CH_3)_2CHCH_2CO$, aiC5 = $CH_3CH_2(CH_3)CHCO$; C. Examples of isomeric S4:17 (2,aiC5,iC5,iC5) metabolites with various substitution positions of acyl groups.*

several acylsugar-containing plant species has been reported [33,34], which revealed branched chain amino acids (BCAAs) function as acyl group precursors that may undergo elongation. According to this result, by tracing isotopic pattern of incorporated ^{13}C signal originating from uniformly ^{13}C labeled Val, Leu, or Ile, branching type of specific acyl groups of targeted acylsugar metabolites could be easily characterized. Due to metabolic oxidative decarboxylation of the α-ketoacid intermediate in this process, only one fewer ^{13}C atom from labeled precursor

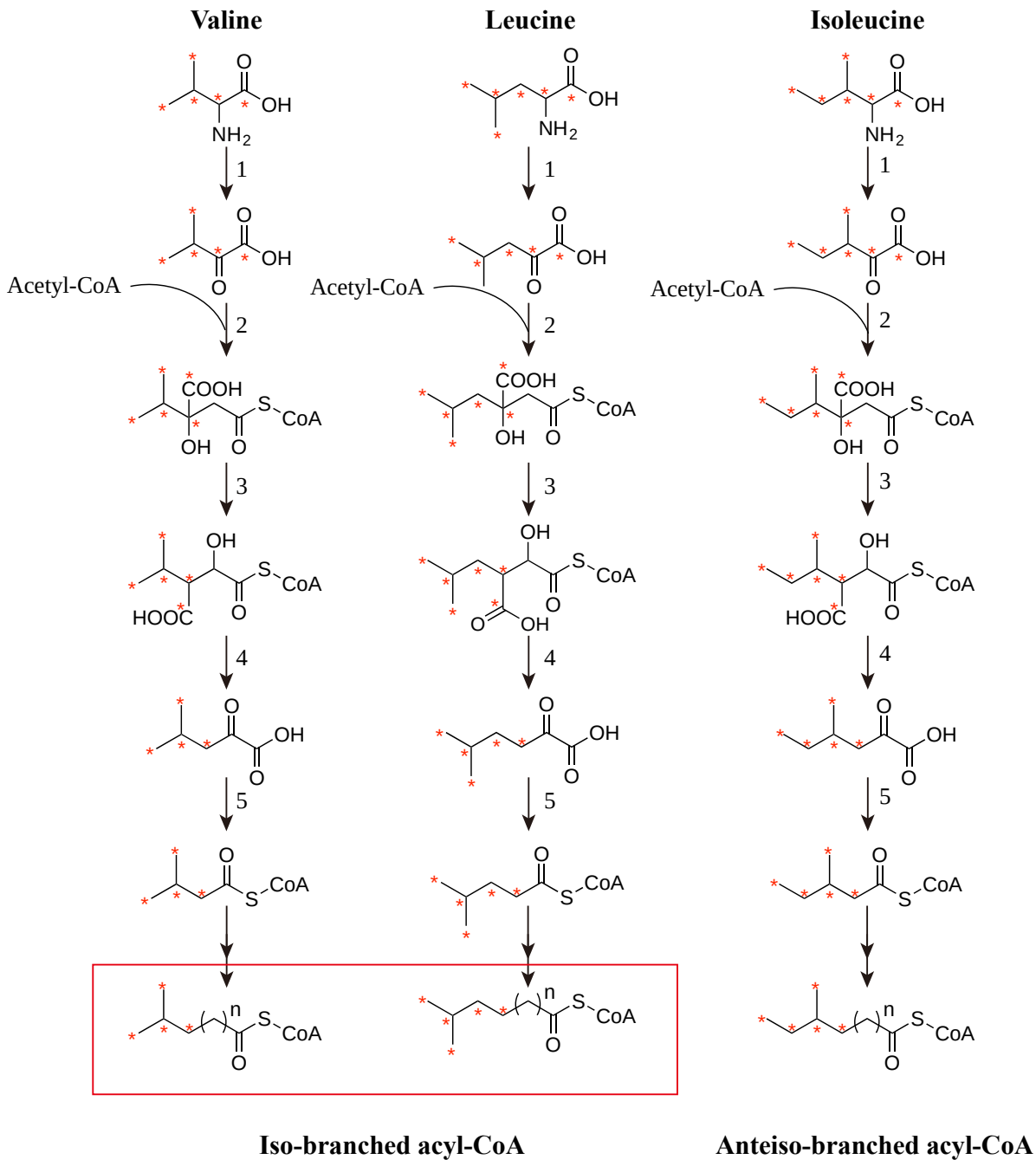


Figure 3.2. Biosynthetic pathway of branched-acyl chains using Branched-Chain Amino Acids (BCAAs) as precursors. *Enzymes involved in branched-acyl chain synthesis: 1. Branched-chain aminotransferase; 2. Isopropylmalate synthases; 3. Isopropylmalate dehydratase; 4. Isopropylmalate dehydrogenase; 5. Branched-chain ketoacid dehydrogenase. ^{13}C atoms are labeled with red stars. Only four ^{13}Cs in Val and five ^{13}Cs in Leu or Ile will be incorporated into branching acyl-CoAs involved in acylsugar synthesis.*

BCAAs will be present in the resulting acyl CoA and incorporated in acylsugar metabolites (Figure 3.2).

Specialized metabolites accumulating in tomato (*S. lycopersicum* M82) trichomes were clearly resolved using UHPLC-MS profiling of leaflet extracts from both labeling (germinated from [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu, or [$^{13}C_6$,^{15}N]Ile containing media) and control groups (grown on unlabeled Val, Leu, Ile in ¼-strength Hoagland solution), indicate that different growth conditions have minimal effect on acylsugar metabolite production (Figure 3.3). Annotation of predominant metabolites using ESI negative mode yielded evidence for the flavonoid glycoside rutin ([M-H]$^-$, *m/z* 609), glycoalkaloid tomatine ([M+HCOO]$^-$, *m/z* 1078.5) and acylsucroses S4:17, S3:22 and S3:24 ([M+HCOO]$^-$, *m/z* 681, 737, and 779). Enrichment of ^{13}C atoms from all

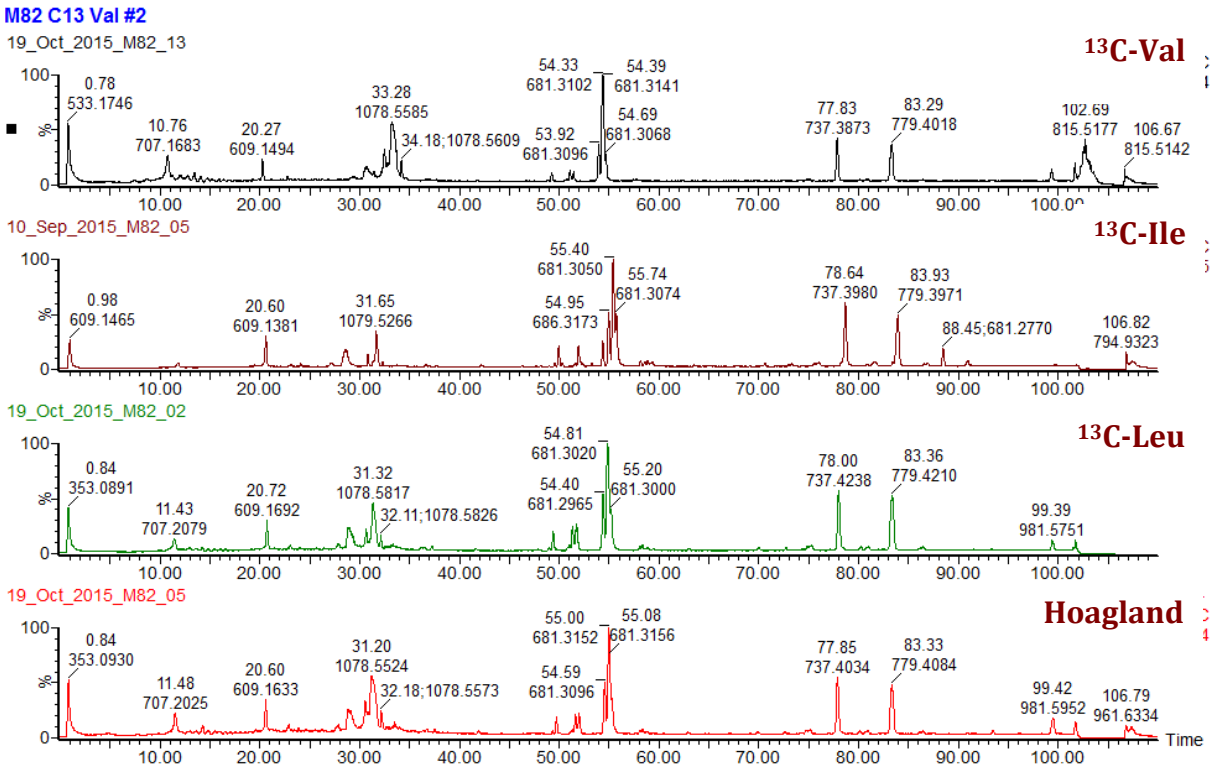


Figure 3.3. Profiling of specialized metabolites of tomato using 110 min LC gradient*. From top to bottom: labeling with [$^{13}C_5$,^{15}N]Val; [$^{13}C_6$,^{15}N]Ile; and [$^{13}C_6$,^{15}N]Leu; control group grown in ¼-strength Hoagland solution.*

three labeled BCAA precursors was observed in tomato acylsugars, and ^{13}C enrichment levels were significantly higher in [$^{13}C_6$,^{15}N]Ile labeled group than for growth on [$^{13}C_5$,^{15}N]Val or [$^{13}C_6$,^{15}N]Leu (Figure 3.4).

Heavy isotope enrichments were traced using both multiplexed CID and MS/MS modes. In the latter case, precursor ions with mass shift by increment of either 4 Da or 5 Da units were selected to correspond to labeling with [$^{13}C_5$,^{15}N]Val and [$^{13}C_6$,^{15}N]Leu or [$^{13}C_6$,^{15}N]Ile, respectively. The presence of ^{13}C-enriched fragment ions in MS/MS spectra provided evidence that ^{13}C-amino acids had been incorporated into acylsugar metabolites in labeled plants. Additionally, isomers of

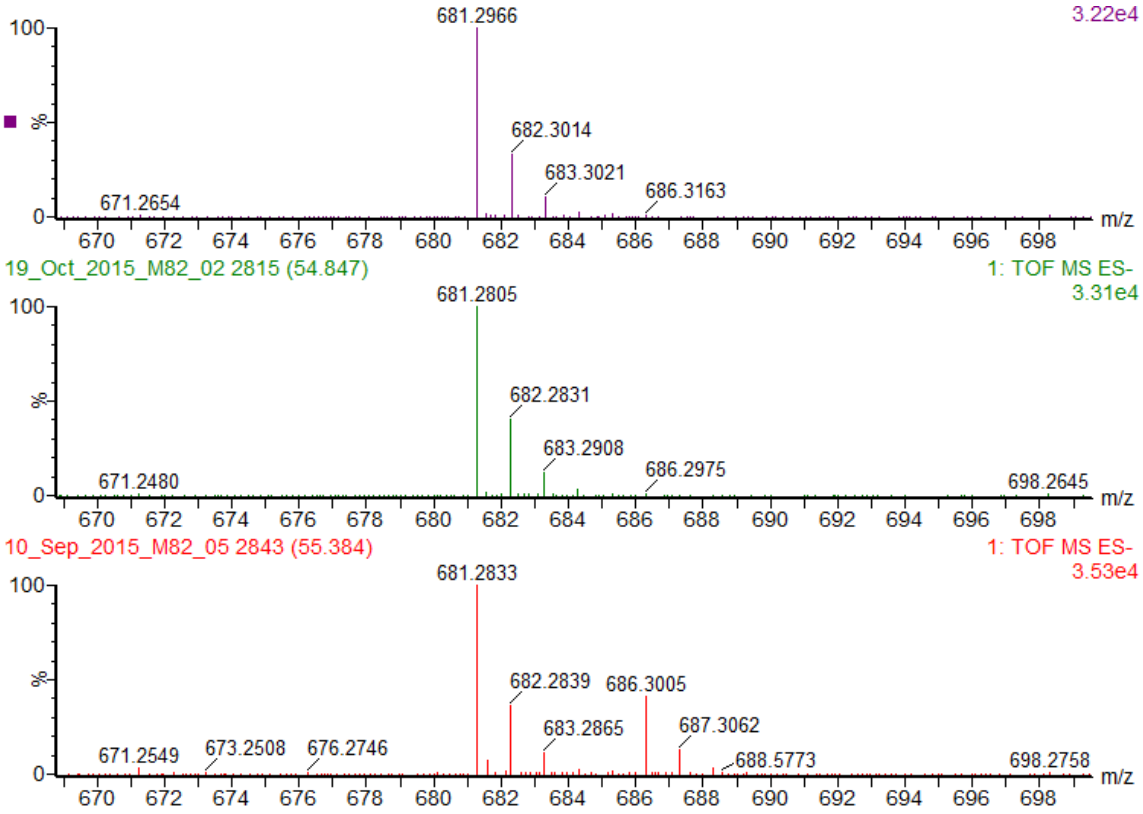


Figure 3.4. Mass spectra of dominant tomato acylsugar S4:17 [3] labeled with three ^{13}C-BCAAs. *From top to bottom: sample labeled with [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu, and [$^{13}C_6$,^{15}N]Ile.*

acylsugars with different branched acyl chain types were distinguished by comparing related isotopolog abundances obtained from plants grown on various ^{13}C-BCAAs. Since iso-branched-acyl groups are derived from corresponding BCAAs with the same branching-type (*e.g.* Val and Leu) and anteiso-branched acyl groups share topology with Ile, anteiso-branched aliphatic chains in plant extracts were annotated as such when ^{13}C was incorporated from [$^{13}C_6$,^{15}N]Ile, whereas [$^{13}C_5$,^{15}N]Val and [$^{13}C_6$,^{15}N]Leu specifically incorporated in iso-branched chains. .

3.3.1 Incorporation of uniformly labeled ^{13}C-Branched Chain Amino Acids (BCAA) in specialized metabolites of wild tomato

To investigate the efficiency of ^{13}C incorporation from labeled BCAA in plant specialized metabolites, a 110-min UHPLC-MS profiling approach was applied to leaf dip extracts of wild tomato, combined with multiplexed CID to analyze isotopologs generated using both labeled and unlabeled growth conditions. Constitutional isomers were resolved by chromatographic retention times. Metabolites enriched with heavy stable isotopes after labeled amino acid incubation are recognized based on altered isotopolog abundances relative to unlabeled controls. For ^{13}C labeled metabolites with sufficient isotope enrichment, multiplexed CID analysis using five parallel data acquisition functions allowed resolution of the locations of incorporated ^{13}C atoms into specific functional groups (Figure 3.5, Figure 3.6). Pseudomolecular [M+formate]$^-$ isotopologs detected at the lowest collision potential show evidence of unlabeled metabolite (*m/z* 681) and metabolite containing five additional neutrons (*m/z* 686) from enrichment of a single [$^{13}C_5$] acyl group (of three total C5 groups). For example, CID spectra of most abundant acylsugar metabolite S4:17[3] (isomers annotated in brackets by elution order) in tomato extracts

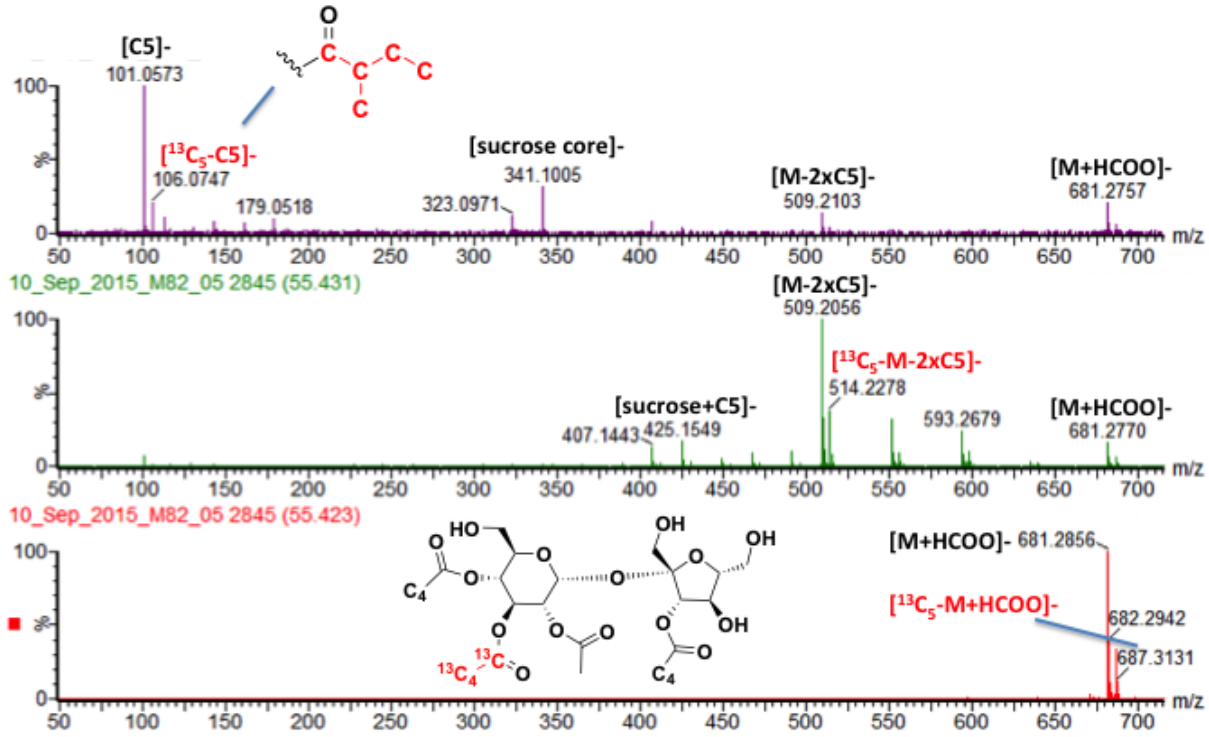
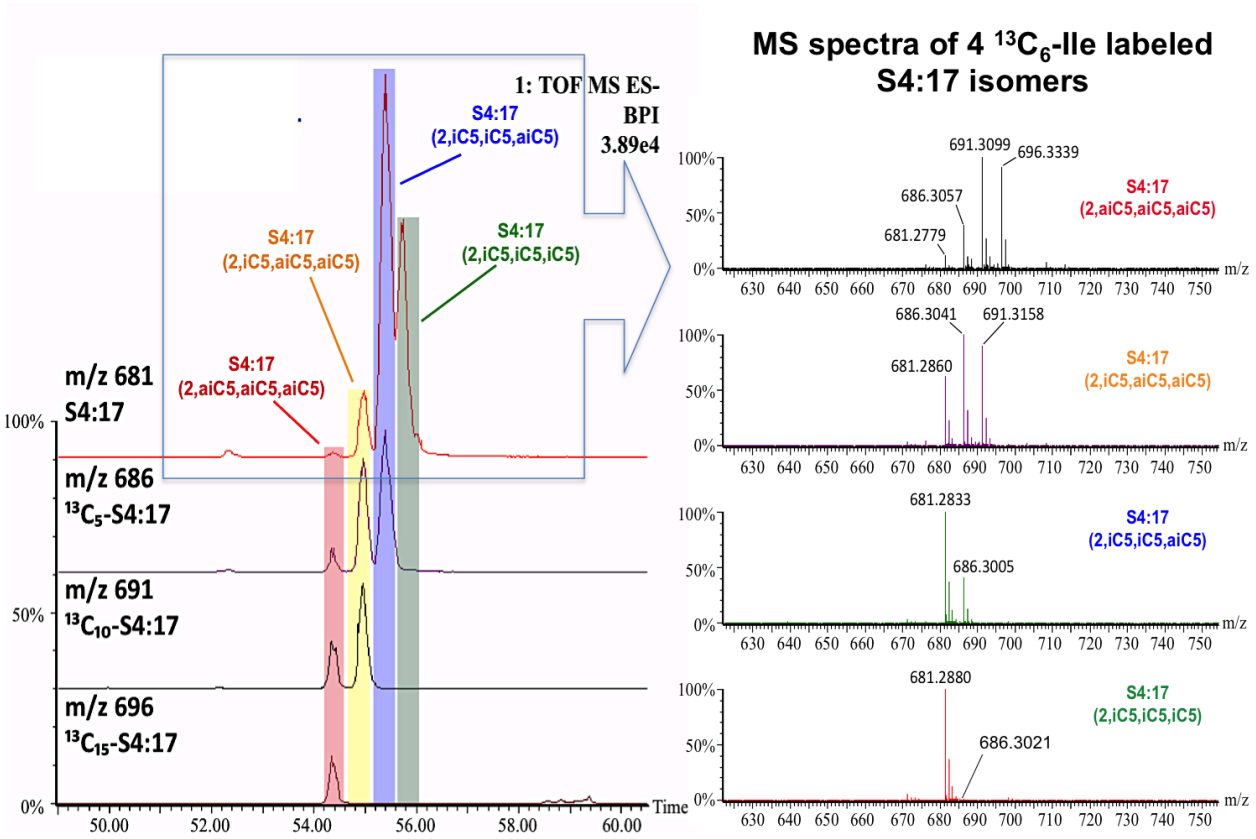


Figure 3.5. Multiplexed CID spectra of the abundant tomato acylsugar S4:17 [3] ([M+formate]⁻ at m/z 681) labeled with [¹³C₆,¹⁵N]Ile. *Collision potentials applied were 10, 55, and 80 V from bottom to top at negative-ion mode ESI. Heavy isotope-containing ions are labeled with red.*

labeled using [$^{13}C_6$,^{15}N]Ile yielded fragment ions at *m/z* 106 that indicate five C atoms of at least one C5 acyl group originated from ^{13}C isotopic tracer, but no enrichment of the sucrose core fragment ion (*m/z* 341) was observed. This assignment is supported further in the fragment ion at *m/z* 514 which is annotated with a data dependent UHPLC-MS/MS strategy was involved in analysis of ^{13}C isotope incorporation in specialized metabolites when extent of labeling was low. Figure 3.7 shows the detailed workflow for UHPLC-MS/MS-based detection of ^{13}C isotope incorporation, using acylsugar metabolite S3:23 (5,6,12) from wild tomato *S. pennellii* LA0716 as example. The target metabolite was found in leaf extracts of LA0716 after 27 days germination from [$^{13}C_6$,^{15}N]Ile contained agar media in growth chamber with 24 hours daylight at 25 °C. From the isotopolog profiling of S3:23, a pseudomolecular [M+formate]$^-$ ion with *m/z*



M82 isotopologue distribution of acylsugar S4:17 isomeric metabolites

	RT (min)	M%	M+5%	M+10%	M+15%
S4:17 [1]	53.43	4.50	18.11	42.01	35.38
S4:17 [2]	53.98	24.97	39.53	35.49	0.01
S4:17 [3]	54.41	75.67	24.31	0.03	0.00
S4:17 [4]	54.72	98.22	1.65	0.12	0.02

Figure 3.6. Heavy isotopolog profiling of isomeric acylsucrose S4:17 in cultivated tomato M82 labeled with [$^{13}C_6$,^{15}N]Ile. *Left: Extracted ion LC/MS chromatograms of S4:17 with 0, 1, 2, 3 [$^{13}C_6$,^{15}N]Ile incorporation (m/z 681, 686, 691, and 696, respectively) from top to bottom. Isomeric peaks with different retention times are color-coded and labeled with proposed acyl chain composition. Right: Isotopologs of isomeric S4:17 [1], S4:17 [2], S4:17 [3], S4:17 [4] with increasing retention time (top to bottom). Percentages of each isotopolog are summarized in the table below the figures.*

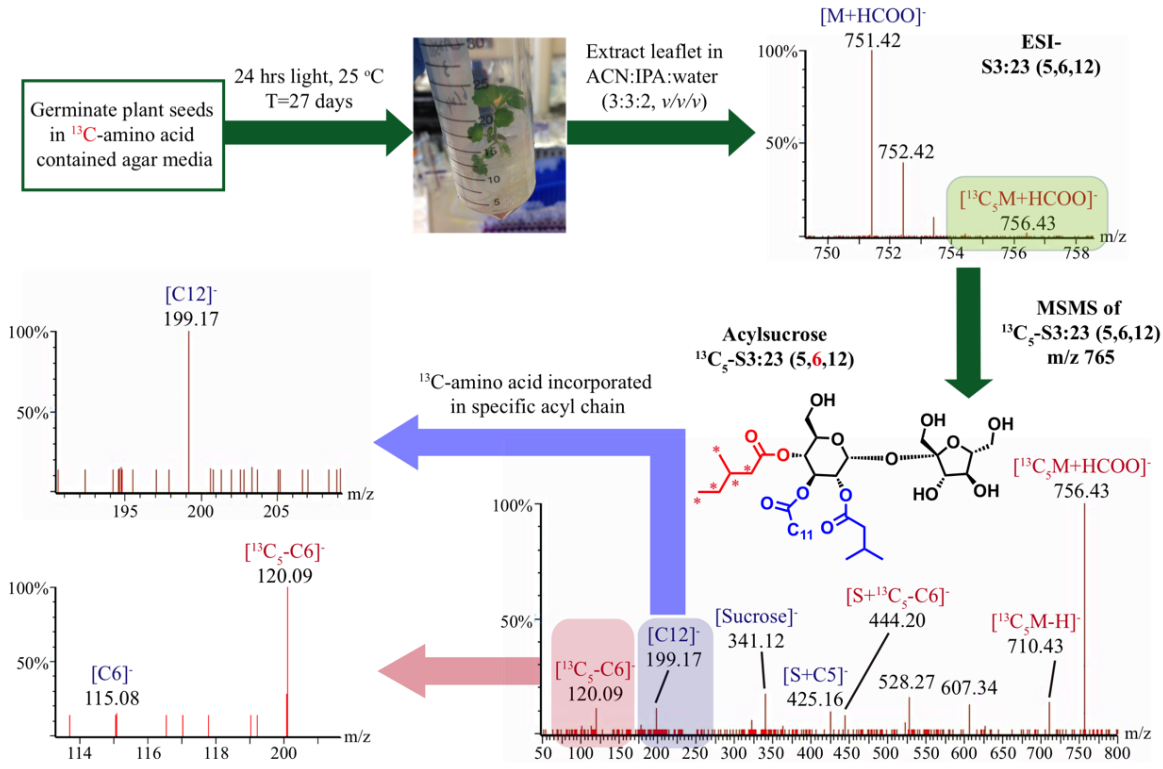


Figure 3.7. Workflow of detecting ^{13}C-BCAA incorporation into acylsucroses using UHPLC-MS/MS. *S3:23 in S. pennellii LA2560 was used as an example to demonstrate the procedure.*

756 was observed in the negative-ion mode mass spectrum, the 5 Da mass increase relative to the monoisotopic adduct $[M + HCOO]^-$ (*m/z* 751) indicated the incorporation of five ^{13}C from labeled Ile during biosynthesis. Multiplexed CID was not suitable to identify isotopolog patterns owing to their low absolute abundances. Therefore, in order to remove interference of unlabeled species and increase sensitivity of detection, this ^{13}C-labeled isotopolog $[^{13}C_5\text{-}M + HCOO]^-$ with *m/z* 756 was chosen as precursor ion for MS/MS analysis. The fate of ^{13}C tracer from $[^{13}C_6,^{15}N]$Ile was revealed in the MS/MS spectrum. The appearance of stable isotope-labeled fragment ions $[^{13}C_5\text{-}M - C12]^-$ (*m/z* 528), $[\text{Sucrose} + {}^{13}C_5\text{-}C6]^-$ (*m/z* 444) demonstrated that five ^{13}Cs from $[^{13}C_6,^{15}N]$Ile were incorporated into the C6 acyl chain of S3:23(5,6,12), whereas C5,

C12 related fragments [M-$^{13}C_5$-C6]$^-$ (*m/z* 607), [Sucrose + C5]$^-$ (*m/z* 425) did not exhibit heavy isotopologs consistent with enrichment. The fragment ion at *m/z* 341 corresponding to deprotonated sucrose core, and its lack of corresponding heavy isotopologs, indicated that carbon atoms from [$^{13}C_6$,^{15}N]Ile were not incorporated into the sugar core of S3:23. More evidence in support of these conclusions is obtained by examining C6 acyl chain carboxylate fragment ion (monoisotopic *m/z* 115) and the corresponding *m/z* 120 ([$^{13}C_5$-C6]$^-$) ion. Since five heavy carbons in the C6 acyl group were incorporated from the labeled Ile precursor, according to BCAA related fatty acid biosynthesis pathway, this particular C6 chain in S3:23(5,6,12) would be anteiso-branched assuming conservation of topology, while C5 and C12 should be either iso-branched or straight chain because they do not contain ^{13}C enrichment and must be derived from other precursors.

3.3.2 Structure diversity assessment of acylsugar related aliphatic chains

A diverse range of aliphatic acyl groups has been identified from acylsugar metabolites in the plant family *Solanaceae*, with either straight or branched (iso- or anteiso-branched) groups ranging in length from C2 to C12. Current understanding holds that aliphatic acyl groups in plant metabolites are synthesized through either fatty acid synthesis pathway or α-ketoacid (αKAE) elongation starting with acetyl CoA, whereas BCAAs are intermediates during biosynthesis that may serve as precursors of branched acyl chains [33,34,35]. Using uniformly ^{13}C-labeled BCAAs as stable isotopic tracers to feed cultivated tomato M82, iso-branched (abbreviated as “i”) acyl groups from iC4 to iC6 were identified by labeling with [$^{13}C_5$,^{15}N]Val and [$^{13}C_6$,^{15}N]Leu, while anteiso-branched (abbreviated as “ai”) aiC5 and aiC6 acyl groups were observed based on incorporation of ^{13}C from [$^{13}C_6$,^{15}N]Ile. C10 and C12 acyl chains not labeled by any of the ^{13}C-

BCAAs were predicted to be straight chain groups derived from linear fatty acid biosynthesis. Other branched acyl groups were observed from labeling of wild tomato relatives and petunia species, where ^{13}C-labeled iC7, iC8 and iC9-11 were detected from several species of *Petunia* and *S. pennellii* species separately (Table 3.1). No anteiso-branched acyl chain longer than C6 has been found in any experimental plant in this investigation.

Table 3.1. Assessment of chemical diversity of acyl groups using 13C-BCAAs labeling approach.

Acyl chains	^{13}C-labeled BCAA precursors	Plant genotypes with labeled acyl groups
iC4	[$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu	*S. lycopersicum* M82, *S. pennellii* LA0716, *S. pennellii* LA2560
iC5	[$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu	*S. lycopersicum* M82, *S. pennellii* LA0716, *S. pennellii* LA2560, *P. exserta*
aiC5	[$^{13}C_6$,^{15}N]Ile	*S. lycopersicum* M82, *S. pennellii* LA0716, *S. pennellii* LA2560, *P. exserta*
iC6	[$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu	*S. lycopersicum* M82
aiC6	[$^{13}C_6$,^{15}N]Ile	*S. lycopersicum* M82, *S. pennellii* LA2560
iC7	[$^{13}C_5$,^{15}N]Val	*P. integrifolia*, *P. exserta*
iC8	[$^{13}C_5$,^{15}N]Val	*P. exserta*
iC9	[$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu	*S. pennellii* LA2560
iC10	[$^{13}C_5$,^{15}N]Val	*S. pennellii* LA0716
nC10	NA	*S. pennellii* LA2560
iC11	[$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu	*S. pennellii* LA2560, *S. pennellii* LA0716
nC12	NA	*S. lycopersicum* M82

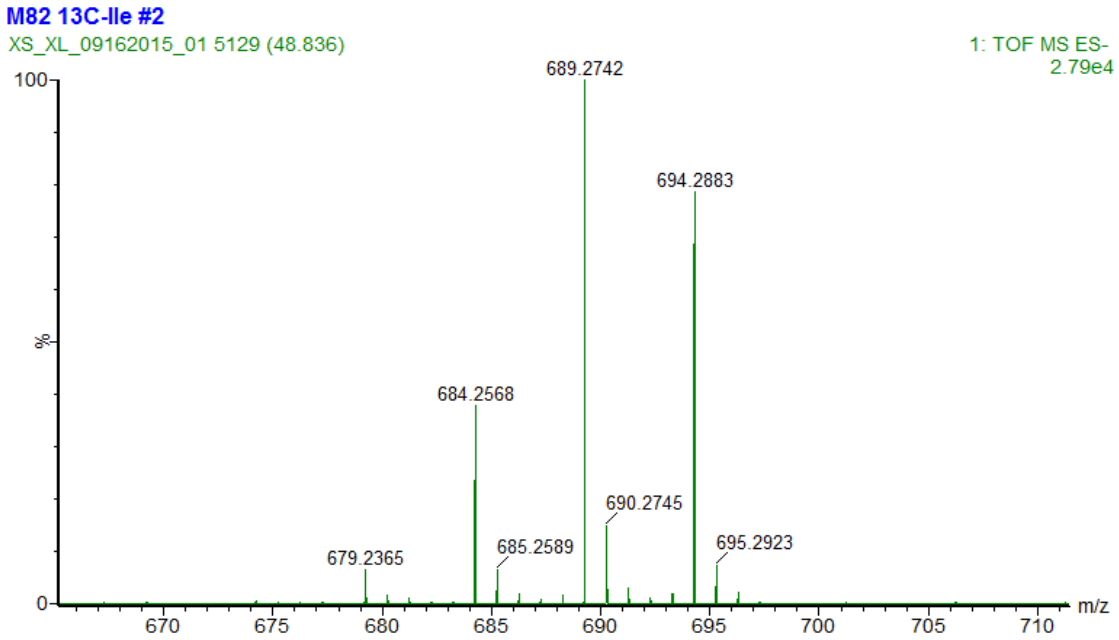
Interestingly, one unsaturated five-carbon (unC5) acyl group has been detected from tomato acylsugar S4:17:1(2,5,5,unC5) detected as [M+formate]$^-$ at *m/z* 679 that was enriched in plants grown on [$^{13}C_6$^{15}N]Ile. Modified nomenclature indicates the number of unsaturations in the third number (*e.g.* S4:17:1). The presence of $^{13}C_5$-unC5 fragments in this acylsugar metabolite was confirmed by data dependent MS/MS, and fragment ions *m/z* 245 and 250 were generated in the

positive-ion mode MS/MS spectra, which indicated that unlabeled and labeled versions of unC5 formed from [$^{13}C_6$,^{15}N]Ile were substituted on the furanose ring. ^{13}C atoms from [$^{13}C_6$,^{15}N]Ile were incorporated in this particular acyl chain, suggesting a five-carbon unsaturated chain with anteiso-branching (Figure 3.8). Additionally, isotopolog profiling of this S4:17:1(2,5,5,unC5) is consistent with incorporation of three anteiso-C5 groups from [$^{13}C_6$,^{15}N]Ile based on $[M+formate+15]^-$ at *m/z* 694 and proposed branching of the unC5 group. This anteiso-unC5 group is putatively assigned as tiglic acid, a natural plant metabolite discovered from croton oil plant [36].

3.3.3 Fast distinguishing of isomeric acylsugar metabolites using multiplexed CID

To distinguish structures of isomeric acylsugars, particularly in terms of branching type of acyl groups, isotopolog profiles of ^{13}C labeled acylsugars were analyzed with multiplexed CID, which gives labeling information in the intact molecules as well as in fragment ions [26]. To demonstrate this strategy, the most abundant acylsugar S4:17(2,5,5,5) in tomato was used as example, and length of each contained acyl groups was verified using multiplexed CID. The three C5 groups in this particular acylsugar could either be branched chains, which should be labeled by [$^{13}C_6$,^{15}N]Ile or [$^{13}C_6$,^{15}N]Leu, or straight chains unlikely to be derived from any ^{13}C-BCAAs. In order to annotate these isomers with various acyl chain branching-types, extracted ion chromatograms (XICs) of fragment ion masses representing all branched acyl chain compositions in S4:17, from all straight chain C5s to all branched chain C5s that have specific number of $^{13}C_6$-Ile/Leu labeling incorporation from 0 to 3, were generated using *m/z* 681 (M), 686 (M+5), 691 (M+10), 696 (M+15), respectively as shown in Figures 3.6 and 3.9.

A.



B.

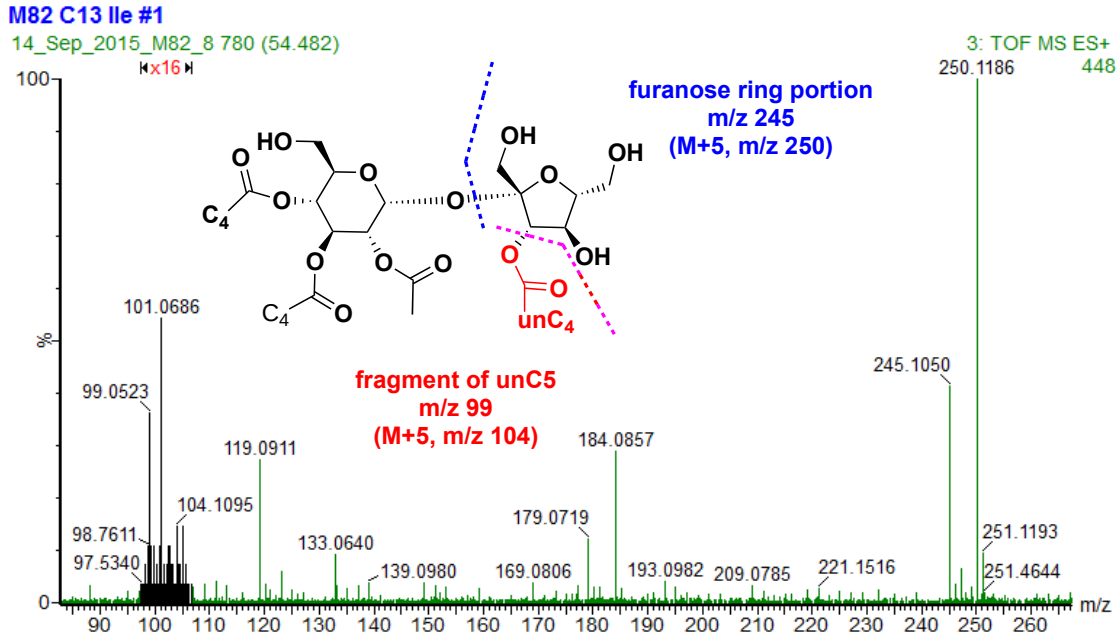


Figure 3.8. CID spectra of tomato S4:17:1 (dominated by [M+formate]⁻ at m/z 679) suggesting an unsaturated C5 chain (unC5) in both positive and negative mode. *A. Isotopolog of S4:17 (monoisotopic [M+formate]⁻, m/z 679); B. Positive mode CID spectrum of S4:17 at function 3 (collision potential at 55 V). Fragment ions at m/z 99, 104 represent unlabeled and one ^{13}C-Ile labeled unC5 (magnified by 16 times). Fragment of furanose ring with m/z 245 and 250 suggest anteiso-unC5 was attached to furanose ring.*

Four isomers of tomato acylsucroses S4:17, designated isomers 1-4, were annotated from ^{13}C-labeling extracts according to chromatographic elution order, and isotopolog profiling of these was performed at the lowest collision energy function of multiplexed CID in negative-ion mode. With [$^{13}C_6$,^{15}N]Ile labeling, S4:17(2,5,5,5) isomer #1 contained ^{13}C-enriched isotopologs containing up to 15 ^{13}C atoms ([$^{13}C_{15}$-M+HCOO]$^-$, *m/z* 696), consistent with incorporation of three chains derived from [$^{13}C_6$,^{15}N]Ile, with labeling indicating that all three C5 chains shared the aiC5 topology. Heavy carbon atoms originating from up to two [$^{13}C_6$,^{15}N]Ile were detected in S4:17 isomer 2 ([$^{13}C_{10}$-M+HCOO]$^-$, *m/z* 691), revealing that two of the three C5 groups were aiC5. Isotopolog of isomers 3 ([$^{13}C_5$-M+HCOO]$^-$, *m/z* 686) and 4 ([M+HCOO]$^-$, *m/z* 681) demonstrated only one or zero [$^{13}C_6$,^{15}N]Ile were incorporated respectively, consistent with one aiC5 in isomer 3 and no aiC5 in isomer 4. Furthermore, ^{13}C-enriched isotopolog profiling produced by [$^{13}C_6$,^{15}N]Leu labeling confirmed that C5 chains not labeled in [$^{13}C_6$,^{15}N]Ile studies were all iC5. By examining isotopologs of isomeric acylsugar metabolites from plants grown on ^{13}C-labeled BCAAs, branching types of acyl groups were characterized from UHPLC-MS data without need for metabolite purification and NMR spectroscopy. A total of 21 abundant acylsugar isomers corresponding to seven different tomato acylsugar formulas (S4:15, S4:16, S4:17, S4:18, S3:22, S3:23, S3:24) were differentiated, in terms of the numbers of acyl groups with specific topologies, using this analysis.

Interestingly, S4:17 labeled by [$^{13}C_6$,^{15}N]Leu and [$^{13}C_6$,^{15}N]Ile generated complementary information regarding order of chromatographic elution of isomers that incorporated different acyl group topologies (Figure 3.9). Acylsugar S4:17 isomer 1, with all three C5 chains being aiC5, eluted first, followed by isomers 2 (two aiC5, one iC5), 3 (one aiC5, two iC5) and four

(three iC5). This finding leads to a more generalized proposal that anteiso-branched acyl chains elute earlier than their iso-branched counterparts using reversed phase liquid chromatography.

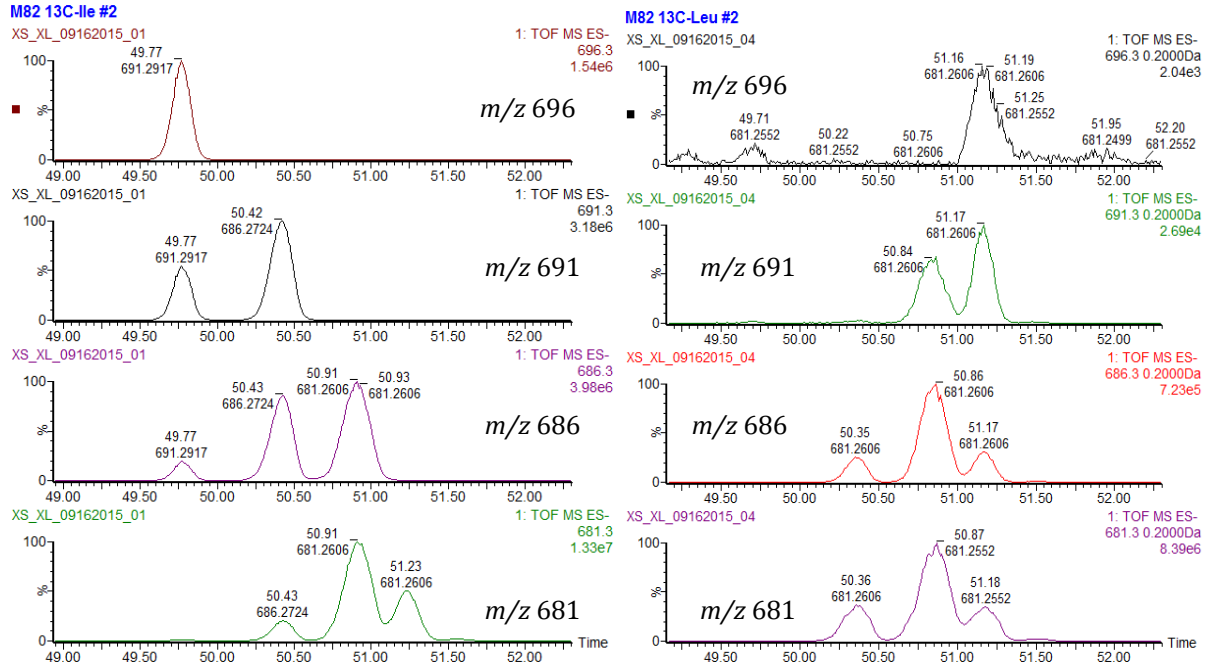


Figure 3.9. Comparison of heavy isotope enrichments in isomers of acylsugar S4:17(2,5,5,5) from tomato plants grown on [$^{13}C_6$,^{15}N]Ile (left) and [$^{13}C_6$,^{15}N]Leu (right) as shown in LC/MS extracted ion chromatograms corresponding to 0, 5, 10, 15 heavy carbons incorporated (bottom to top).

3.3.4 Structure prediction of isomeric acylsugar metabolites

One of the greatest challenges faced when performing LC-MS based plant metabolomics is to establish high-throughput methods for structure elucidation of novel specialized metabolites without requiring tedious purification and NMR interpretation. The results presented above demonstrated the feasibility and success of using fast UHPLC-MS/MS coupled with *in vivo* labeling using ^{13}C-labeled BCAAs to resolve chemical diversity of isomeric specialized metabolites from crude plant extracts. However, one drawback to this approach is that diversity of positions of acylation positions across sugar cores of has yet to be resolved using mass

spectrometry alone, though assignments of groups to the pyranose and furanose rings can be made from positive-ion CID mass spectra. Though some substitution-related structural information, such as the distribution of acyl groups between pyranose and furanose rings, has been derived from previous multiplexed CID analysis of fragment patterns of acylsugars [31], extension of these analyses to assignments of acyl groups to specific hydroxyl positions has been limited by the lack of distinguishing fragment ion masses, as most fragments are formed by

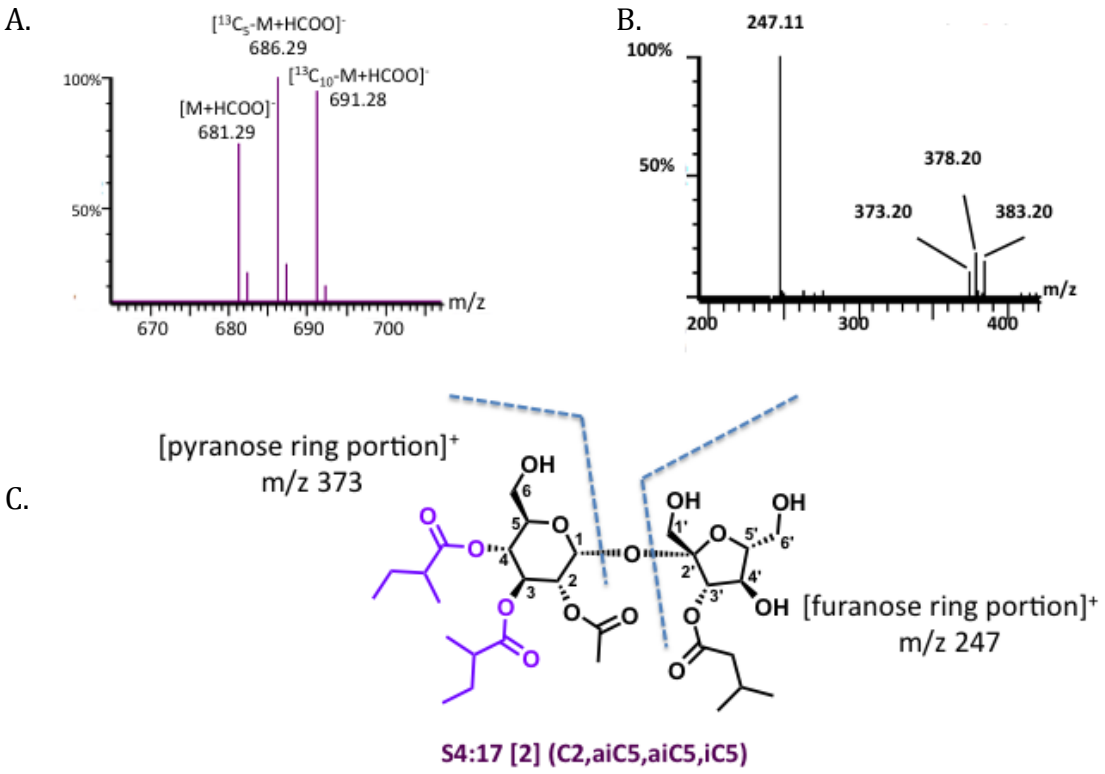


Figure 3.10. Structure annotation of tomato acylsugar S4:17 [2] by in vivo labeling using [$^{13}C_6$,^{15}N]Ile and UHPLC-MS analysis. *A. ESI negative-ion mass spectrum (collision potential of 0 V) of Isotopologs of S4:17 [2] labeled by [$^{13}C_6$,^{15}N]Ile. The [M+formate]- isotopolog with m/z 691 indicates incorporation of 10 heavy carbons consistent with two aiC5 esters. B. ESI positive-ion mass spectrum of S4:17 [2] fragments by glycoside bond cleavage. The fragment ion of m/z 247 corresponds to the furanose ring fragment containing one C5 but no heavy carbon atoms; the fragment ion with m/z 373 corresponds to the monoisotopic pyranose ring portion with two C5 groups, and heavy isotopologs at m/z 378 and 383 represent incorporation of one and two [$^{13}C_6$,^{15}N]Ile into heavy C5 chains. C. Structure of S4:17 [2] (2,5,5,5)*

losses of acyl groups. Fortunately, with our deeper understanding of activity and substrate selectivity of acylsugar biosynthesis related enzymes, acylsucrose acyltransferases (ASATs), the first step of structure prediction of isomeric acylsugar metabolites in cultivated tomato M82 was taken [37]. The most abundant acylsucrose S4:17 was used as an example to describe this attempt (Figure 3.10).

Preferable acylation positions of tomato acylsucrose metabolites have been assigned to positions 2, 3, 4 and 3′, based on *in vitro* reconstruction study of acylsucrose sequential biosynthesis using four major tomato ASATs. Exclusive acylation at the 2 position has been established as C2 (acetyl) owing to SlASAT-4 substrate selectivity [37]. According to results from isotopolog profiling of [$^{13}C_6$,^{15}N]Ile-labeled S4:17, structures of isomer #1 and isomer #4 of S4:17, equipped with either three aiC5 chains or three iC5 groups, separately, were proposed as one C2 attached as R2, and other branched C5 substituted on R3, R4 and R3′. In contrast, predictions of all acylations of isomers #2 and #3 were more challenging since branching types of the three C5 groups were varied. A structure of S4:17[2](2,iC5,aiC5,aiC5) was proposed with the help of ESI positive mode MS/MS data from enrichment using [$^{13}C_6$,^{15}N]Ile, with iC5 assigned to R3′ on the furanose ring in order to generate the furanose fragment ion of *m/z* 247 with no ^{13}C incorporation. Acylated pyranose fragments with five and ten ^{13}C enrichment (*m/z* 378, *m/z* 383) provided more evidence to support assignment of aiC5s to both the R3 and R4 positions. (Figure 3.10). Positive mode spectrum of S4:17[3](2,iC5,iC5,aiC5) also yielded the unlabeled acylfuranose fragment ion (*m/z* 247) and a single $^{13}C_6$-labeled pyranose (*m/z* 378) fragment ion, indicative that iC5 was acylated at R3′. Since the enzyme activity of tomato ASAT2, which adds the second acyl chain to the R3 position of acylsucrose intermediates using a variety of acyl CoA substrates but showed poor activity using iC5-CoA [8], it is proposed that aiC5 is the substituent

at R3 and iC5 at R4. Proposed structures of 17 abundant tomato isomeric acylsugars using the same approach are summarized in Table 3.2. Furthermore, this prediction of structure of isomer #3 of S4:17 was consistent with a previous reported structure of a tomato acylsucrose metabolite that matched chromatographic retention time and molecular ion mass, and had its structure established using NMR analysis [3]. With in-depth understanding of enzyme bioactivity related to synthesis of tomato acylsucroses, UHPLC-MS/MS coupled with ^{13}C-BCAAs labeling approach provides a useful pathway to annotate structures of plant specialized metabolites without pre-purification and NMR.

Table 3.2. Proposed structure annotations of the dominant tomato acylsugar isomers using in vivo ^{13}C-BCAAs labeling combined with SlASATs bioactivity. *Isomeric acylsugars are numbered based on chromatographic elution order.*

Acylsugar	Isomer number	R2	R3	R4	R3′	Branching type (i/ai)
S4:17 (2,5,5,5)	#1	C2	aiC5	aiC5	aiC5	3 ai
	#2	C2	aiC5	aiC5	iC5	2 ai
	#3	C2	aiC5	iC5	iC5	1 ai
	#4	C2	iC5	iC5	iC5	no ai
S4:18 (2,5,5,6)	#1	C2	aiC5	aiC5	aiC6	3 ai
	#2	C2	aiC5	iC5	aiC6	2 ai
	#3	C2	aiC5	aiC6	iC5	2 ai
	#4	C2	aiC5	iC6	iC5	1 ai
	#5	C2	iC6	iC5	iC5	no ai
S3:22 (5,5,12)	#1	NA	nC12	aiC5	aiC5	2 ai
	#2	NA	nC12	aiC5	iC5	1 ai
S3:23 (5,6,12)	#1	NA	nC12	aiC5	aiC6	2 ai
	#2	NA	nC12	aiC6	iC5	1 ai
	#3	NA	nC12	iC6	iC5	no ai
S4:24 (2,5,5,12)	#1	C2	nC12	aiC5	aiC5	2 ai
	#2	C2	nC12	aiC5	iC5	1 ai
	#3	C2	nC12	iC5	iC5	no ai

3.3.5 Performance evaluation and prospective applications

The combination of UHPLC-MS/MS coupled with ^{13}C-BCAAs labeling offers potential to advance untargeted discovery and chemical diversity characterization of all specialized metabolites derived from BCAA metabolic pathways. In cultivated tomato M82, stable isotopic atoms of ^{13}C-BCAAs were detected only from acylsugar metabolites, not in alkaloids (tomatine) or flavonoids (rutin), which was consistent with our expectations since BCAAs pathway was not related with their biosynthesis. The discovery of $^{13}C_6$-Ile labeled tiglic acid residue within low abundance acylsugar S4:17:1 (*m/z* 679) demonstrated the possibility to expand this approach into identification of novel BCAA related specialized metabolites, as well as investigation of unrevealed BCAA related metabolic pathway. A total of six genotypes from the family *Solanaceae* selected for this investigation and incorporated with detectable ^{13}C labeling pattern, which means this technique should be suitable for metabolic analysis of wide range of plants.

Stable isotope enrichment levels in plant specialized metabolites is one of the dominant factors that could affect the utility of this ^{13}C-BCAAs labeling coupled with UHPLC-MS/MS approach, mainly because low (< 1 mol%) ^{13}C isotope enrichment in specific compound cannot be distinguished during analysis of crude extracts when coeluting interfering ions at the same mass are present. Levels of isotope incorporation into specific metabolites are likely affected by stage of plant and tissue development, environmental conditions, and ^{13}C tracers used. According to our investigation, healthy plants with 21 and 27 days of incubation showed almost the same ^{13}C enrichment level and appeared to reach a steady state relative to photosynthetic carbon fixation (of natural isotope abundance), whereas plants germinated from [$^{13}C_6$,^{15}N]Ile-containing media showed greater ^{13}C incorporation than for the other two isotope tracers (Table 3.3). Though longer light exposure would be expected to increase extent of photosynthetic fixation of carbon

at natural abundance isotope ratios, rates of nutrient uptake and rates of biosynthesis of specialized metabolites may also be affected. Since no ^{13}C enrichment was observed in plant central metabolites including sucrose and glucose, it is interpreted that recycling of carbon from ^{13}C-enriched precursors into central metabolism was minimal under the conditions used in this investigation.

Table 3.3. Stable isotope enrichment in tomato acylsugars. *Calculated relative abundances of [M+formate]$^-$ isotopologs $A_{[m+5]}/A_{[m]}$, $A_{[m+4]}/A_{[m]}$ represented ^{13}C enrichment from [$^{13}C_6$,^{15}N]Ile, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_5$,^{15}N]Val labeled samples separately. Hoagland solution samples were used in growth media for the unlabeled control group.*

Tomato Acylsugars	Control (m+4/m%)	Control (m+4/m%)	[$^{13}C_6$,^{15}N] Ile (m+5/m%)	[$^{13}C_6$,^{15}N] Leu (m+5/m%)	[$^{13}C_5$,^{15}N] Val (m+4/m%)
S3:15	4.2E-01	7.6E-02	1.6E+01	1.9E+00	3.6E+00
S4:16	2.3E-01	7.0E-02	9.7E+00	2.4E+00	6.0E+00
S4:17	4.8E-01	1.1E-01	3.8E+01	3.2E+00	5.1E+00
S4:18	5.2E-01	2.2E-01	5.6E+01	2.8E+00	4.2E+00
S4:22	9.3E-01	3.8E-01	3.5E+01	2.3E+00	8.2E+00
S3:22	4.8E-01	6.2E-02	9.0E+00	2.8E+00	1.6E+00
S4:24	5.2E-01	1.2E-01	1.5E+01	9.0E+00	1.2E+01
Total acylsugar	4.8E-01	1.1E-01	2.6E+01	3.7E+00	4.7E+00

3.4 Conclusions

Advances in strategies for identifying and distinguishing structures of novel metabolites have lagged far behind improvements in modern genomic DNA sequencing technologies. It is now faster and less costly to sequence an entire genome than to identify a novel specialized metabolite. Metabolite identification now represents one of the most significant bottlenecks to discovery of novel genes and enzymes that give rise to natural products of potential importance to medicine and agriculture. One attempt to overcome these limitations has been described here, involving *in vivo* labeling with ^{13}C-labeled amino acid and UHPLC-MS analysis. This approach

is particularly effective for discovering aliphatic chain lengths and branching in acylated natural products. Based on accurate masses, relative retention times and characteristic fragment ions generated by UHPLC-MS with different collision potentials, a reliable and sensitive method was established that improved throughput for metabolite annotation. Meanwhile, stable isotope labeling was introduced in order to annotate isomeric metabolites with same pseudomolecular and fragment ion masses but different molecular topologies. Combination of the two methods accelerated investigation of structural diversity of plant specialized metabolites without need for compound purification. As more plant acyltransferase substrate selectivities become known, we hope to propose structures of some isomeric acylsugars without purification and NMR analysis.

Furthermore, one advantage of labeling with precursors for which all carbons are ^{13}C derives from anticipation that each derived product will experience a distinct molecular mass shift (*e.g.* M+4, M+5), that avoids interference from peaks corresponding natural isotopic peaks of unenriched species, even when the labeled form has relative abundance below 1%. In such cases, even with low ^{13}C enrichment level, stable isotope labeled isotopologs can be selected and analyzed for their isotopic labeling pattern by using MS/MS analysis.

Although this investigation focused on isomeric acylsugar metabolites, the analytical approach described in this report is expected to have broad application to a wide range of amino acid-derived specialized metabolites not limited to the plant kingdom. Since ^{13}C amino acids are universal precursors involved in biosynthesis of large groups of metabolites, all the relevant intermediates and by-products could be labeled and investigated providing precursor uptake is significant. We anticipate the information obtained will lead to a more global recognition and

understanding of valuable biosynthetic pathways that employ BCAA-derived metabolite biosynthetic pathways across a wide range of organisms.

REFERENCES

REFERENCES

1. Fiehn, O., Combining Genomics, Metabolome Analysis, and Biochemical Modelling to Understand Metabolic Networks. *Comparative and Functional Genomics* **2001,** *2* (3), 155-168.

2. Pichersky, E.; Lewinsohn, E., Convergent evolution in plant specialized metabolism. *Annual Review of Plant Biology* **2011,** *62*, 549-66.

3. Duke, S. O.; Canel, C.; Rimando, A. M.; Telle, M. R.; Duke, M. V.; Paul, R. N., Current and potential exploitation of plant glandular trichome productivity. *Advances in Botanical Research* **2000,** *31*, 121-151.

4. Wagner, G. J., Secreting glandular trichomes: more than just hairs. *Plant Physiology* **1991,** *96* (3), 675-679.

5. Balandrin, M. F.; Klocke, J. A.; Wurtele, E. S.; Bollinger, W. H., Natural plant chemicals: sources of industrial and medicinal materials. *Science(Washington)* **1985,** *228* (4704), 1154-1159.

6. Cragg, G. M.; Newman, D. J., Plants as a source of anti-cancer agents. *Journal of Ethnopharmacology* **2005,** *100* (1), 72-79.

7. Veeresham, C., Natural products derived from plants as a source of drugs. *Journal of Advanced Pharmaceutical Technology & Research* **2012,** *3* (4), 200.

8. Last, R. L.; Jones, A. D.; Shachar-Hill, Y., Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* **2007,** *8* (2), 167-174.

9. Rolin, D.; Deborde, C.; Maucourt, M.; Cabasson, C.; Fauvelle, F.; Jacob, D.; Canlet, C.; Moing, A.; Rolin, D., High-resolution 1H-NMR spectroscopy and beyond to explore plant metabolome. *Academic Press: London* **2013,** *67*, 99-158.

10. Nicholson, J. K.; Lindon, J. C., Systems biology: metabonomics. *Nature* **2008,** *455* (7216), 1054-1056.

11. Hartmann, T.; Kutchan, T. M.; Strack, D., Evolution of metabolic diversity. *Phytochemistry:* **2005**, *66* (11), 1198-1199

12. Benderoth, M.; Textor, S.; Windsor, A. J.; Mitchell-Olds, T.; Gershenzon, J.; Kroymann, J., Positive selection driving diversification in plant secondary metabolism. *Proceedings of the National Academy of Sciences of the USA* **2006,** *103* (24), 9118-9123.

13. MacBeath, G.; Saghatelian, A., The promise and challenge of '-omic' approaches. *Current Opinion in Chemical Biology* **2009**. *13*, 501-502.

14. Taggi, A. E.; Meinwald, J.; Schroeder, F. C., A new approach to natural products discovery exemplified by the identification of sulfated nucleosides in spider venom. *Journal of the American Chemical Society* **2004,** *126* (33), 10364-10369.

15. Vinayavekhin, N.; Saghatelian, A., Discovery of a Protein–Metabolite Interaction between Unsaturated Fatty Acids and the Nuclear Receptor Nur77 Using a Metabolomics Approach. *Journal of the American Chemical Society* **2011,** *133* (43), 17168-17171.

16. Lv, H.; Hung, C. S.; Chaturvedi, K. S.; Hooton, T. M.; Henderson, J. P., Development of an integrated metabolomic profiling approach for infectious diseases research. *Analyst* **2011,** *136* (22), 4752-4763.

17. Hagel, J. M.; Facchini, P. J., Plant metabolomics: analytical platforms and integration with functional genomics. *Phytochemistry Reviews* **2008,** *7* (3), 479-497.

18. Kind, T.; Wohlgemuth, G.; Lee, D. Y.; Lu, Y.; Palazoglu, M.; Shahbaz, S.; Fiehn, O., FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry* **2009,** *81* (24), 10038-10048.

19. Wishart, D. S., Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery* **2016**., *15*, 473-484

20. Ghosh, B.; Westbrook, T. C.; Jones, A. D., Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014,** *10* (3), 496-507.

21. Molinski, T. F., NMR of natural products at the 'nanomole-scale'. *Natural Product Reports* **2010,** *27* (3), 321-329.

22. Schroeder, F. C.; Taggi, A. E.; Gronquist, M.; Malik, R. U.; Grant, J. B.; Eisner, T.; Meinwald, J., NMR-spectroscopic screening of spider venom reveals sulfated nucleosides as major components for the brown recluse and related species. *Proceedings of the National Academy of Sciences of the USA* **2008,** *105* (38), 14283-14287.

23. Dettmer, K.; Aronov, P. A.; Hammock, B. D., Mass spectrometry- based metabolomics. *Mass Spectrometry Reviews* **2007,** *26* (1), 51-78.

24. Ekanayaka, E. A. P.; Celiz, M. D.; Jones, A. D., Relative mass defect filtering of mass spectra: a path to discovery of plant specialized metabolites. *Plant Physiology* **2015,** *167* (4), 1221-1232.

25. Hegeman, A. D., Plant metabolomics—meeting the analytical challenges of comprehensive metabolite analysis. *Briefings in Functional Genomics* **2010,** *9* (2), 139-148.

26. Wang, Z.; Jones, A. D., Profiling of stable isotope enrichment in specialized metabolites using liquid chromatography and multiplexed nonselective collision-induced dissociation. *Analytical Chemistry* **2014,** *86* (21), 10600-10607.

27. Meissen, J. K.; Pirman, D. A.; Wan, M.; Miller, E.; Jatkar, A.; Miller, R.; Steenwyk, R. C.; Blatnik, M., Phenotyping hepatocellular metabolism using uniformly labeled carbon-13 molecular probes and LC-HRMS stable isotope tracing. *Analytical Biochemistry* **2016,** *508*, 129-137.

28. Creek, D. J.; Chokkathukalam, A.; Jankevics, A.; Burgess, K. E. V.; Breitling, R.; Barrett, M. P., Stable isotope-assisted metabolomics for network-wide metabolic pathway elucidation. *Analytical Chemistry* **2012,** *84* (20), 8442-8447.

29. Gómez-Cortés, P.; Brenna, J. T.; Sacks, G. L., Production of Isotopically Labeled Standards from a Uniformly Labeled Precursor for Quantitative Volatile Metabolomic Studies. *Analytical Chemistry* **2012,** *84* (12), 5400-5406.

30. Huang, K.; Huang, L.; Van Breemen, R. B., Detection of reactive metabolites using isotope-labeled glutathione trapping and simultaneous neutral loss and precursor ion scanning with ultra-high-pressure liquid chromatography triple quadrupole mass spectrometry. *Analytical Chemistry* **2015,** *87* (7), 3646-3654.

31. Ghosh, B.; Jones, A. D., Profiling, characterization, and analysis of natural and synthetic acylsugars (sugar esters). *Anal. Methods* **2017,** *9* (6), 892-905.

32. Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Jones, A. D.; Last, R. L., Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. *The Plant Journal* **2010,** *62* (3), 391-403.

33. Kroumova, A. B.; Xie, Z.; Wagner, G. J., A pathway for the biosynthesis of straight and branched, odd-and even-length, medium-chain fatty acids in plants. *Proceedings of the National Academy of Sciences of the USA* **1994,** *91* (24), 11437-11441.

34. Kroumova, A. B.; Wagner, G. J., Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants. *Planta* **2003,** *216* (6), 1013-21.

35. Ohlrogge, J. B.; Jaworski, J. G., Regulation of fatty acid synthesis. *Annual Review of Plant Biology* **1997,** *48* (1), 109-136.

36. Attygalle, A. B.; Wu, X.; Will, K. W., Biosynthesis of tiglic, ethacrylic, and 2-methylbutyric acids in a carabid beetle, *Pterostichus* (Hypherpes) *californicus*. *Journal of Chemical Ecology* **2007,** *33* (5), 963-970.

37. Fan, P.; Miller, A. M.; Schilmiller, A. L.; Liu, X.; Ofner, I.; Jones, A. D.; Zamir, D.; Last, R. L., In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. *Proceedings of the National Academy of Sciences of the USA* **2016,** *113* (2), E239-E248.

Chapter Four.

Investigation of acylsugar related aliphatic acid elongation using *in vivo* ^{13}C-amino acid labeling and UHPLC-QTOF-MS/MS analysis

4.1 Introduction

The dramatic growth of world population within the last century brings tremendous challenges to modern agriculture. Metabolic engineering of crop plants provides novel solutions to increase crop productivity, reduce cost of agricultural production and decrease adverse impacts on the environment [1]. Large numbers of structurally diverse specialized metabolites, many of which are commercially valuable compounds with ecological and therapeutic importance, accumulate within plants and contribute to plant chemical defenses against herbivores and environmental stress [2-4]. Additionally, since biosynthetic pathways of specialized metabolites have been modified by evolution and natural selection, plants can be viewed as efficient "factories" that produce bioactive chemicals in a more cost-effective manner than artificial synthesis [5]. In many cases, specialized metabolism occurs in specific plant cell types such as storage/secreting glandular trichomes (SGTs), which are epidermal appendages that play critical roles in plant chemical defenses [6]. SGTs contribute to plant resistance through several ways including synthesis and accumulation of a variety of volatile and nonvolatile specialized metabolites such as terpenes, acylsugars, methylketones, and flavonoids, many of which are toxic or repellent to insects and pathogens. There is significant chemical diversity of specialized metabolites not only within different type of SGTs, but across different species of the genus *Solanum*, which includes tomato, potato, and eggplant. Additionally, as aerial surface cells, glandular trichomes are accessible and can be selectively sampled for metabolites and mRNA. Therefore, SGTs have been used as model tissues to investigate specialized metabolism in plants and its genetic basis. A current challenge in plant metabolism is that biosynthesis and accumulation mechanisms of specialized metabolites in plant tissue are still poorly understood [7,8], largely because the identities of only a small fraction of specialized metabolites have been defined. Hence, a

promising approach to overcome this challenge involves renewed efforts to identify specialized metabolites and the enzymes and genes that are related to metabolite biosynthesis.

Acylsugars (also called sugar esters), which are usually produced by type I/IV glandular trichomes in the family *Solanaceae*, are the most abundant class of specialized metabolites in many Solanaceous plants, and have been documented to function as defenses against insect herbivores [9,10]. Though acylsugars are the dominant specialized metabolites in several Solanaceous species, limited numbers of structures of acylglucoses [11-13] and acylsucroses [14-18] have been proposed and reported in previous reports. Substantial chemical diversity of acylsugars has been observed among species, usually arising from different acyl chain lengths (C2-C12), straight or branched aliphatic chains, number of ester groups, and positions of substitution [19]. Acylsugars, especially those substituted by acyl chains longer than eight carbons, play key roles in plant resistance against several insects including psyllids, aphids, whiteflies and spider mites [20-23]. Synthetic variants of these specialized metabolites also have commercial applications as food additives and cosmetic products [23,24]. However, in contrast to the enormous diversity in acylsugars within and across species, our understanding of the factors that drive acylsugar diversity, particularly with respect to identifications of biosynthetic intermediates, the genes and enzymes that produce them, and dynamics of metabolic fluxes remains limited.

According to current understanding of acylsugar structures, three major steps in the biosynthesis of acylsugar metabolites in Solanaceous plants contribute to their chemical diversity: biosynthesis of sugar cores; biosynthesis of acylsugar-related aliphatic acyl-CoAs; and acylation of specific hydroxyl groups on the sugar core from select acyl-CoAs. Glucose and sucrose are common sugar cores of acylsugar metabolites in the *Solanaceae*. A previous study showed UDP-glucose was used as precursor to produce triacylglucose catalyzed by *S. pennellii* extracts [25], though ongoing and unpublished research in the laboratory of Professor Robert Last has suggested alternative pathways to acylglucoses. More recently, several trichome-specific BAHD (**B**EAT, **A**HCT, **H**CBT, **D**AT) [26,27] class acyltransferases, named AcylSugar AcylTransferases (ASATs) were identified from Solanaceous species, and these catalyzed

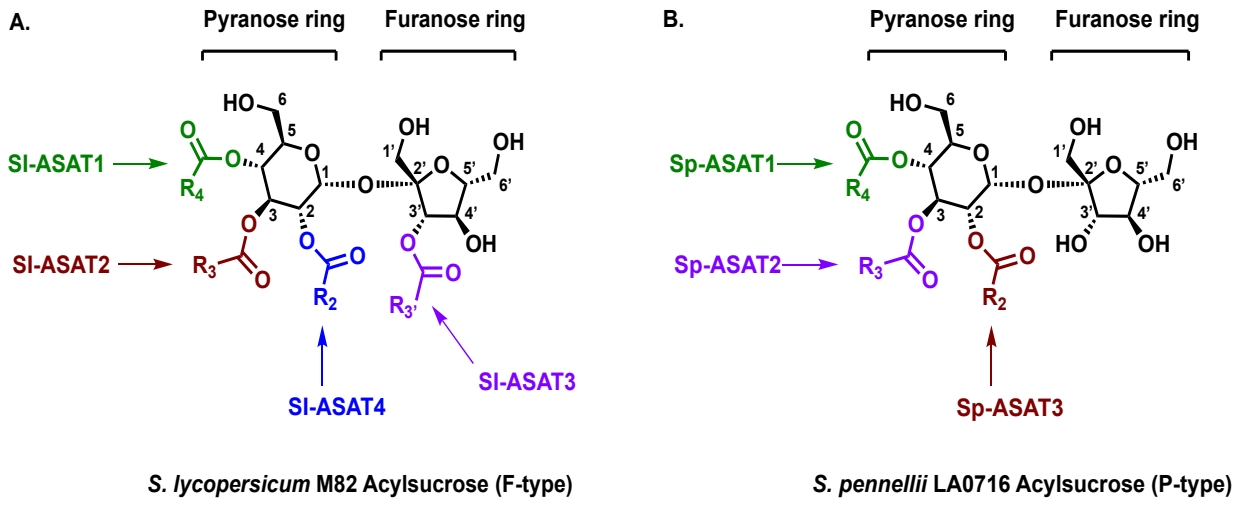


Figure 4.1. Examples of acylsugar structures and their flipped biosynthesis pathways from S. lycopersicum M82 (F-type) and S. pennellii LA0716 (P-type). *A. Structure of tetra-acylsucrose from S. lycopersicum M82 with acylations on both the pyranose and furanose rings, which has a single acyl groups at R3′ position. Four acylsugar acyltransferases (SI-ASATs) were used to sequentially add acyl groups to specific positions on the sucrose core. Orders of 1-4 steps of reaction are color coded as green, red, purple and blue, respectively. B. Structure of tri-acylsucrose from S. pennellii LA0716 with three acylations on the pyranose and no furanose ring substitution. Three acylsugar acyltransferases (Sp-ASATs) participated in biosynthesis of the P-type acylsucroses via a "flipped pathway", where Sp-ASAT3 catalyzed production of di-acylsucrose, not tri-acylsucrose, in S. pennellii LA0716. Color codes for reaction order were the same as in A.*

acylsugar assembly by sequentially adding acyl groups onto specific hydroxyl groups on sucrose cores. The acylsucrose biosynthetic network in cultivated tomato (*Solanum lycopersicum*) is relatively simple, involving sequential actions of four *S. lycopersicum* acyltransferases (*Sl*ASATs) that acylate sucrose using various acyl-CoA substrates [28] (Figure 4.1a). *Sl*ASAT1 and *Sl*ASAT2 catalyze the first two steps of acylsugar biosynthesis, which decorate R4 and R3 positions on the 6-membered (pyranose) ring of sucrose, yielding di-acylsucrose intermediates. These, in turn, serve as substrates for *Sl*ASAT3 which adds a third acyl chain at the R3′ position on the 5-membered (furanose) ring to make tri-acylsucrose [29]. At last, *Sl*ASAT4 (formerly AT2) is responsible for tetra-acylsucrose biosynthesis by using acetyl-CoA to acetylate the R2 position in *S. lycopersicum* and *S. habrochaites* species (by *Sh*ASAT4) [30]. These enzymes, especially *Sl*ASAT2 and *Sl*ASAT3, produce diverse acylsucrose metabolites in cultivated tomato, due in part to their promiscuity toward using acyl-CoA substrates with different branching types and chain lengths [14, 28-30]. Moreover, diversification of ASAT variants plays an important role in shaping the structure diversity of acylsugars across plant genotypes. For example, a variant of *Sl*ASAT3 detected in *S. pennellii* (*Sp*ASAT3) uses mono-acylsucrose, rather than di-acylsucrose, as acyl acceptor to make di-acylated sugar esters, with small differences in amino acid sequence resulting in a “flipped biosynthetic pathway” that generates an unusual set of tri-acylsugars with only six-membered pyranose ring decoration [29] (Figure 4.1b).

Beyond the roles of sugar core acylation by ASATs, understanding of acylsugar chemical diversification still requires solving of another piece of the acylsugar biosynthetic puzzle. Metabolic networks of acyl-CoAs also contribute to this variation, yielding a wide range of molecules that serve as a rich pool of aliphatic acyl chain precursors of acylsugars. Radioisotope

feeding studies with ^{14}C- and ^{3}H- labeled amino acids indicated that branched-chain acyl chains of C4 to C12 were derived from branched-chain amino acids (BCAAs; Val, Leu and Ile) through either of two elongation mechanisms [34-36]. Additionally, a recent report described how mutations in the BCAA metabolism-related enzyme isopropylmalate synthase 3 (IPMS3) changed acyl chain composition in acylsugar metabolites from accessions of the wild tomato *Solanum pennellii* [37]. Previous studies of acylsugar biosynthesis mechanism in Solanaceous plants described two pathways of acyl chain elongation via various precursors [30-33] (Figure 4.2): 1. fatty acid synthase (FAS)-mediated acyl chain elongation, which generates the majority of straight chain and membrane fatty acids, involves a two-carbon elongation mechanism by adding two acetate carbon atoms per elongation cycle via acetyl carrier protein and 2. α-ketoacid elongation (αKAE) yields both straight and branched acyl chains. This latter mechanism involves condensation between acetyl CoA and an α-ketoacid precursor, followed by isomerization and decarboxylation to produce an α-ketoacyl CoA or acid elongated by one carbon. In the family *Solanaceae*, wild tomato species *S. pennellii* and *Datura metel* have been proposed to form acylsugar related acyl chains via FAS biosynthesis pathway, while αKAE mechanism is involved in acyl chain elongation in several Petunia and Tobacco species [31,32]. However, mutations are common in specialized metabolism genes, and often lead to functional in specialized metabolism within a species.

Identification and characterization of unknown specialized metabolites serves as an important prerequisite to functional genomic discoveries of novel gene and enzyme functions. In this study, a specific branched acyl chain derived from 3-methylpentanoic acid (aiC6), not previously detected in tomato or its wild relatives, was identified from metabolite profiling of *S. pennellii* accession LA2560. The structures of two aiC6-contained acylglucoses have been established

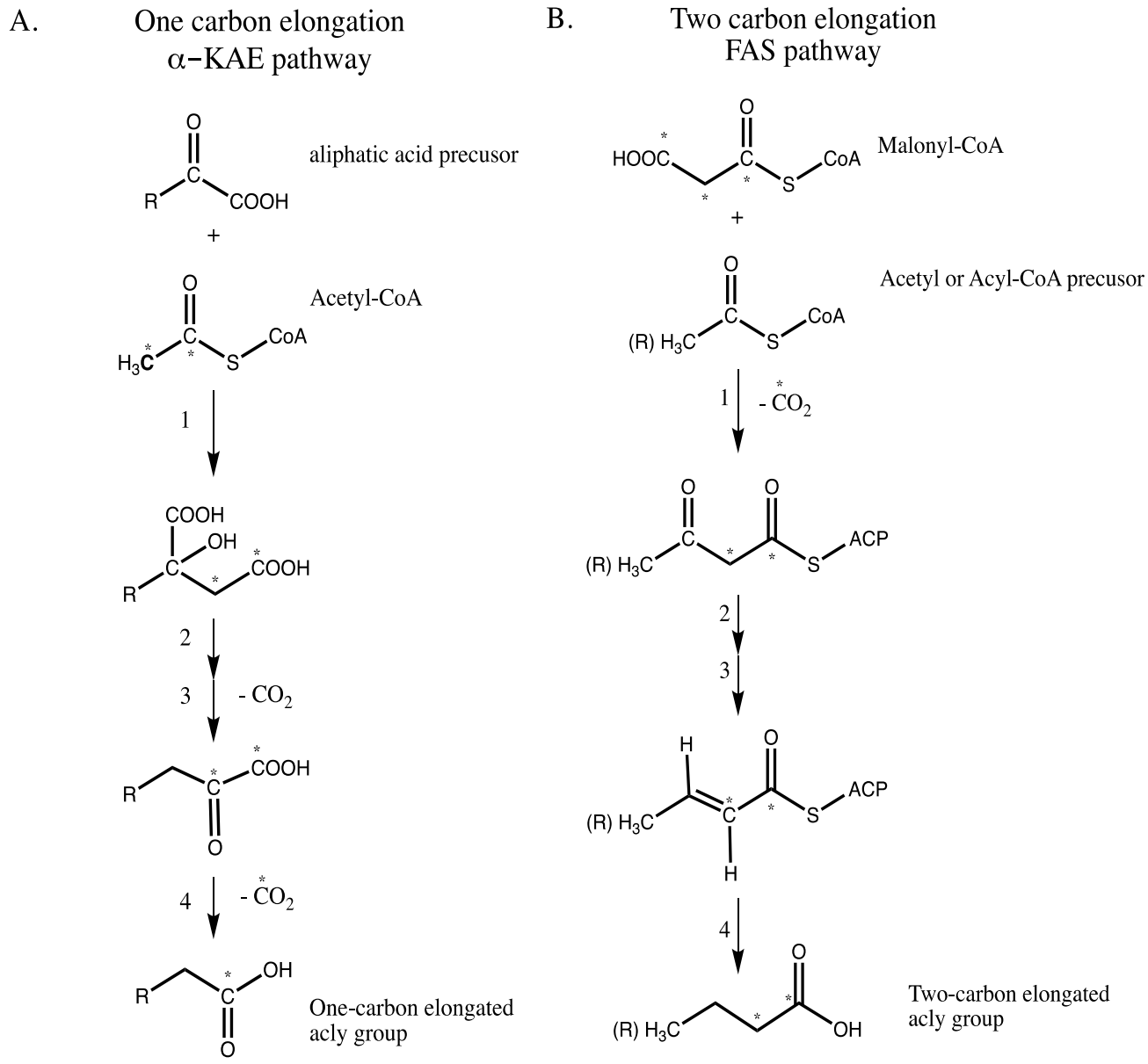


Figure 4.2. Diagram of aliphatic acid one-carbon and two-carbon elongation pathways*. A. One-carbon elongation (α-KAE) pathway. Carbon atoms originated from acetyl-CoA was labeled with stars and only one carbon will incorporated into product through this mechanism. Enzymes catalyzed 1-4 reactions assumed to be: 2-isopropylmalate synthase (IPMS), isopropylmalate dehydratase, 3-isopropylmalate dehydrogenase, 2-oxovalerate dehydrogenase, respectively; B. Two-carbon elongation (FAS) pathway. Carbon atoms originated from malonyl-CoA was labeled with stars, while two carbon atoms will incorporated into elongated acyl group. Enzymes involved in 1-4 reactions assumed to be: 3-ketoacyl-ACP synthase, 3-ketoacyl-ACP reductase, 3-hydroxyacyl ACP dehydratase, enoyl-ACP reductase, respectively.*

using LC-MS and NMR. However, anteiso-branching type of C6 suggested isoleucine should be used as precursor to produce this particular acyl group, but this elongation from Ile to aiC6 required participation of one-carbon elongation biosynthesis mechanism in *S. pennellii* species, which has not been evident from other *S. pennellii* acylsugar profiling according to previous literatures. To address this issue, wild tomato *S. pennellii* accessions LA2560 and LA0716 were grown with agar media containing uniformly labeled [$^{13}C_6$,^{15}N]Ile, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_5$,^{15}N]Val, which are putative precursors of branched acyl chains. Metabolite labeling was analyzed by liquid chromatography coupled with quadrupole time-of-flight tandem mass spectrometry (LC-QTOF-MS/MS). This analytical approach, coupled with *in vivo* labeling using ^{13}C-amino acids, was also applied to other plant species, providing an efficient and practical way to reveal biosynthetic mechanisms of branched aliphatic acids in the plant family *Solanaceae*, but also extends our understanding of the factors that drive complexity of acylsugar metabolism within and across plant species.

4. 2 Materials and methods

4.2.1 Materials and chemicals

HPLC grade methanol, acetonitrile, 2-propanol and ACS grade formic acid (88%, aqueous solution, J.T. Baker), and ammonium hydroxide (14.8 M, aqueous solution, EMD) were purchased from VWR Scientific. Chloroform-D (D, 99.8%), L-valine ($^{13}C_5$, 99%; ^{15}N, 99%), L-leucine ($^{13}C_6$, 99%; ^{15}N, 99%) and L-isoleucine ($^{13}C_6$, 99%; ^{15}N, 99%) were ordered from Cambridge Isotope laboratories, Inc. Hoagland nutrient solution at 1/4 strength was diluted from 1/2 X Hoagland solution provided by the Plant Growth Chamber Facility of Michigan State

University. Germicidal Bleach (Clorox) was obtained from the Biochemistry and Molecular Biology Research Store at Michigan State University.

Seed germination agar media (0.8 % agar, containing 3 μg/mL amino acid) was prepared by mixing 1.6 g agar with 200 mL MilliQ water. After autoclaving, 3 mL of specific BCAA stock solution (0.1 mg/mL, filtered using MILLEX GV filter unit, 0.22 μm) was added into liquid state media. The above solutions were shaken well and divided into 15 mL aliquots for each plant seed.

Tomato seeds (*S. pennellii* LA0716, LA2560 and *Solanum lycopersicum* M82) were provided by the UC Davis C. M. Rick Tomato Genetics Resource Center. Petunia seeds (*P. exserta, P. axillaris, P. integrifolia*) were obtained from Dr. Cornelius Barry at Department of Horticulture, Michigan State University. Growth conditions for 13C-BCAAs labeling experiments were described in section 4.2.4 and the growth conditions for *S. pennellii* accessions used for metabolites profiling and purification were described in next section 4.2.2.

4.2.2 UHPLC-MS profiling of acylsugar metabolites from *S. pennellii* LA0716 and LA2560

S. pennellii accessions LA0716 and LA2560 were grown from cuttings of mature plants in growth chamber with 17 hours daylight, 7 hours light-off and 86% humidity at 28 °C. After 2 weeks, ten terminal leaflets of each plant were harvested and extracted by dipping into 10 mL of isopropanol: acetonitrile: water (3:3:2, *v/v/v*) for 2 min. Three biological replicates were sampled for each accession. Extracts were concentrated by drying under vacuum with SpeedVac (no heat applied). Residues were redissolved in 500 μL acetonitrile: water (4:1, *v/v*). Extracts were vortexed for 30 s and centrifuged for 10 min (25 °C, 10000 g), followed by transferring 200 μL

supernatant to glass autosampler insert in LC vials. Telmisartan was added at 5 μM as internal standard to obtain accurate mass (lock mass at *m/z* 513.2296 in negative mode; *m/z* 515.2442 in positive mode). Extracts were stored at -20 °C until they were thawed for UHPLC/MS analyses.

Specialized metabolite profiling was performed using a Waters Xevo G2-S QTOF mass spectrometer coupled with Waters Acquity UHPLC pumps. Crude plant extracts were resolved using UHPLC separation on an Ascentis Express fused core silica C18 UHPLC column (10 cm×2.1 mm; 2.7 μm; Supelco, USA) at 40 °C, using aqueous 10 mM ammonium formate, pH 2.73, as mobile phase A and acetonitrile as mobile phase B. The total chromatographic run time was 110 min and flow rate was set at 0.3 mL/min. The linear gradient was: hold at 1% B for 1 min, linear gradient to 80% B over 1-100 min, followed by increase to 100% B during 100-101 min, hold at 100% B from 101-105 min, followed by return to 1% B over 105-106 min. The system was further equilibrated by a hold at 1% B from 106-110 min. A 10 μL aliquot was injected for each analysis. Mass spectra were acquired using ESI in both positive- and negative-ion mode with multiplexed CID. Mass resolution (full width half maximum) was 22000 and data acquisition was performed over *m/z* 50 to 1500 using dynamic range extension with 0.1 s acquisition time for each function. Five parallel functions were applied by switching the collision cell voltage between 10, 20, 40, 60, and 80 V. Several source parameters were: capillary voltage 2.14 kV in both negative mode and positive mode, sample cone voltage 35 V, desolvation temperature 280 °C, source temperature 90 °C, cone gas flow 0 L/h and desolvation gas flow 800 L/h. MassLynx software v4.1 (Waters Inc.) was used to monitor the instrument as well as process mass spectra data.

4.2.3 Isolation and structure elucidation of acylglucoses from *S. pennellii* LA2560

About 120 leaflets (12.4 g, fresh weight) of *S. pennellii* LA2560 plants were harvested and extracted in 200 mL HPLC grade methanol. The mixture was stirred with a glass rod for about 5 min, then quickly transferred into a 1L glass bottle through filter paper using a Buchner funnel. Rotary evaporation was used to remove all the solvent under vacuum. The residue was re-dissolved in 3 mL acetonitrile/ water (4/1, v/v) using ultrasonication for 10 min. Re-dissolved extract was centrifuged at 12857x*g*, 25°C for 5 min, and supernatant was collected and transferred to LC vials for purification.

Metabolite purification was performed using a Waters 2795 HPLC Separations Module and a Dionex Acclaim 120 C18 semi-preparative HPLC column (15 cm×4.6 mm; 5 μm). LKB fraction collector was used to collect eluted fractions in 1-min intervals during the 30-min gradient. An injection volume of 100 μL was used for each injection, and 20 injections were made and fractions pooled. A 30 min-gradient LC method was developed to purify acylglucoses from *S. pennellii* LA2560 crude extracts, using aqueous 0.15% formic acid aqueous solution as mobile phase A and acetonitrile as mobile phase B with 2 mL/min flow rate. The linear gradient method used was: linear gradient of 1-45% B over 0-3 min, hold at 45% B for 2 min (3-5 min), linear gradient to 52% B over 5-25 min, hold at 100% B for 3 min (26-28 min), return to 1% B at 29 min and hold 1 min (30 min). Two acylglucoses were isolated using this separation method, and composition of each fraction was checked using the LCT Premier mass spectrometer. Fractions containing purified acylsugars were evaporated to dryness under vacuum using SpeedVac. Acylglucoses G3:15(5,5,5) (fractions 14,15,16) and G3:16(5,5,6) (fractions 19,20,21) yielded white powders that were completely dissolved separately using 600 μL deuterated chloroform for NMR analysis.

To elucidate structures of purified *S. pennellii* LA2560 acylglucoses, one-dimensional (^{1}H and ^{13}C) and two-dimensional (HSQC, COSY, HMBC) NMR spectra were obtained using a 500 MHz (Agilent) NMR spectrometer at the Michigan State University Max T. Rogers NMR Facility. Deuterated chloroform ($CDCl_3$) was used as NMR solvent and chemical shifts reference (δ = 7.24 ppm for ^{1}H, δ = 77.0 ppm for ^{13}C) for individual acylsugar. Generally, ^{1}H and ^{13}C chemical shifts were resolved using 1D NMR data, whereas 2D NMR results were applied to annotate acyl chains onto different substitution positions along glucose ring.

4.2.4 ^{13}C-branched chain amino acids (BCAAs) labeling procedure

Seeds of tomato species *S. lycopersicum* M82, *S. pennellii* LA2560, LA0716, IL-7-4, IL-12-2 Petunia species *P. axillaris, P. exserta, P. integrifolia* were germinated in growth chamber for 27 days (21 days for tomato) at 24 hours light, 25 °C. Concentrations of ^{13}C-branched chain amino acids (^{13}C-Val, ^{13}C-Leu and ^{13}C-Ile) in media were 3 μg/mL. Unlabeled amino acid and ¼-strength Hoagland solutions were used as controls. Leaflets of each plant were sampled and extracted by dipping into 2 mL extraction solvent for 2 min. Extracts were concentrated under vacuum, and specialized metabolites were resolved by UHPLC separation, using acetonitrile and aqueous 10 mM ammonium formate, pH 2.71, as mobile phase components. Additional labeling procedures were presented in Section 3.2 of Chapter 3.

4.2.5 UHPLC-MS/MS analysis of stable isotope enriched acylsugar metabolites

Data dependent UHPLC-MS/MS analysis was carried out using both Waters Xevo G2-S and G2-XS QTOF mass spectrometers. Stable isotope labeled precursor ions with 4 or 5 Da upward mass shift of target acylsugars were selected in Function 1 MS (MS1) survey scan for MS/MS

fragmentation using collision cell potential ramp from 15-80 V. Mass range for MS1 scan was *m/z* 50-1500. UHPLC Chromatographic run time was 60 min for extracts of *S. pennellii* accessions LA2560 and LA0716 using a linear LC gradient method as: hold at 1% B for 1 min, linear gradient from 1-80% B over 1-54 min, hold at 100% B for 3 min from 54.01-57 min, return to 1% B at 57.01 min and hold 3 min to 60 min. UHPLC chromatographic run time was 60 min for *P. axillaris, P. exserta* and *P. integrifolia* with linear LC gradient method as: hold at 1% B for 1 min, linear gradients from 1-45% B over 1-5 min, 45-65% B over 5-50 min, to 100% B over 50-55 min, hold at 100% B from 3-58 min, return to 1% B at 58.01 min and hold 2 min until 60 min. UHPLC Chromatographic 110-min method for tomato M82 was described in Section 3.2 of Chapter 3. Aqueous 10 mM ammonium formate, pH 2.73, was used as mobile phase A and acetonitrile was used as mobile phase B for all methods mentioned above.

4.3 Results and discussion

Two accessions of *S. pennellii* LA0716 and LA2560 were subjected to the UHPLC-MS based deep profiling of acylsugar metabolites and chemical diversity analysis. In this study, twenty major acylglucoses (twelve in LA0716, eight from LA2560) and thirteen acylsucroses (four in LA0716, nine in LA2560), with over 80 isomers, were annotated from these two wild tomato species according to distinct retention times, pseudomolecular masses, and fragment ion masses. A summary of the acylsugar metabolite identities with mass to charge ratio, retention time and putative acyl chain composition is displayed in Table 4.1. Additionally, detailed structures of two acylglucoses purified from *S. pennellii* LA2560 were proposed according to LC-MS and NMR analysis, which provided evidence of the presence of 3-methylpentanoic acid esters (anteios-C6, abbrev. as aiC6) in acylglucoses. However, to our knowledge, biosynthesis of aiC6 in *S.*

pennellii species is not well explained using BCAA related two-carbon elongation mechanism of *S. pennellii* species that reported in previous literatures [31-34].

To address this issue, *S. pennellii* LA0716 and LA2560 were germinated with uniformly labeled [$^{13}C_5$,^{15}N]Valine, [$^{13}C_6$,^{15}N]Leucine and [$^{13}C_6$,^{15}N]Isoleucine, and incorporation of stable isotopes in acylsugar metabolites was analyzed using data dependent MS/MS. In the case of acylsugars containing aiC6 esters, this acyl group was labeled by growth on [$^{13}C_6$,^{15}N]Ile, which is consistent with both BCAA elongation and α-KAE mechanism were involved in aiC6 biosynthesis in *S. pennellii* LA2560 (Figure 4.3). Interestingly, evidence to suggest this pathway was not observed in the other *S. pennellii* accession, LA0716. Moreover, this labeling strategy was applied to other tomato species as well as petunia plants within the Solanaceae family, with the purpose of revealing diversification of acylsugar related aliphatic acid group biosynthesis between plant species. Results of these analyses were discussed in the section 4.4, 4.5, 4.6.

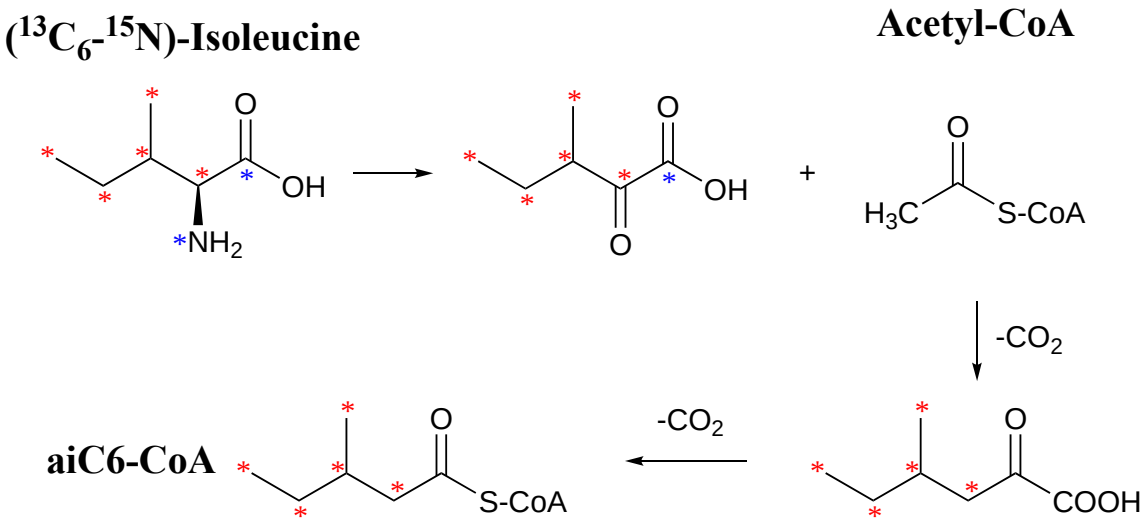


Figure 4.3. Proposed biosynthesis of aiC6 acyl chain through one-carbon (α-KAE) pathway using ^{13}C-Ile as precursor. *Stable isotopes incorporated into final aiC6-CoA were labeled with red stars*

4.3.1 Acylsugar profiling in *S. pennellii* LA0716 and LA2560

Over 90 chromatographic peaks corresponding to *S. pennellii* acylsugar metabolites were demonstrated using UHPLC-MS analysis and annotated based on retention times, as well as accurate pseudomolecular and fragment ion masses. Twenty-two distinct acylsugar masses were identified from *S. pennellii* LA0716 and LA2560, consisting of 9 acylsucroses and 13 acylglucoses (Figure 4.4). Interestingly, only tri-acylsucroses and tri-acylglucoses were resolved in the analyses, with all acyl groups substituted on the pyranose ring as judged from fragment ions generated in positive-ion mode using non-selective multiplexed CID. No tetra- or penta-acylsugars were detected in the two *S. pennellii* accessions. Additionally, more than 80 distinct acylsugar isomers were resolved using a 110-min LC gradient and extracted ion chromatograms. Molecular masses, retention times and predicted acyl group composition of annotated isomeric acylsugar metabolites from *S. pennellii* species are summarized in Table 4.1.

Table 4.1. *Acylsugars of S. pennellii LA0716 and LA2560 detected using UHPLC-MS*

***S. pennellii* LA0716 acylsugars**				
Acylsugar type	Proposed annotation	Retention time (min)	Theoretical *m/z*	Experimental *m/z*
Acylglucoses	G3:12 (4,4,4)	41.10	435.1886	435.186
		41.95	435.1886	435.1859
	G3:13 (4,4,5)	45.32	449.2023	449.2013
		45.99	449.2023	449.2015
		46.31	449.2023	449.2018
		46.84	449.2023	449.201
		47.07	449.2023	449.2004
	G3:14 (4,5,5)	49.79	463.2179	463.2183
		50.27	463.2179	463.2195
		50.74	463.2179	463.2171
		51.14	463.2179	463.2172
		51.45	463.2179	463.2173
		51.81	463.2179	463.216
	G3:15 (5,5,5)	54.63	477.2336	477.2309
		55.38	477.2336	477.2304
		56.00	477.2336	477.2345
		56.31	477.2336	477.235
	G3:16 (4,4,8)	61.10	491.2492	491.2502
		62.21	491.2492	491.2506
		63.09	491.2492	491.2502
	G3:17 (4,5,8)	64.93	505.2649	505.2677
		66.07	505.2649	505.2664
		67.13	505.2649	505.2668
		68.17	505.2649	505.269
	G3:18 (4,4,10)	70.84	519.2805	519.2802
		71.98	519.2805	519.2799
		73.12	519.2805	519.2809
	G3:19 (4,5,10)	74.59	533.2962	533.2964
		75.78	533.2962	533.2962
		76.9	533.2962	533.2959
	G3:20 (4,5,11)	78.57	547.3118	547.3114
		79.68	547.3118	547.3121
		80.83	547.3118	547.3119
	G3:20 (4,4,12)	81.9	547.3118	547.3124
		83.15	547.3118	547.3115

Table 4.1. (*cont'd)*

***S. pennellii* LA0716 acylsugars**				
Acylsugar type	Proposed annotation	Retention time (min)	Theoretical *m/z*	Experimental *m/z*
Acylglucoses	G3:21 (5,5,11)	84.63	561.3275	561.3309
	G3:21 (4,5,12)	85.63	561.3275	561.3282
		86.89	561.3275	561.3276
	G3:22 (5,5,12)	89.45	575.3431	575.3411
		90.64	575.3431	575.3428
Acylsucroses	S3:18 (4,4,10)	60.59	681.3334	681.3372
		61.52	681.3334	681.3381
	S3:19 (4,5,10)	63.69	695.349	695.3527
		64.62	695.349	695.3534
	S3:20 (4,4,12)	70.42	709.3647	709.3691
	S3:21 (4,5,12)	73.64	723.3803	723.3833
***S. pennellii* LA2560 acylsugars**				
Acylsugar type	Proposed annotation	Retention time (min)	Theoretical *m/z*	Experimental *m/z*
Acylglucoses	G3:14 (4,5,5)	51.55	463.2179	463.2195
		51.8	463.2179	463.2183
		52.13	463.2179	463.2198
	G3:15 (5,5,5)	55.16	477.2336	477.2333
		55.61	477.2336	477.2334
		56.29	477.2336	477.2337
		56.53	477.2336	477.233
	G3:16 (5,5,6)	59.86	491.2492	491.2504
		60.85	491.2492	491.2488
	G3:19 (5,5,9)	75.09	533.2962	533.2977
	G3:21 (5,5,11)	83.61	561.3275	561.3282
		84.73	561.3275	561.3283
	G3:21 (4,5,12)	85.67	561.3275	561.3314
		86.81	561.3275	561.3323
	G3:22 (5,5,12)	89.48	575.3431	575.3452
		90.64	575.3431	575.3441
	G3:23 (5,6,12)	93.14	589.3588	589.3617
		94.18	589.3588	589.3611
Acylsucroses	S3:15 (5,5,5)	46.77	639.2864	639.28
		47.19	639.2864	639.2803
	S3:16 (5,5,6)	50.25	653.3021	653.3065

Table 4.1. (*cont'd)*

***S. pennellii* LA2560 acylsugars**				
Acylsugar type	Proposed annotation	Retention time (min)	Theoretical *m/z*	Experimental *m/z*
Acylsucroses	S3:16 (5,5,6)	50.46	653.3021	653.3073
		50.63	653.3021	653.3091
		50.86	653.3021	653.3080
	S3:18 (5,5,8)	59.35	681.3334	681.3374
		59.67	681.3334	681.3378
		59.95	681.3334	681.3382
	S3:19 (5,5,9)	63.10	695.3490	695.3513
		63.33	695.3490	695.3528
	S3:19 (4,5,10)	64.81	695.3490	695.3517
	S3:20 (5,5,10)	66.56	709.3647	709.3701
		67.64	709.3647	709.3691
		68.58	709.3647	709.3692
	S3:21 (5,6,10)	70.92	723.3803	723.3864
	S3:21 (5,5,11)	72.01	723.3803	723.3844
	S3:21 (4,5,12)	73.73	723.3803	723.3849
	S3:22 (5,6,11)	75.29	737.396	737.4018
	S3:22 (5,5,12)	76.99	737.396	737.4013
		77.62	737.396	737.4019
	S3:23 (5,6,12)	80.97	751.4116	751.4176
	S3:24 (6,6,12)	84.38	765.4273	765.4336
		85.80	765.4273	765.4294
		86.94	765.4273	765.4266

Non-selective multiplexed CID spectra can provide important information that supports structure annotation of acylsugar metabolites, and can be used to narrow the possible structures of isomeric acylsugars. Annotation of acyl chains and relevant acylation positions may be proposed based on fragment ion masses of a specific acylsugar. In negative-ion mass spectra, compositions of acyl chains and esterification positions in acylsugar were suggested by fragment ions corresponding to neutral losses of acyl groups as ketenes (*e.g.* 84 Da for C5 ketenes, 140 Da for

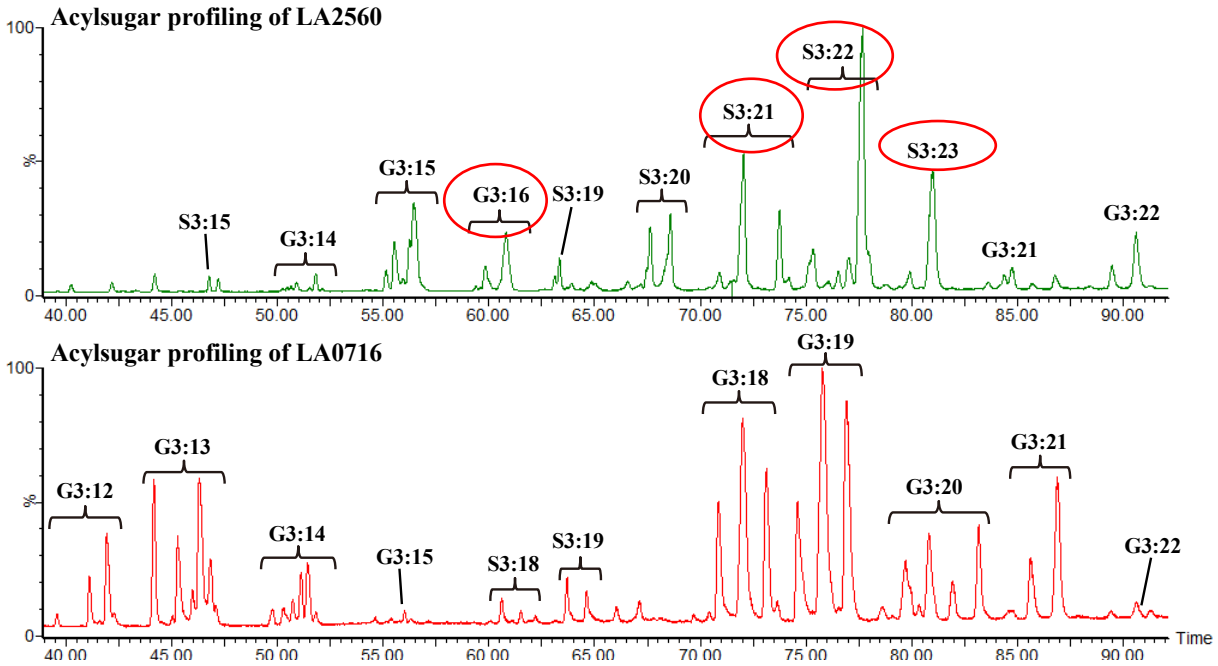


Figure 4.4. Acylsugar profiling of leaflet extracts from S. pennellii LA2560 and LA0716 using UHPLC/MS in negative-ion mode*. Total chromatographic time is 110 min. Dominant acylsugars are annotated. Acylsugars containing ai-C6 chain are circled on the LA2560 chromatogram*

C9 ketenes) formed at intermediate collision potentials (20, 40 and 60 V). Ions corresponding to deprotonated fatty acids (*e.g. m/z* 101 Da for C5 carboxylates and 157 Da for C9 carboxylates) were detected at the highest collision potential (80 V). In contrast, positive-ion mass spectra yielded fragment ions corresponding to cleavage of the glycosidic bond in acylsucroses as the collision voltage increased. The masses of key abundant fragment ions facilitated assignments of esters on either the pyranose or furanose rings. For *S. pennellii* acylsucroses, the most abundant fragment ion was often derived from retention of charge on the furanose ring, and its mass reflected the sum of acyl substitutions. Figure 4.5A illustrates how negative and positive mode mass spectra were applied to establish carbon number of acyl chains and their distribution on sucrose ring of S3:22 (5,5,12) from *S. pennellii* LA2560. In negative-ion CID mass spectra, formate adduct $[M+HCOO]^-$ (*m/z* 737) was found as the most abundant ion with the lowest

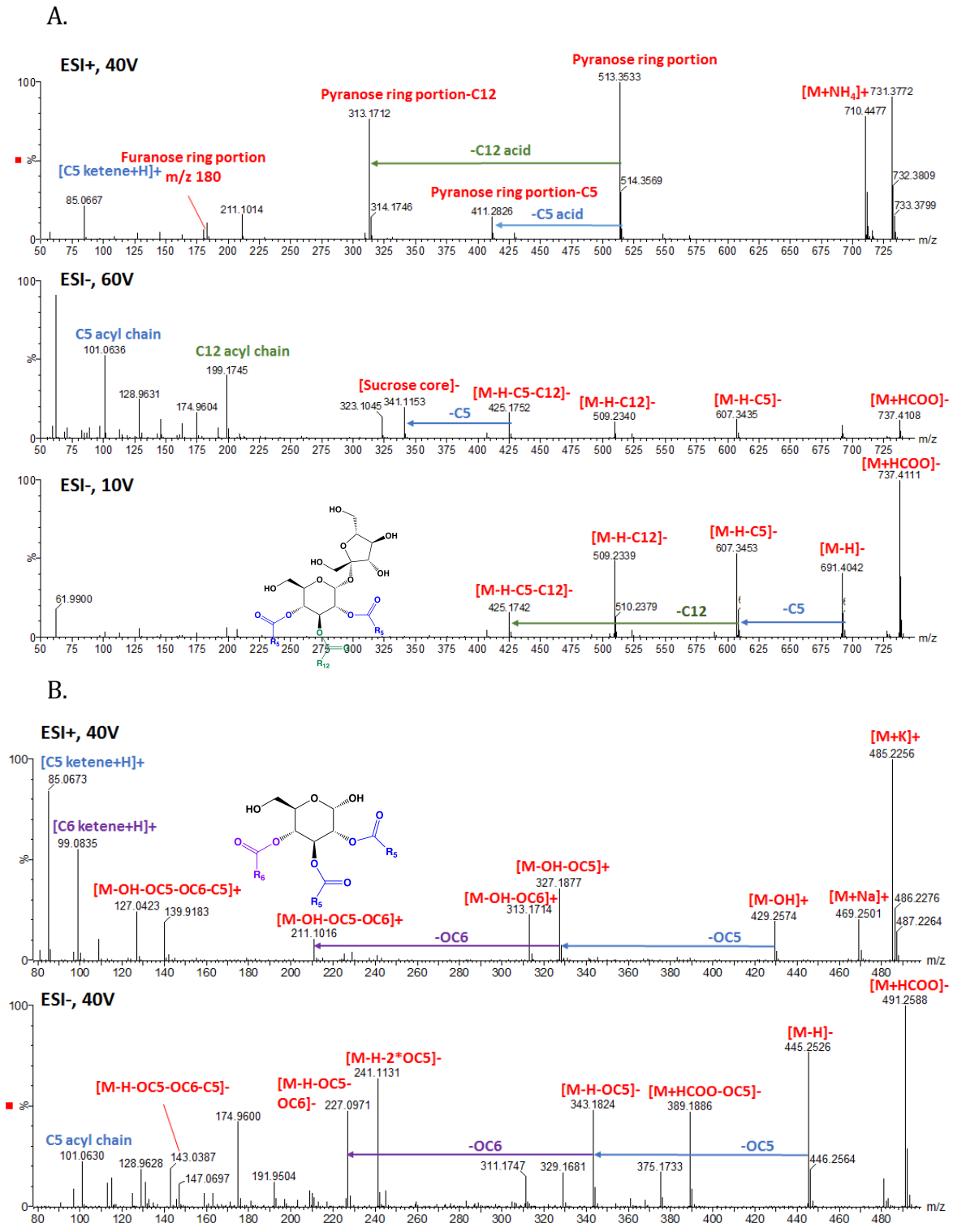


Figure 4.5 Putative assignments of pseudomolecular and fragment ions of acylsucrose S3:22 and acylglucose G3:16 from S. pennellii LA2560*. Mass spectra acquired in both negative and positive mode using multiplexed CID. A. CID spectra of S3:22 (5,5,12) at 10 and 60 V collision potential in negative-ion mode, and 40 V in positive-ion mode. (RT=77.62 min); B. CID spectra of G3:16 (5,5,6) at 10 and 60 V collision potential in negative mode and 40 V in positive mode. (RT=60.85 min).*

collision energy at 10 V. By increasing the CID potential up to 60 V, fragments corresponding to neutral losses of ketenes, including $[M-H-C5]^-$ (*m/z* 607), $[M-H-C12]^-$ (*m/z* 509) and $[M-H-C5-C12]^-$ (*m/z* 425, also corresponding to $[sucrose+C5]^-$ residue) were observed, as well as fragments corresponding to the deprotonated sucrose core (*m/z* 341), aliphatic acid C5 carboxylate (*m/z* 101) and C12 carboxylate (*m/z* 199), supporting assignment of acyl groups of S3:22 as two C5 and one C12. In positive-ion CID mass spectra, high abundance fragment ion $[M-furanose]^+$ (*m/z* 513) was observed at a collision potential of 40 V, consistent with substitution of three acyl esters on the 6-membered ring and no aliphatic chains attached to the furanose ring.

In contrast to acylsucrose behavior, fragment ions of acylglucoses from *S. pennellii* were formed by neutral loss of carboxylic acids (*e.g.* 102 Da for C5, 116 Da for C6) instead of ketenes (*e.g.* 84 Da or 98 for C5 or C6) in both positive and negative modes (Figure 4.5B). The above results show that information of acylsugar type, carbon numbers of acyl groups and substitutions on each sugar ring are revealed by mass spectra generated by multiplexed CID in both positive and negative modes. However, mass spectra have yet to distinguish the exact substitution position of each acyl group or distinguish acyl chain branching or the α- or β- configuration of the 1-position hydroxyl of acylglucoses. For chromatographically-resolved isomers, multiplexed CID mass spectra have limited utility for distinguishing their structures because the isomers share the same fragment ion masses. Therefore, one-dimensional and two-dimensional NMR (1D/2D NMR) techniques were employed for more complete structure assignments.

4.3.2 Structure elucidation of acylglucose metabolites purified from *S. pennellii* LA2560

Compared with acylsucrose metabolites [14-18], few acylglucose structures have been reported in the literature [11-13], which has limited our understanding of acylsugar chemical diversity as

well as biosynthetic pathways leading to acylglucose metabolites. In this investigation, milligram quantities of two tri-acylglucoses G3:15(5,5,5) and G3:16(5,5,6) were purified from *S. pennellii* LA2560, and structures are proposed based on 1D (^{1}H and ^{13}C) and 2D (COSY, HSQC and HMBC) NMR spectra (Table 4.2).

Similarly, as was the case for ^{1}H-NMR analyses of acylsucrose metabolites, there are three distinct chemical shift regions for proton resonances in ^{1}H spectra of purified acylglucoses: 1) 0.8~1.9 ppm (□) is the region for aliphatic H from acyl groups; 2) 2.1~2.5 ppm is the region for α-protons adjacent to ester group carbonyls; 3) 3.5~5.7 ppm is the region for protons on the pyranose ring. Protons attached to carbons in glucose rings with ester groups attached fell in 4.8~5.7 ppm, and 3.5~4.6 ppm is the region for protons on sugar ring carbons lacking attached ester groups. Acylations cause a downfield shift in resonances for protons attached at the esterified carbon positions, and this information was used to help assign substituted positions. ^{1}H and COSY spectra were applied to assign protons on glucose carbons based on ^{1}H-^{1}H coupling, starting with the downfield doublet peaks (δ 5.5 ppm) that correspond to the 1-position protons on the glucose ring. ^{13}C and HSQC spectra were used to not only assign carbon atoms on glucose core and for acyl chains, but also helped to distinguish overlapping signals of ring Hs that

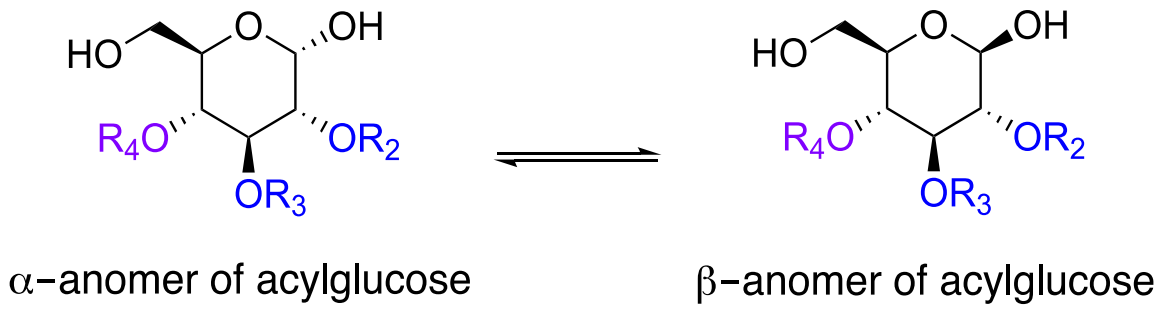


Figure 4.6. Rapid equilibrium between α- and β-anomers of acylglucose.

exhibited distinct carbon chemical shifts. Substitution positions for each acyl chain on the sucrose core were recognized using HMBC spectra, which showed correlations between H and C nuclei that are 2-4 bonds away from each other. Carbonyl carbons (δ 170~180 ppm) generated cross peaks with both glucose ring protons and α-position protons on aliphatic acids to demonstrate specific esterification positions for individual acyl groups. According to NMR data, the two acylglucoses G3:15 and G3:16 presented very similar chemical shift and splitting

Table 4.2. NMR chemical shifts and structures of purified acylglucoses G3:15 and G3:16 from S. pennellii LA2560

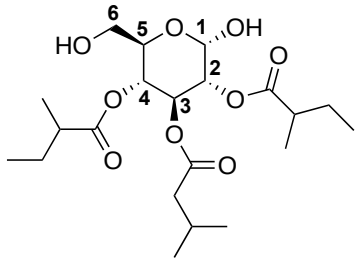
	**α-G3:15 (5,5,5)** **Purified from *S. pennellii* LA2560** **HRMS: (ESI) *m/z* calculated for $C_{22}H_{37}O_{11}^-$ ([M+HCOO$^-$]): 477.2336, found: 477.2339** **Material recovered: ~1-2 mg** **NMR solvent: CDCl3**	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.52 (d, *J*=3.7 Hz)	90.26 ($^{1}J_{CH}$ = 175.0 Hz)
2 (CH)	4.88 (dd, *J*=3.7, 10.0 Hz)	71.33
2-O-	–	–
-1 (CO)	–	177.04
-2 (CH)	2.40 (m)	40.86
-2′ (CH_3)	1.13 (d, *J*=8.2 Hz)	16.44
-3 (CH_2)	1.45 (m), 1.65 (m)	26.40
-4 (CH_3)	0.88 (m)	11.48
3 (CH)	5.67 (t, *J*=10.0 Hz)	68.78
3-O-	–	–
-1 (CO)	–	171.84
-2 (CH_2)	2.20 (m)	43.08
-3 (CH)	2.06 (m)	25.47
-4 (CH_3) x 2	0.94 (d, *J*=7.5 Hz)	22.32
4 (CH)	5.03 (t, *J*=10.0 Hz)	68.60
4-O-	–	–
-1 (CO)	–	175.79
-2 (CH)	2.40 (m)	41.15
-2′ (CH_3)	1.15 (d, *J*=8.2 Hz)	16.44
-3 (CH_2)	1.45 (m), 1.65 (m)	26.40
-4 (CH_3)	0.88 (m)	11.48
5 (CH)	4.06 (m)	69.54
6 (CH_2)	3.74 (m), 3.59 (m)	61.17

Table 4.2. (cont'd)

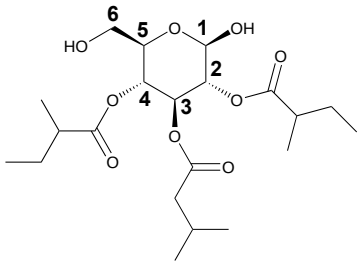
	**β-G3:15 (5,5,5)** **Purified from *S. pennellii* LA2560** **HRMS:** (ESI) *m/z* calculated for $C_{22}H_{37}O_{11}^-$ ([M+HCOO$^-$]): 477.2336, found: 477.2339 **Material recovered:** 1~2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	4.74 (d, *J*=9.5 Hz)	95.86 ($^1J_{CH}$ = 163.0 Hz)
2 (CH)	4.91 (m)	73.49
2-O-	–	–
-1 (CO)	–	177.04
-2 (CH)	2.40 (m)	40.86
-2' (CH_3)	1.13 (d, *J*=8.2 Hz)	16.44
-3 (CH_2)	1.45 (m), 1.65 (m)	26.40
-4 (CH_3)	0.88 (m)	11.48
3 (CH)	5.37 (t, *J*=10.0 Hz)	71.19
3-O-	–	–
-1 (CO)	–	171.85
-2 (CH_2)	2.20 (m)	43.08
-3 (CH)	2.06 (m)	25.47
-4 (CH_3) x 2	0.94 (d, *J*=7.5 Hz)	22.32
4 (CH)	5.06 (t, *J*=10.0 Hz)	68.60
4-O-	–	–
-1 (CO)	–	172.45
-2 (CH)	2.40 (m)	41.15
-2' (CH_3)	1.15 (d, *J*=8.2 Hz)	16.44
-3 (CH_2)	1.45 (m), 1.65 (m)	26.40
-4 (CH_3)	0.88 (m)	11.48
5 (CH)	3.56 (m)	74.59
6 (CH_2)	3.74 (m), 3.59 (m)	61.17

Table 4.2. (cont'd)

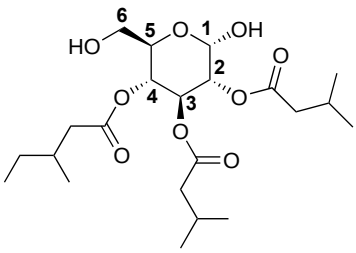
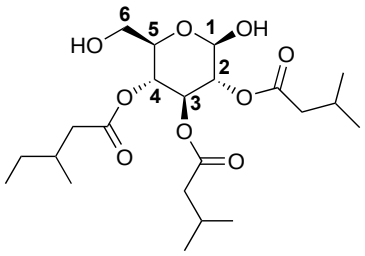
	**α-G3:16 (5,5,6)** **Purified from *S. pennellii* LA2560** **HRMS:** (ESI) *m/z* calculated for $C_{23}H_{39}O_{11}^-$ ([M+HCOO$^-$]): 491.2492, found: 491.2486 **Material recovered:** ~1-2 mg **NMR solvent:** $CDCl_3$	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	5.52 (d, *J*=3.3 Hz)	90.23
2 (CH)	4.85 (dd, *J*=3.3, 10.0 Hz)	71.38
3-O-	–	–
-1 (CO)	–	172.25
-2 (CH_2)	2.15 (m)	43.05
-3 (CH)	2.07 (m)	25.47
-4 (CH_3) x 2	0.94 (m)	22.31
3 (CH)	5.67 (t, *J*=10.0 Hz)	68.80
3-O-	–	–
-1 (CO)	–	171.85
-2 (CH_2)	2.15 (m)	43.05
-3 (CH)	2.07 (m)	25.47
-4 (CH_3) x 2	0.94 (m)	22.31
4 (CH)	5.03 (t, *J*=10.0 Hz)	68.61
4-O-	–	–
-1 (CO)	–	172.91
-2 (CH_2)	2.35 (m), 2.08[a] (m)	41.14
-3 (CH)	1.86 (m)	31.71
-3′ (CH_3)	0.95 (m)	19.22
-4 (CH_2)	1.34 (m), 1.22 (m)	29.19
-5 (CH_3)	0.87[a] (m)	11.52
5 (CH)	4.06 (m)	69.54
6 (CH_2)	3.74 (dd, *J*=3.9, 12.5 Hz), 3.59 (dd, *J*=3.9, 12.5 Hz)	61.12
a. Determined by HMBC		

Table 4.2. (cont'd)

	**β-G3:16 (5,5,6)** **Purified from *S. pennellii* LA2560** **HRMS:** (ESI) *m/z* calculated for $C_{23}H_{39}O_{11}^-$ ([M+HCOO$^-$]): 491.2492, found: 491.2486 **Material recovered:** 1~2 mg	
Carbon # (group)	^{1}H (ppm)	^{13}C (ppm)
1 (CH)	4.74 (d, *J*=8.0 Hz)	95.89
2 (CH)	4.88 (m)	73.47
3-O-	–	–
-1 (CO)	–	172.6
-2 (CH_2)	2.15 (m)	43.05
-3 (CH)	2.07 (m)	25.47
-4 (CH_3) x 2	0.94 (m)	22.31
3 (CH)	5.38 (t, *J*=10.0 Hz)	71.19
3-O-	–	–
-1 (CO)	–	171.85
-2 (CH_2)	2.15 (m)	43.05
-3 (CH)	2.07 (m)	25.47
-4 (CH_3) x 2	0.94 (m)	22.31
4 (CH)	5.06 (t, *J*=10.0 Hz)	68.61
4-O-	–	–
-1 (CO)	–	172.91
-2 (CH_2)	2.35 (m), 2.08[a] (m)	41.14
-3 (CH)	1.86 (m)	31.71
-3' (CH_3)	0.95 (m)	19.22
-4 (CH_2)	1.34 (m), 1.22 (m)	29.19
-5 (CH_3)	0.87[a] (m)	11.52
5 (CH)	3.56 (m)	69.54
6 (CH_2)	3.74 (dd, *J*=3.9, 12.5 Hz), 3.59 (dd, *J*=3.9, 12.5 Hz)	61.12

patterns in the glucose ring region (3.5~5.7 ppm in H, 63~100 ppm in C), which suggested that these two purified acylglucoses are esterified on the same positions.

Due to the rapid equilibration between α and β anomers of acylglucose (Figure 4.6), all purified acylglucoses consisted of mixtures of anomers by the time NMR spectra were generated, and the more complex NMR data brought unexpected challenges when elucidating structures of acylglucoses, especially during assignment of Hs on glucose ring. The two most downfield doublet peaks were annotated as H1 peaks, where α-H1s were differentiated from β-H1s according to larger chemical shifts, ^{1}H-^{13}C coupling constants, and smaller ^{13}C chemical shifts and ^{1}H-^{1}H coupling constants. For example, G3:15 from *S. pennellii* LA2560 exhibited two doublet peaks at 5.52 (d, *J*=3.7 Hz), 4.74 ppm (d, *J*=9.5 Hz), which represented α- and β-anomeric protons respectively. Coupled HSQC spectra showed the heteronuclear (^{1}H-^{13}C) coupling constant was 175 Hz for α-anomer and 163 Hz for β-anomer (Table 4.2).

Interestingly, an acyl group assigned as 3-methylpentanoic acid (anteiso-C6 or aiC6) was identified in the NMR analysis of *S. pennellii* LA2560 acylglucose G3:16(5,5,6), which has not been described in earlier reports of acylsugar structures of wild tomato species [14-18]. 2D NMR data, especially the HMBC spectrum, provided clear evidence for the anteiso-C6 group. Figure 4.7 displays the HMBC spectra of G3:16 associated with aiC6 chain. Correlations between most Hs and Cs (except carbonyl carbon at δ 172.91 ppm) within 2-4 bonds were detected from these HMBC spectra. α-C (labeled as C2 in figure 4.7) of aiC6 acyl chain (δ 41.14 ppm) was correlated with protons attached to C3 (CH, δ 1.86 ppm), C3' (CH_3, δ 0.95 ppm) and C4 (CH_2, δ 1.34, 1.22 ppm), but was not correlated with Hs attached to C5 (CH_3, δ 0.87 ppm). This result

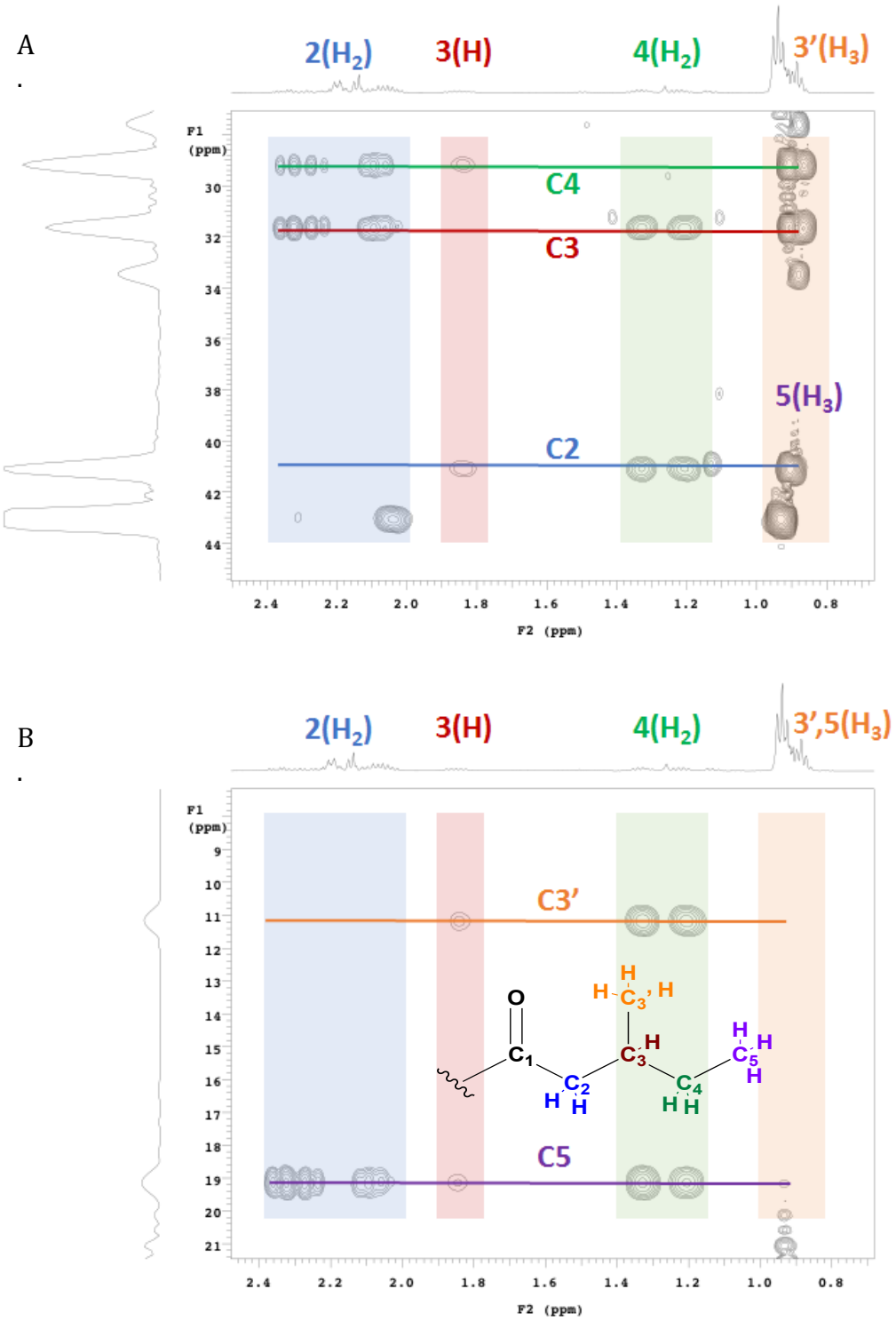


Figure 4.7. HMBC spectra of aiC6 acyl chain of G3:16 (5,5,6) in LA2560. *A. HMBC spectrum showing correlations between C2, C3, C4 of aiC6 vs. Hs within 2-4 bonds. B. HMBC spectrum of correlations between two methyl carbons C3′ and C5 of aiC6 vs. Hs within 2-4 bonds.*

could be used as evidence to distinguish Hs between 3′(CH_3) and 5(CH_3) that had very similar chemical shifts around δ 1ppm. Furthermore, chemical shift of β-C (C3, δ 31.71 ppm) confirmed the presence of aiC6 groups, which showed correlations with Hs of positions 2(CH_2, δ 2.35, 2.08 ppm), 4(CH_2), 3′(CH_3) and 5(CH_3, δ 0.87 ppm). Assignments of protons on the aiC6 ester groups and their attached Cs were confirmed from HSQC spectra.

4.3.3 Comparison of acylsugar profiles and structure diversity within two *S. pennellii* accessions

Profiling of acylsugar metabolites from *S. pennellii* LA0716 and LA2560 showed notable differences between accessions, especially regarding to the ratios of acylglucoses to acylsucroses. Most acylsugar metabolites identified from accession LA0716 were acylglucoses , while acylsucroses were the dominating metabolites detected in LA2560. In addition, only four acylsucroses were detected among sixteen major LA0716 acylsugars according to molecular masses and chromatographic retention times (Table 4.1), with acylsucroses accounting for approximately 4% of total acylsugar LC/MS peak area for this accession. In contrast, acylsucrose levels in accession LA2560 were about 75% of total acylsugars, and correspond to eleven of the seventeen major acylsugars masses detected in this accession. Furthermore, the total acylsugar level in accession LA0716 was about 1.7-fold higher than LA2560 based on integrated extracted ion chromatograms from UHPLC/MS profiles. These results were consistent with previous *S. pennellii* acylsugar characterization that explored geographic distribution of the species, in that the southern accession (LA0716) produced more acylsugars, as well as more acylglucoses than acylsucroses, than the northern accession LA2560 [38].

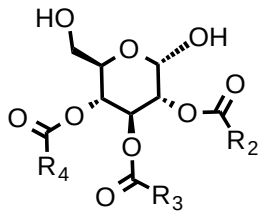
All acylsucrose and acylglucose metabolites identified in *S. pennellii* were tri-substituted sugar esters, and showed diversity of acyl chain composition. Masses of pseudomolecular and fragment ions obtained by CID spectra demonstrated that acyl groups ranging from C4 to C12 were present in both LA0716 and LA2560 accessions. However, no C7, C8 acyl groups were detected in either accession. No acyl chain longer than 12 carbons was observed, which was consistent with our previous profiling of other *Solanum* species [14]. No acetyl group was detected in any *S. pennellii* acylsugar, in accord with the identification and documented substrate selectivity of *SpASATs* [28,30], as well as the lack of orthologs in *S. pennellii* of the tomato enzyme *SlASAT4*, known to acetylate the 2-position hydroxyl.

The structural diversity of acyl groups and the position conservation of specific acyl chains were demonstrated using NMR spectra. A total of 11 acylglucoses and 7 acylsucroses were purified and structures are summarized in Table 4.3 (NMR spectra of 16 of the 18 acylsugars were analyzed by Dr. Banibrata Ghosh). Differing from cultivated tomato M82 and wild tomato species including *S. habrochaites*, only 2,3,4-substituted acylsugars were detected in *S. pennellii* accessions, and mass and NMR spectra yielded no evidence for acylation on any furanose ring hydroxyls of acylsucroses. The dominating acyl groups were iC4 and aiC5 in acylsugars of LA0716, while in LA2560 the major ester groups were iC5 and aiC5. Anteiso-C6 chains were only discovered in acylsugars from accession LA2560. In all cases where NMR spectra were available, the R4 positions of both acylglucoses and acylsucroses in LA0716 were conserved by esterification by iC4, while the R4 position was mostly occupied by aiC6 in accession LA2560 (except three acylsugars not containing C6 esters: G3:15, G3:22 and S3:20, all of which had C5 esters at R4). The greatest diversity of acylation groups at specific positions on sugar core was observed at the 3-position on acylsucrose pyranose rings in the two *S. pennellii* species, which

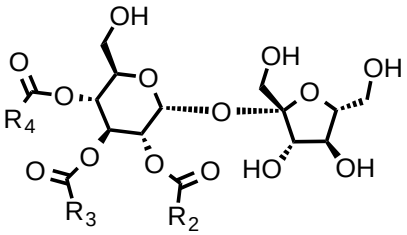
included acyl chains ranging from iC4 to nC12, but none showed evidence of anteiso-branching groups. Substituent diversity was also observed in R2 positions of acylsugars, with either iC4 or aiC5 detected at this site in LA0716, as iC5, aiC5 and one aiC6 (S3:24(6,6,12)) detected at R2 in LA2560. Moreover, the longer acyl esters (C10, C11 and C12) were only observed at the 3-position, while the preferred attachment of the shorter acyl chains (C4, C5 and C6) was at positions 2 and 4 of the pyranose ring. These observations were consistent with activities of *SpASAT*s using various acyl-CoA as substrates [28,30].

Despite the identification of acylsucrose acyltransferases in *S. pennellii*, some enzymes and pathways involved in acylglucose biosynthesis remained unclear. Interestingly, the substitution

Table 4.3. NMR elucidated structures of acylsugars purified from S. pennellii LA2560 and LA0716.



Acylglucose



Acylsucrose

LA2560				LA0716			
	R2	R3	R4		R2	R3	R4
G3:15	aiC5	iC5	aiC5	**G3:12**	iC4	iC4	iC4
G3:16	iC5	iC5	**aiC6**	**G3:13**	aiC5	iC4	iC4
G3:22	iC5	nC12	iC5	**G3:14**	aiC5	iC5	iC4
G3:23	iC5	nC12	**aiC6**	**G3:18**	iC4	iC10	iC4
S3:21	iC5	iC10	**aiC6**	**G3:19 [1]**	aiC5	iC10	iC4
S3:22	iC5	iC11	**aiC6**	**G3:19 [2]**	aiC5	nC10	iC4
S3:23	iC5	nC12	**aiC6**	**S3:18**	iC4	iC10	iC4
S3:24	**aiC6**	nC12	**aiC6**	**S3:19**	aiC5	iC10	iC4

positions in both acylglucoses and acylsucroses were consistent, with exactly the same acyl groups attached to the 2-, 3-, and 4-positions of the pyranose ring were observed in metabolites from both *S. pennellii* accessions. For example: both G3:18 ($iC4^{R2}$, $iC4^{R4}$, $iC10^{R3}$) and S3:18 ($iC4^{R2}$, $iC4^{R4}$, $iC10^{R3}$) from accession LA0716 had two iC4 groups attached at 2- and 4-positions, and iC10 acylated at the 3-position of the pyranose ring. These observations lend support to a hypothesis that sucrose derivatives are substrates for *S. pennellii* acyltransferases, and acylglucoses are formed by hydrolytic removal of the furanose ring. This suggestion stands in contrast to a proposed mechanism for acylglucose biosynthesis that proposed that formation of a 1-*O*-acylglucose was the initial intermediate that underwent subsequent intramolecular transesterification reactions that allowed acyl groups to migrate to other hydroxyl positions [39,40]. However, this current investigation found no evidence from either UHPLC-MS profiling or NMR spectra of R1 acylation.

4.3.4 Investigation of acylsugar related aliphatic acid elongation in wild tomato *S. pennellii* accessions

Investigation of ^{13}C-labeling patterns in *S. pennellii* LA2560 and LA0716 acylsugar acyl groups derived from branched chain amino acids provided evidence to suggest that both one-carbon and two-carbon elongation pathways are involved in acylsugar biosynthesis in accession LA2560 (Figure 4.9), while LA0716 esters were consistent with only two-carbon elongation reactions (Figure 4.11). Heavy isotope labeling patterns were analyzed using both multiplexed CID and MS/MS, in which precursor ions $[M+formate+4]^-$, $[M+formate+5]^-$ were selected as acylsugars labeled by $[^{13}C_5,^{15}N]$Val and $[^{13}C_6,^{15}N]$Ile or $[^{13}C_6,^{15}N]$Leu respectively.

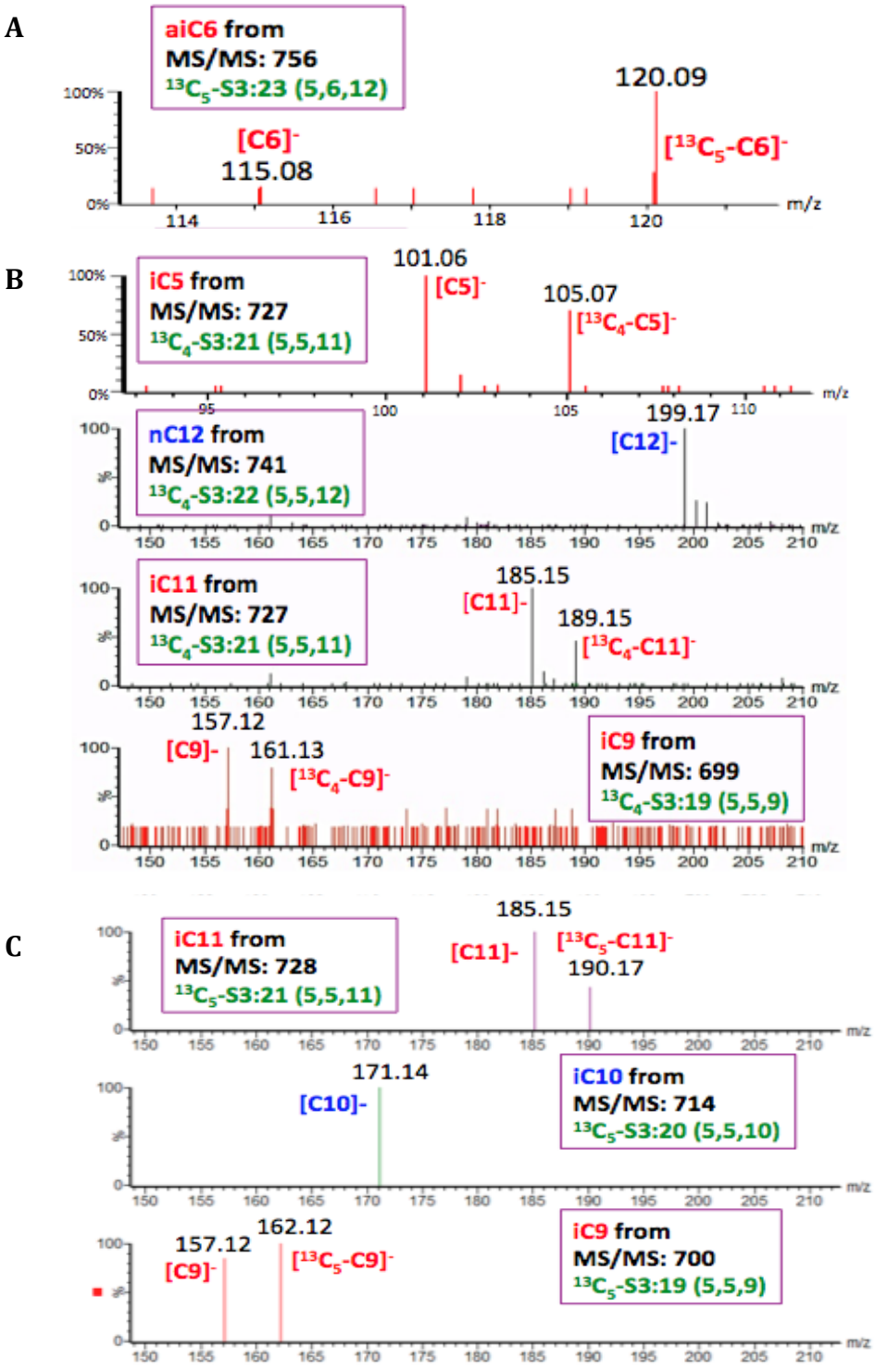


Figure 4.8. MS/MS spectra of aliphatic acyl groups labeled with ^{13}C-BCAAs in S. pennellii LA2560*. A. Product ion spectra of heavy C6 acyl chains in [$^{13}C_6$,^{15}N]Ile labeled acylsugars; B. Product ion spectra showing heavy C5, C9, C11 and light C12 acyl chain fragments in [$^{13}C_5$,^{15}N]Val labeled acylsugars; C. product ion spectra showing heavy C9, C11 and light C10 in [$^{13}C_6$,^{15}N]Leu labeled acylsugars from accession LA2560 (top to bottom).*

Evidence for one-carbon elongation in LA2560 came from incorporation of five heavy isotopes from [$^{13}C_6$,^{15}N]Ile in the aiC6 acylsugars and four ^{13}C from [$^{13}C_5$,^{15}N]Val in iC5 acyl chains (Figure 4.8). Stable isotopes originating from [$^{13}C_6$,^{15}N]Ile were only incorporated into aiC5 ([$^{13}C_5$-C5]$^-$, *m/z* 106) and aiC6 ([$^{13}C_5$-C6]$^-$, *m/z* 120), and not detected in any acyl long chains (C9-C12), which was consistent with the NMR result that no anteiso-acyl long chains were observed in LA2560 acylsugars. Furthermore, in this *S. pennellii* accession, heavy isotopes from [$^{13}C_5$,^{15}N]Val were observed in the same odd-numbered carbon iso-long chains iC9 ([$^{13}C_4$-C9]$^-$,

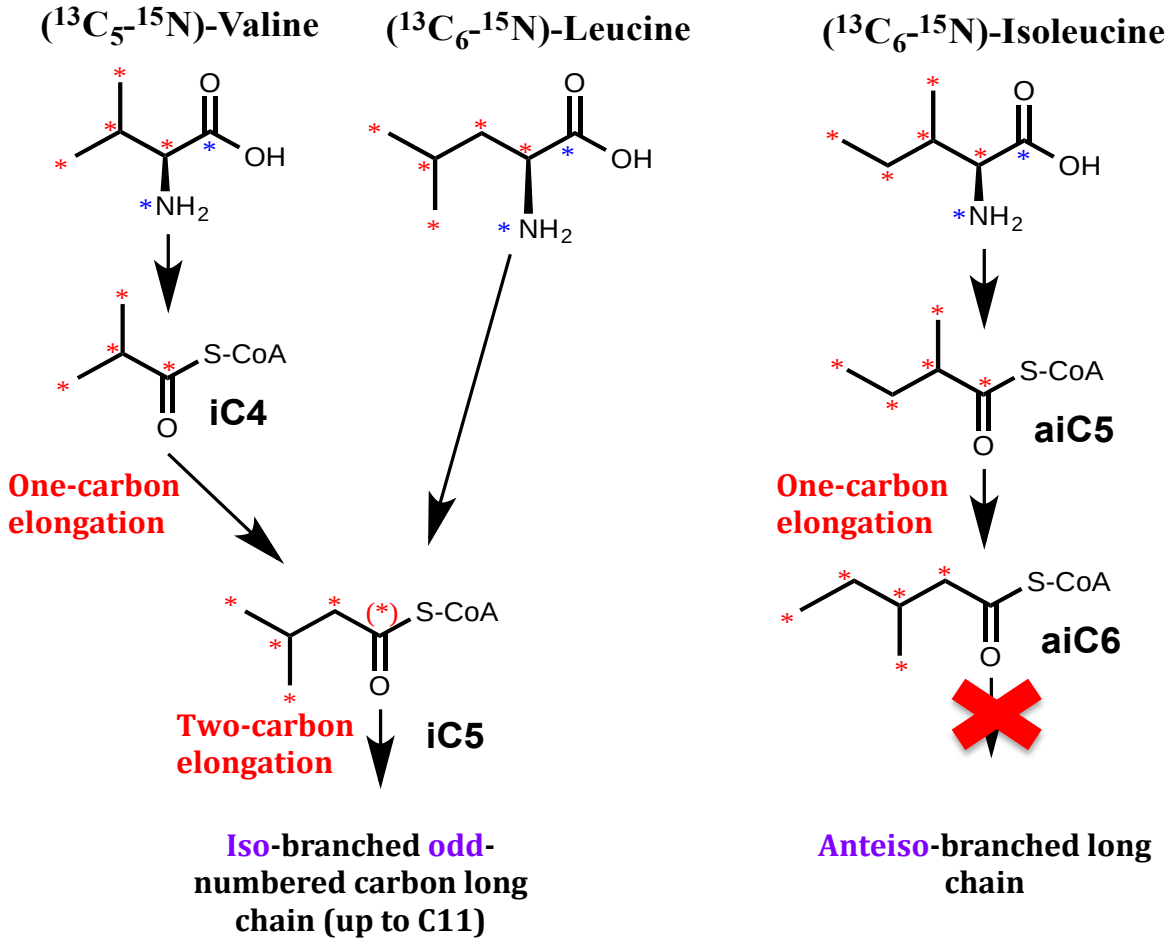


Figure 4.9. Proposed elongation pathway of aliphatic acyl chain with different branching types in S. pennellii LA2560. *Heavy isotopic atoms were labeled with stars. Red star labeled atoms were incorporated into acyl groups.*

m/z 161) and iC11 ([$^{13}C_4$-C11]$^-$, *m/z* 189), which was identical to results obtained using [$^{13}C_6$,^{15}N]Leu, providing evidence that these odd-numbered iso-branching long acyl chains were synthesized from iC5 through two-carbon elongation, meanwhile valine-derived intermediates were elongated by one-carbon elongation mechanism to make iC5, which was them elongated via two-carbon additions to yield the odd-carbon iC9 and iC11 chains. However, this study still did not yield conclusive evidence of pathways yielding the even-numbered acyl long chain iC10, nC10 and nC12 since none of these acyl groups shown detectable 13C enrichment derived from any ^{13}C-BCAAs precursors.

Two-carbon elongation was consistent with the major biosynthesis pathway involved in making aliphatic chains in *S. pennellii* LA0716. Unlike in LA2560, because neither heavy aiC6 nor iC5 was found from [$^{13}C_6$,^{15}N]Ile or [$^{13}C_5$,^{15}N]Val labeled samples, no evidence for one-carbon elongation was observed for this accession. Evidence for two-carbon elongation in accession LA0716 was obtained from incorporation of five heavy isotopes from [$^{13}C_6$,^{15}N]Leu in the odd-numbered iso-branched acyl chains including iC9, iC11 and four ^{13}C from [$^{13}C_5$,^{15}N]Val in iC10, the latter of which represents even-numbered iso-aliphatic chains (Figure 4.10). Similar as in accession LA2560, stable isotopes originating from [$^{13}C_6$,^{15}N]Ile were only incorporated into aiC5 ([$^{13}C_5$-C5]$^-$, *m/z* 106) while no longer heavy anteiso-acyl chains (C6-C12) were observed in acylsugars from accession LA0716.

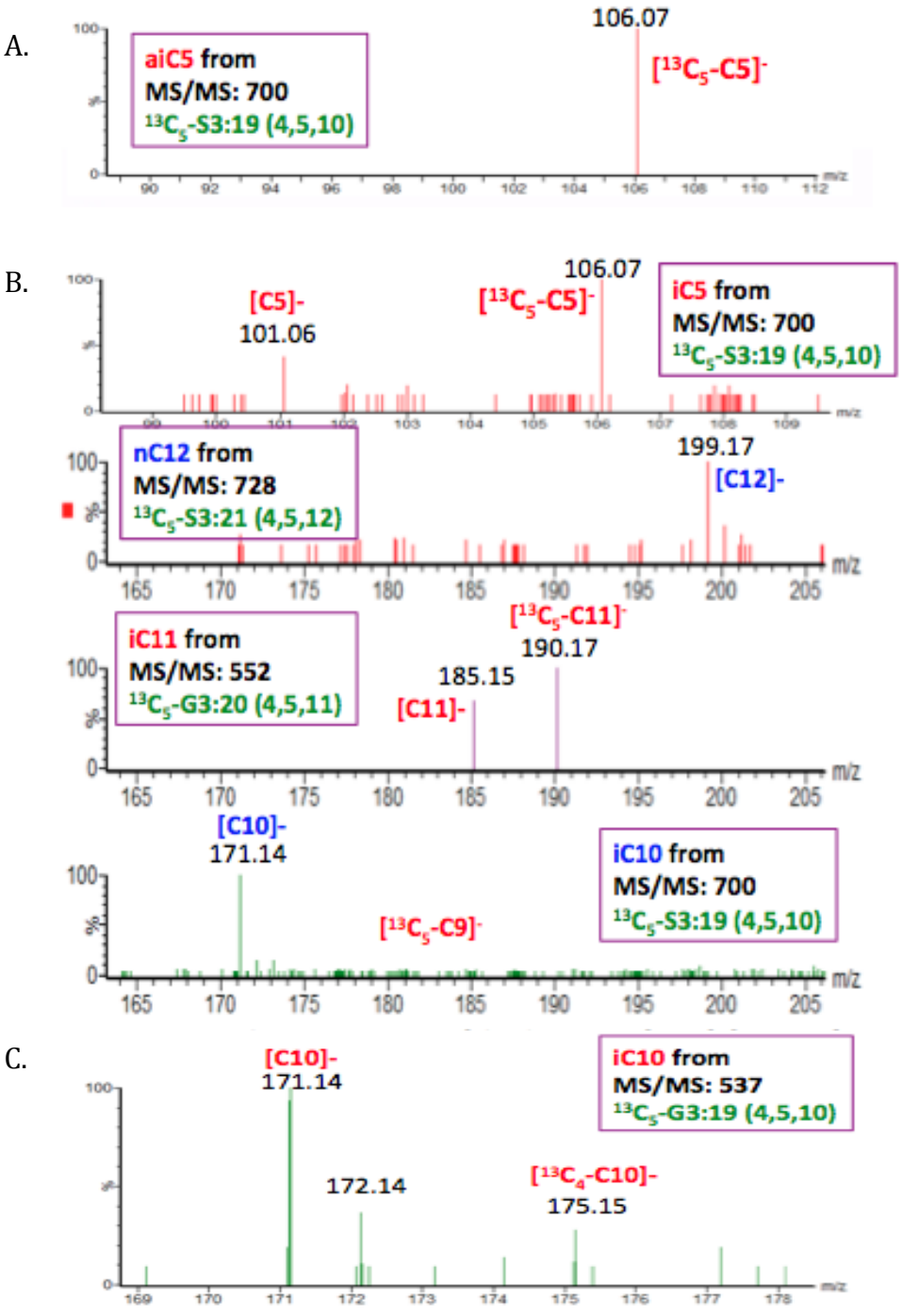


Figure 4.10. MS/MS spectra of acylsugar aliphatic acyl groups labeled with ^{13}C-BCAAs in S. pennellii LA0716*. A. Negative-ion product ion MS/MS spectra of heavy C5 acyl chain in [$^{13}C_6$,^{15}N]Ile labeled acylsugars; B. Product ion spectra showing heavy C5, C9, C11 and light C12 acyl chain fragments in [$^{13}C_5$,^{15}N]Val labeled acylsugars; C. product ion spectra showing heavy C10 in [$^{13}C_6$,^{15}N]Leu labeled acylsugars from LA0716 (top to bottom).*

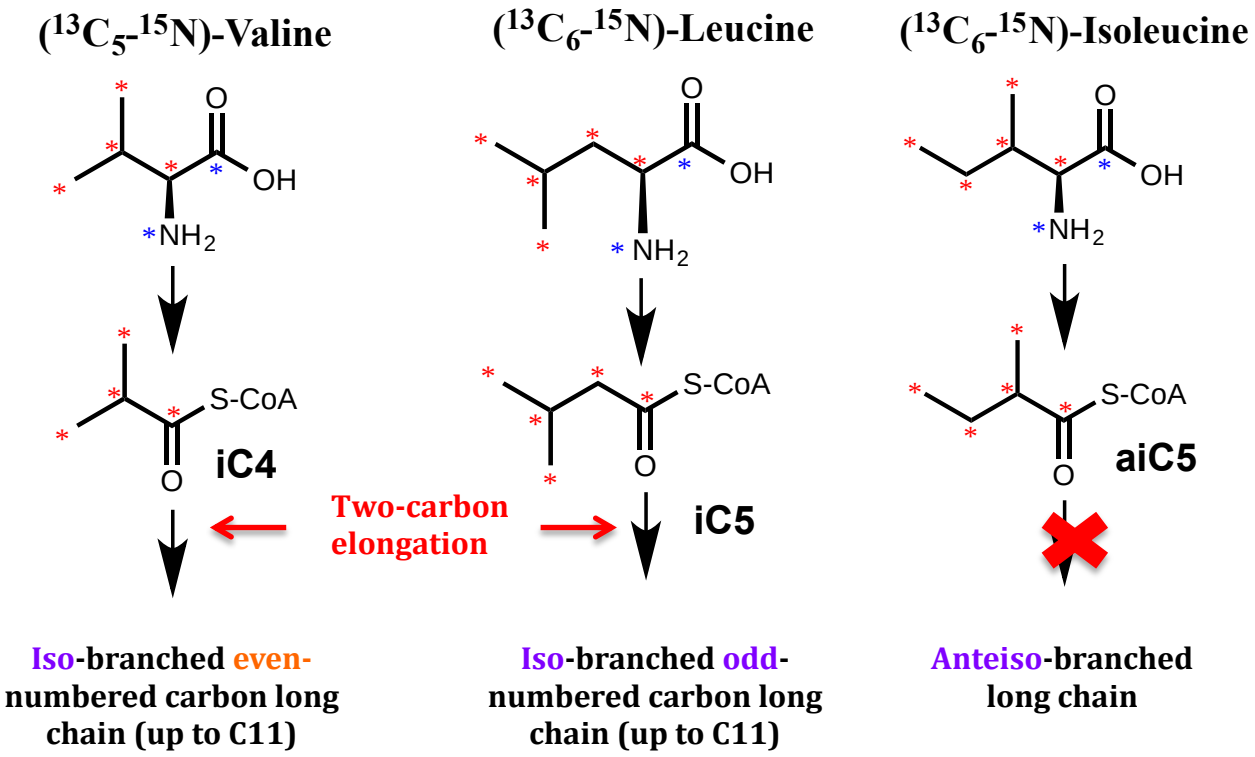


Figure 4.11. Proposed elongation pathway of aliphatic acyl chain with different branching types in S. pennellii LA0716. *Heavy isotopic atoms were labeled with stars. Red star labeled atoms were incorporated into acyl groups.*

Findings from a recent investigation of isopropylmalate synthase (IPMS) alleles in wild and cultivated tomato species can be applied to explain the disparate metabolic pathways of acyl chains within *S. pennellii* accessions [37]. Compared with the *SlIPMS3* gene encoded in cultivated tomato M82 and orthologs in northern *S. pennellii* accessions (including LA2560), *SpIPMS3* in LA0716 encodes an inactive enzyme that has a longer C-terminal truncation, which correlates not only with iC4 and iC5 acyl chain abundances in acylsugars, but also disabled the one-carbon elongation in LA0716 that usually operates by the combination of IPMS and isopropylmalate isomerase (IPMI). Nevertheless, since iC6 is not synthesized from iC5 or labeled by [$^{13}C_6$,^{15}N]Leu, and one-carbon elongation is only involved in short acyl chain (C5, C6) biosynthesis, it seems like the IPMS3 or other enzymes catalyzed one-carbon elongation are prefer to use short branched acyl-CoAs (*e.g.* iC4-CoA, iC5-CoA, aiC5-CoA) as substrates to

making acyl groups no longer than C6 through 1-2 cycle of reactions. However, more chemical and biological evidence is needed to test this hypothesis.

4.3.5 Investigation of acylsugar related aliphatic acid elongation in cultivated tomato M82

The acyl chain elongation pathways of cultivated tomato M82, analyzed using the ^{13}C-BCAAs labeling coupled with UHPLC-MS/MS approach, shares similarity with mechanism in *S. pennellii* LA2560 (Figure 4.13). Since tomato expressed the active longer form *SlIPMS3* enzyme in trichomes, normal one-carbon elongation mechanism is active in biosynthesis of short iso-branched acyl groups (C4-C6). Iso-branching in iC4, iC5 and anteiso-branching in aiC5 were labeled individually by [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_6$,^{15}N]Ile, while heavy iC5 ([$^{13}C_4$-C5]$^-$, *m/z* 105), iC6 ([$^{13}C_5$-C6]$^-$, *m/z* 120) and aiC6 ([$^{13}C_5$-C6]$^-$, *m/z* 120) fragment ions were identified in MS/MS spectra of S4:18(2,5,5,6), suggesting one-carbon elongation, based on [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_6$,^{15}N]Ile as precursors (Figure 4.12). It was notable that heavy ^{13}C from [$^{13}C_5$,^{15}N]Val was incorporated into iC6 chains ([$^{13}C_4$-C6]$^-$, *m/z* 119), indicating biosynthesis of iC6 might undergo either one cycle of two-carbon elongation or two cycles of one-carbon mechanism, which cannot be distinguished from these results.

No clear evidence for two-carbon elongation in tomato (M82) can be derived from these results, especially for that involved in acyl long chain (C10-12) synthesis. C12 fragment ions from three major tomato acylsucroses S3:22(5,5,12), S3:23(5,6,12) and S4:24(2,5,5,12) did not show isotope enrichment in plants grown on any of the three ^{13}C-BCAAs precursors, which was consistent with previous report that C12 was straight chain [14]. In addition, C11 containing

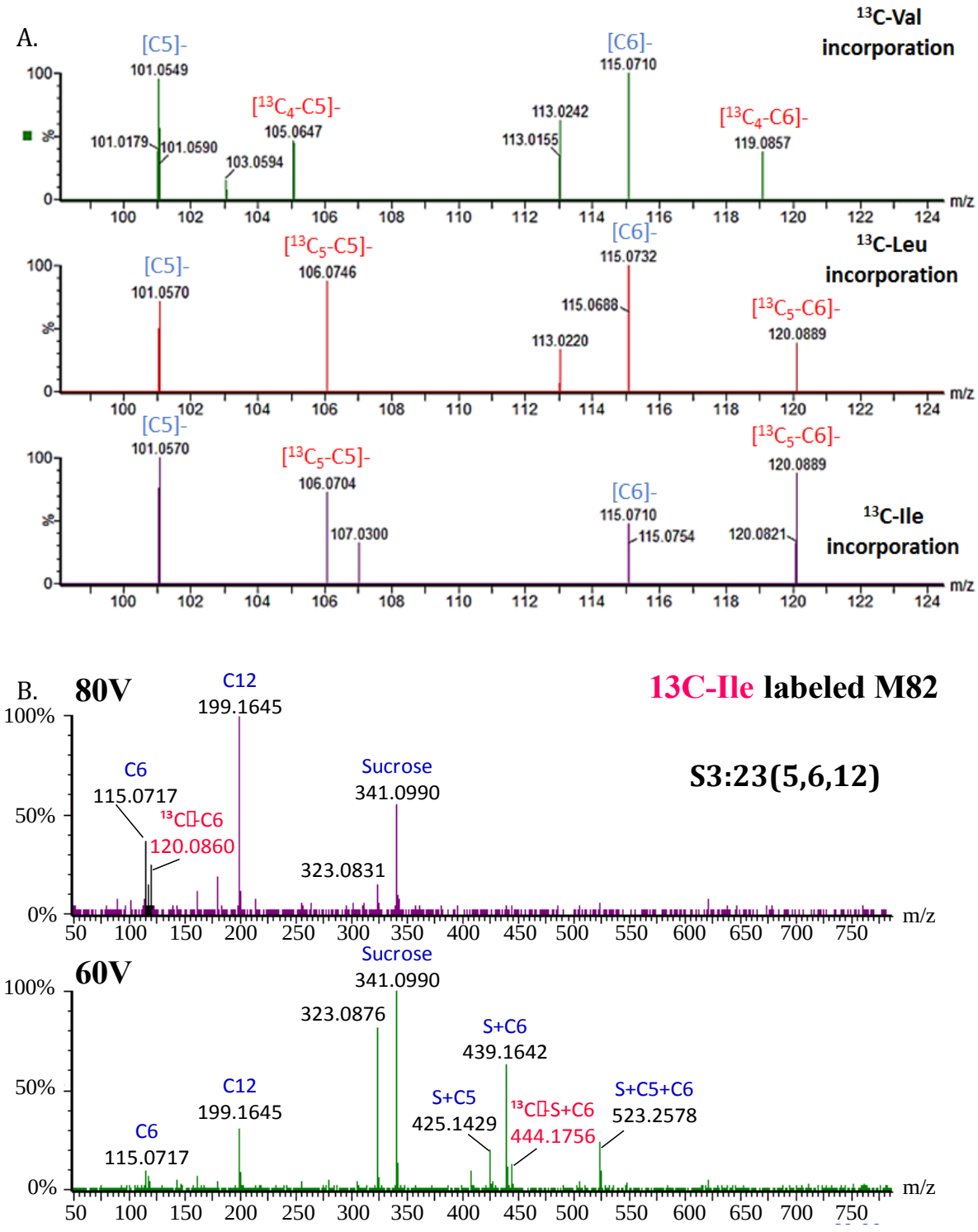


Figure 4.12. Negative-ion MS/MS spectra showing aliphatic acyl groups labeled with ^{13}C-BCAAs in cultivated tomato M82*. A. Stable isotope incorporation pattern of C5 and C6 acyl groups in M82 labeled using [$^{13}C_5$,^{15}N]Val, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_6$,^{15}N]Ile (top to bottom) B. CID spectra of S3:23 (5,6,12) in tomato indicated [$^{13}C_6$,^{15}N]Ile was not incorporated into C12 acyl groups.*

acylsugars were barely detected in tomato, while C10 from S4:22(2,5,5,10) also was not labeled using any of the three heavy isotopic BCAA precursors. To avoid false negative result due to the low acylsugar intensities in sample, two C10-acylsugar enriched introgression lines (*S. lycopersicum x S. pennellii*) IL7-4 and IL12-2 were subjected to this analysis, which produced the same stable isotopic labeling pattern of C10 as in tomato. Therefore, as neither heavy iso-branched ([$^{13}C_5$,^{15}N]Val, [$^{13}C_5$,^{15}N]Leu) nor ^{13}C anteiso-branched amino acids ([$^{13}C_6$,^{15}N]Ile)

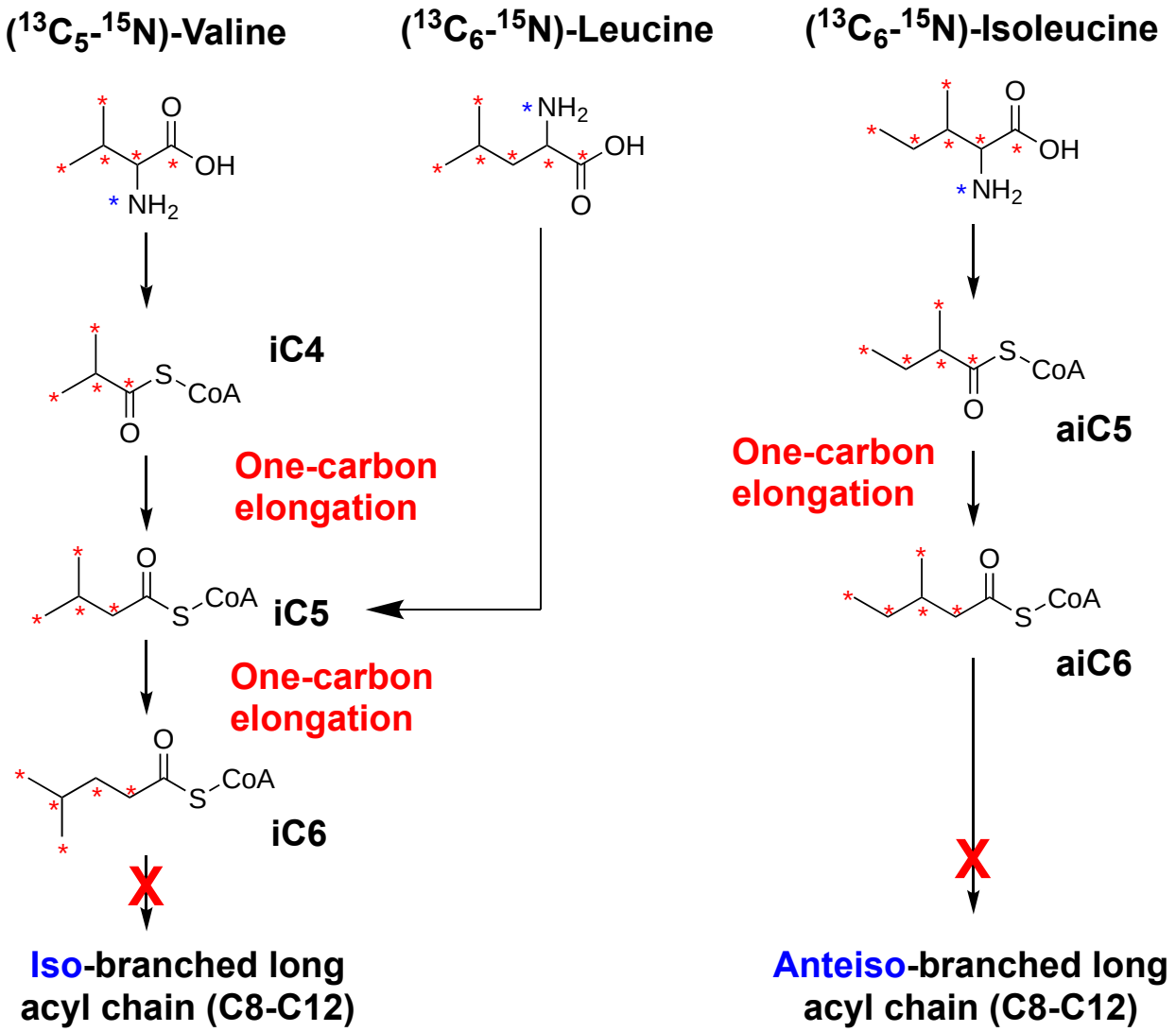


Figure 4.13. Proposed elongation pathway of aliphatic acyl chain with different branching types in cultivated tomato M82. *Heavy isotopic atoms were labeled with stars. Red star labeled atoms were incorporated into acyl groups.*

were incorporated into C10 or C12, it is hypothesized that acyl long chains in tomato acylsugars and the two introgression lines, were straight chains and derived from normal fatty acid biosynthetic intermediates rather than BCAAs.

4.3.6 Investigation of acylsugar related aliphatic acid elongation in *Petunia exserta*

The analysis of [$^{13}C_5$,^{15}N]Val incorporation of acylsugar related aliphatic chains in *P. exserta* provided evidence for one-carbon elongation in the genus *Petunia*. Heavy fragment ions of acyl groups C5 ([$^{13}C_4$-C5]$^-$, *m/z* 105), C6 ([$^{13}C_4$-C6]$^-$, *m/z* 119), C7 ([$^{13}C_4$-C7]$^-$, *m/z* 133), C8 ([$^{13}C_4$-C8]$^-$, *m/z* 147) were observed in MS/MS spectra of acylsugar metabolites S4:21(5,5,5,6), S(m)5:25(m,5,5,5,7) and S(m)5:26(m,5,5,5,8); (Figure 4.14), indicating that [$^{13}C_5$,^{15}N]Val was used as precursor of a series of short- to medium-length iso-branched acyl chains through one-carbon elongation (Figure 4.15). Additionally, heavy isotope incorporation into C7 acyl group fragments generated by [$^{13}C_5$,^{15}N]Val labeling were also detected in *P. integrifolia* and *P. axillaris*.

Label incorporation into petunia acylsugars failed to yield evidence that Leu or Ile metabolites are involved in one-carbon elongation mechanism was acquired from relevant labeling investigations. No heavy isotopes derived from [$^{13}C_6$,^{15}N]Leu or [$^{13}C_6$,^{15}N]Ile were incorporated into C6 acyl groups in any petunia acylsugars, while only fragments of [$^{13}C_4$]iC5, [$^{13}C_5$]iC5 labeled by [$^{13}C_6$,^{15}N]Val and [$^{13}C_6$,^{15}N]Leu and [$^{13}C_5$]aiC5 labeled with [$^{13}C_6$,^{15}N]Ile were observed in MS/MS spectra of *P. exserta* acylsugars. Similar results regarding the labeling of branching C5 and C6 were observed in the other two petunia species as well. Though the low ^{13}C incorporation level may result in a false negative conclusion of this analysis, the bioactive preference of acyl-CoAs, as well as BCAAs precursors, of enzymes related to one-carbon

elongation should play a key role in driving such outcomes. Since [$^{13}C_6$,^{15}N]Val could be incorporated into most of the iso-branched medium-length acyl groups in *P. exserta*, it suggested that valine-derived acyl-CoAs are preferred by one-carbon elongation enzymes. Meanwhile, Table 4.4 shows ^{13}C enrichment of major acylsugars detected in *P. exserta* labeled with different

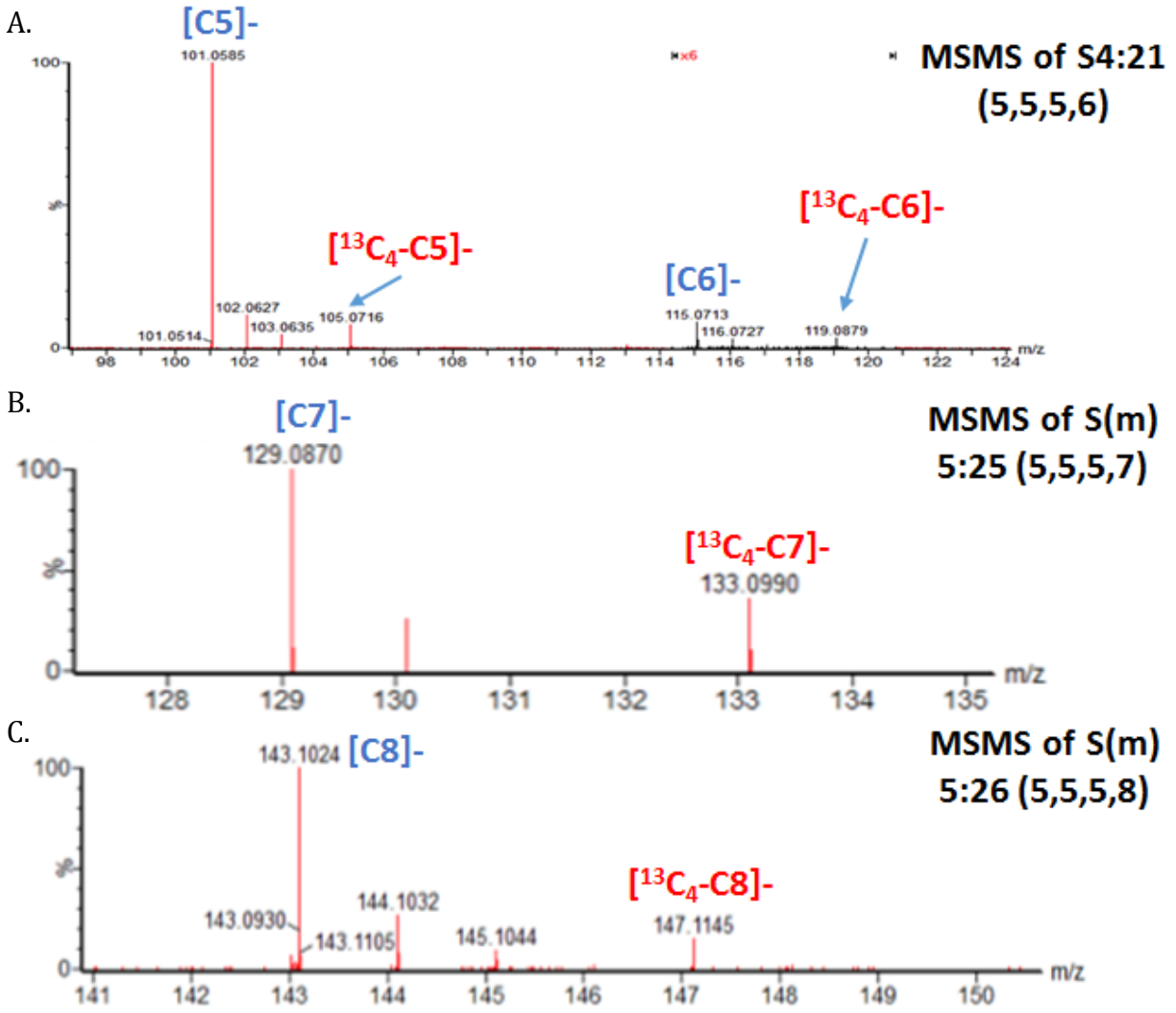


Figure 4.14. Negative-ion MS/MS spectra of aliphatic acyl groups labeled with [$^{13}C_5$,^{15}N]Val in P. exserta*. A. Stable isotope labeling pattern of C5 and C6 acyl groups in P. exserta S4:21, C6 range with 6X magnification. B. Heavy C7 acyl group was labeled by [$^{13}C_5$,^{15}N]Val in malonylated acylsugar S(m)5:25. C. Heavy C8 acyl chain observed in S(m)5:26 of P. exserta labeled with [$^{13}C_5$,^{15}N]Val.*

^{13}C-BCAAs. The total incorporation of ^{13}C enriched in all *P. exserta* acylsugars from [$^{13}C_5$,^{15}N]Val (~1% [m+4]/[m]) was highest among the three ^{13}C-BCAAs labeled samples (~0.2% [m+5]/[m] ^{13}C enrichment for [$^{13}C_5$,^{15}N]Leu and [$^{13}C_5$,^{15}N]Ile), which was consistent with the hypothesis that valine was the major precursor involved in acyl chain biosynthesis in Petunia species. It is worth mentioning that although one-carbon elongation has been identified

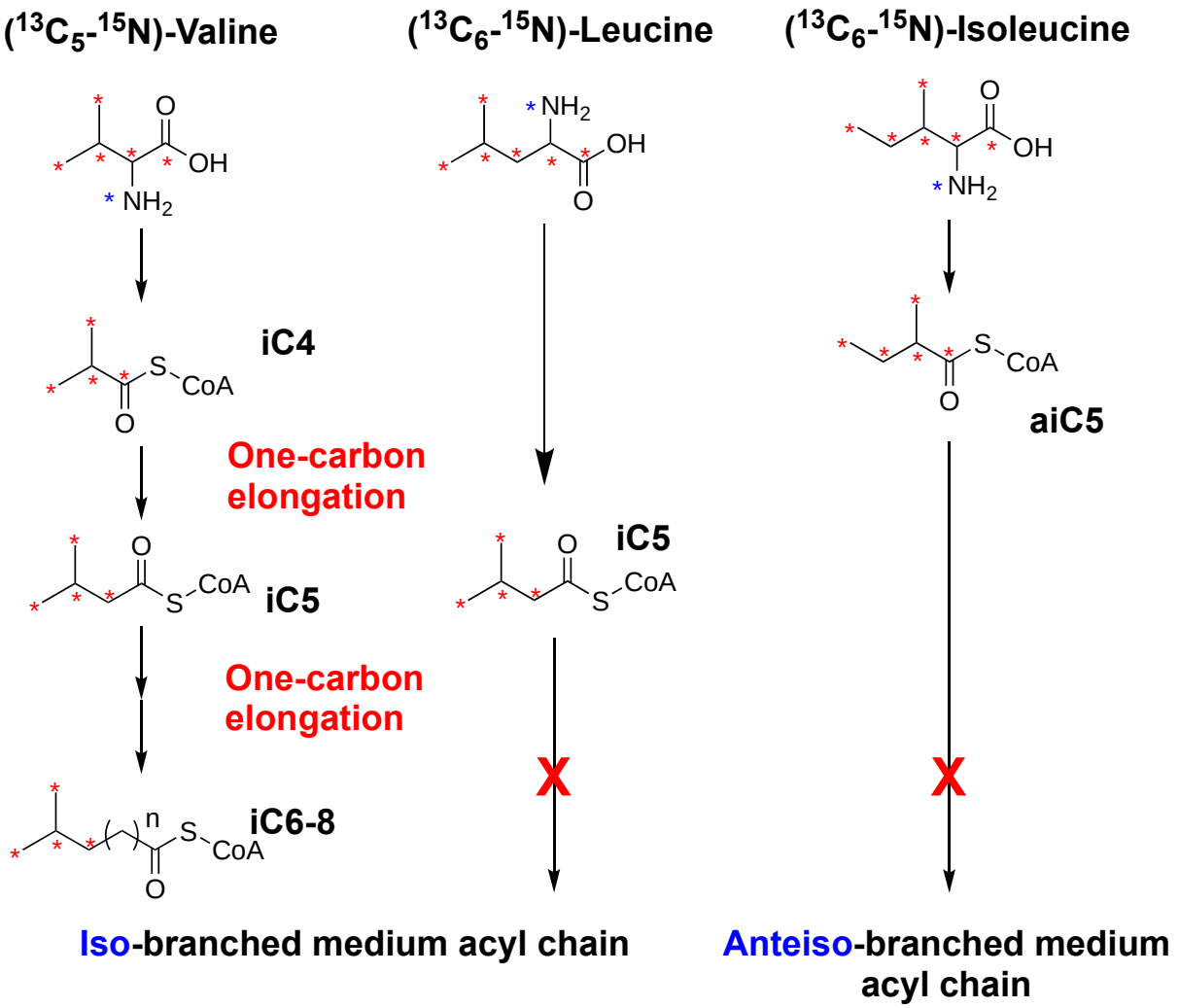


Figure 4.15. Proposed elongation pathway of aliphatic acyl chain with different branching types in Petunia exserta. *Heavy isotopic atoms were labeled with stars. Red star labeled atoms were incorporated into acyl groups.*

from both petunia and tomato (M82 and *S. pennellii* LA2560), in petunia, branched medium acyl chains are synthesized through several cycles (up to 4) of one-carbon reaction using valine as precursor; while in tomato, the biosynthesis of C6 derived from leucine or valine only experienced one cycle or up to two cycles of one-carbon elongation. One hypothesis proposed to explain these observations is that enzymes involved in one-carbon elongation in different plant species may have distinct bioactive preference of acyl-CoA substrates, which need to be more deeply investigated in the future.

Table 4.4. Stable isotope enrichment in P. exserta acylsugars. *Calculated relative abundances of [M+formate]$^-$ isotopologs $A_{[m+5]}/A_{[m]}$, $A_{[m+4]}/A_{[m]}$ represented ^{13}C enrichment from [$^{13}C_6$,^{15}N]Ile, [$^{13}C_6$,^{15}N]Leu and [$^{13}C_5$,^{15}N]Val labeled samples separately. Hoagland solution samples were used in growth media for the unlabeled control group.*

Acylsugars	Control (m+5/m%)	Control (m+4/m%)	13C Ile (m+5/m%)	13C Leu (m+5/m%)	13C Val (m+4/m%)
S4:20	0.19	0.45	0.25	0.30	1.20
S4:21	0.08	0.32	0.17	0.22	1.03
S4:22	0.04	0.40	0.20	0.12	0.79
S4:23	0.15	0.55	0.42	0.24	0.83
S(m)5:23	0.16	0.57	0.32	0.20	0.98
S(m)5:24	0.09	0.41	0.22	0.18	0.93
S(m)5:25	0.06	0.39	0.29	0.13	0.89
Total acylsugars	0.11	0.44	0.28	0.17	0.91

4.4 Conclusions

Structures of two acylglucoses from wild tomato *S. pennellii* LA2560 were proposed using UHPLC-MS coupled with NMR analysis, with one of them containing an anteiso (ai) C6 alkyl ester, which was not previously detected during our profiling of other *Solanum* accessions [4,6]. Considering all 18 proposed structures, particularly with respect to the lengths and position conservation of acyl chains, chemical diversity of acylsugar metabolites is observed not only

between two *S. pennellii* accessions, but also across wild and cultivated tomato species. Metabolite profiling presented here, particularly with regard to demonstration of selectivity in positions of ester substitution, provides an important foundation for subsequent *in vivo* and *in vitro* testing of hypotheses regarding the evolution of biosynthetic genes involved in acylsugar diversity and accumulation.

A total of six aiC6-containing acylsugars were identified from *S. pennellii* LA2560, and the presence of the aiC6 acyl chain suggested one-carbon elongation mechanism, which had not been clearly documented in wild tomato species in previous publications [11-18, 40,41]. ^{13}C-amino acid labeling patterns in *S. pennellii* LA2560 and LA0716 acylsugars were analyzed using UHPLC-MS/MS, and provided evidence that both one-carbon and two-carbon elongation pathways are involved in acylsugar biosynthesis in *S. pennellii* LA2560, while esters in accession LA0716 were consistent with only two-carbon fatty acid elongation and a negligible role for one-carbon elongation. This ^{13}C-BCAA labeling coupled with UHPLC-MS/MS approach was further applied to other plants in the family *Solanaceae* to deeply investigate the acyl groups elongation pathways in cultivated tomato M82 and petunia, which offered a simplified but powerful tool to reveal differences in aliphatic ester biosynthesis phenotypes among genotypes.

According to the results of this study, acylsugar related acyl chain biosynthesis may be more complex than a choice between one-carbon or two-carbon elongation for a given genotype. Mechanisms involved in producing even-numbered iso-branched long chain acyl groups (*e.g.* iC10) in LA2560, as well as straight long chain acyl groups (*e.g.* nC10, nC12) in tomato were revealed by the labeling results. Furthermore, besides IPMS3 alleles identified in *S. pennellii*

accessions and cultivated tomato [37], our understanding of acyl chain biosynthesis related enzymes, particularly regarding to the relative importance of acyltransferase selectivity for acyl-CoAs substrates (*e.g.* iC4-CoA, iC5-CoA and aiC5-CoA) vs. metabolic differences that influence fluxes into specific CoA esters was limited. It is recommended that additional efforts be made to resolve these issues in the near future.

REFERENCES

REFERENCES

1. Milo, R.; Last, R. L., Achieving diversity in the face of constraints: lessons from metabolism. *Science* **2012**, *336* (6089), 1663-1667.

2. Cragg, G. M.; Newman, D. J., Plants as a source of anti-cancer agents. *Journal of Ethnopharmacology* **2005**, *100* (1), 72-79.

3. Wurtzel, E. T.; Kutchan, T. M., Plant metabolism, the diverse chemistry set of the future. *Science* **2016**, *353* (6305), 1232-1236.

4. Glas, J. J.; Schimmel, B. C. J.; Alba, J. M.; Escobar-Bravo, R.; Schuurink, R. C.; Kant, M. R., Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *International Journal of Molecular Sciences* **2012**, *13* (12), 17077-17103.

5. DellaPenna, D.; O'Connor, S. E., Plant gene clusters and opiates. *Science **2012***, *336* (6089), 1648-1649.

6. Last, R. L.; Jones, A. D.; Shachar-Hill, Y., Towards the plant metabolome and beyond. *Nature Reviews Molecular Cell Biology* **2007**, *8* (2), 167-174.

7. Schilmiller, A. L.; Last, R. L.; Pichersky, E., Harnessing plant trichome biochemistry for the production of useful compounds. *The Plant Journal* **2008**, *54* (4), 702-711.

8. Arrendale, R. F.; Severson, R. F.; Sisson, V. A.; Costello, C. E.; Leary, J. A.; Himmelsbach, D. S.; Van Halbeek, H., Characterization of the sucrose ester fraction from *Nicotiana glutinosa*. *Journal of Agricultural and Food Chemistry* **1990**, *38* (1), 75-85.

9. Tissier, A., Glandular trichomes: what comes after expressed sequence tags? *The Plant Journal* **2012**, *70* (1), 51-68.

10. Burke, B. A.; Goldsby, G.; Mudd, J. B., Polar epicuticular lipids of *Lycopersicon pennellii*. *Phytochemistry* **1987**, *26* (9), 2567-2571.

11. King, R. R.; Calhoun, L. A., 6 2, 3-Di-O-and 1, 2, 3-tri-O-acylated glucose esters from the glandular trichomes of *Datura metel*. *Phytochemistry* **1988**, *27* (12), 3761-3763.

12. Slocombe, S. P.; Schauvinhold, I.; McQuinn, R. P.; Besser, K.; Welsby, N. A.; Harper, A.; Aziz, N.; Li, Y.; Larson, T. R.; Giovannoni, J., Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiology* **2008**, *148* (4), 1830-1846.

13. Ghosh, B.; Jones, A. D., Profiling, characterization, and analysis of natural and synthetic acylsugars (sugar esters). *Analytical Methods* **2017**, *9*, 892-905.

14. Ghosh, B.; Westbrook, T. C.; Jones, A. D., Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014**, *10* (3), 496-507.

15. King, R. R.; Singh, R. P.; Boucher, A., Variation in sucrose esters from the type B glandular trichomes of certain wild potato species. *American Journal of Potato Research* **1987**, *64* (10), 529-534.

16. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Sucrose esters associated with glandular trichomes of wild *Lycopersicon* species. *Phytochemistry* **1990**, *29* (7), 2115-2118.

17. King, R. R.; Singh, R. P.; Calhoun, L. A., Elucidation of structures for a unique class of 2, 3, 4, 3′-tetra-O-acylated sucrose esters from the type B glandular trichomes of *Solanum neocardenasii* Hawkes & Hjerting (PI 498129). *Carbohydrate Research* **1988**, *173* (2), 235-241.

18. King, R. R.; Calhoun, L. A.; Singh, R. P.; Boucher, A., Characterization of 2, 3, 4, 3′-tetra-O-acylated sucrose esters associated with the glandular trichomes of *Lycopersicon typicum*. *Journal of Agricultural and Food Chemistry* **1993**, *41* (3), 469-473.

19. McKenzie, C. L.; Puterka, G. J., Effect of sucrose octanoate on survival of nymphal and adult *Diaphorina citri* (Homoptera: Psyllidae). *Journal of Economic Entomology* **2004**, *97* (3), 970-975.

20. Rodriguez, A. E.; Tingey, W. M.; Mutschler, M. A., Acylsugars of Lycopersicon pennellii deter settling and feeding of the green peach aphid (Homoptera: Aphididae). *Journal of Economic Entomology* **1993**, *86* (1), 34-39.

21. Hill, K.; Rhode, O., Sugar- based surfactants for consumer products and technical applications. *European Journal of Lipid Science and Technology* **1999**, *101* (1), 25-33.

22. Schilmiller, A.; Shi, F.; Kim, J.; Charbonneau, A. L.; Holmes, D.; Jones, A. D.; Last, R. L., Mass spectrometry screening reveals widespread diversity in trichome specialized metabolites of tomato chromosomal substitution lines. *The Plant Journal* **2010**, *62* (3), 391-403.

23. Chortyk, O. T.; Pomonis, J. G.; Johnson, A. W., Syntheses and characterizations of insecticidal sucrose esters. *Journal of Agricultural and Food Chemistry* **1996**, *44* (6), 1551-1557.

24. Weinhold, A.; Baldwin, I. T., Trichome-derived O-acyl sugars are a first meal for caterpillars that tags them for predation. *Proceedings of the National Academy of Sciences of the USA* **2011**, *108* (19), 7855-7859.

25. Ghangas, G. S.; Steffens, J. C., UDPglucose: Fatty acid transglucosylation and transacylation in triacylglucose biosynthesis. *Proceedings of the National Academy of Sciences of the USA* **1993**, *90* (21), 9911-9915.

26. St-Pierre, B.; De Luca, V., Chapter nine evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. *Recent Advances in Phytochemistry* **2000**, *34*, 285-315.

27. D'Auria, J. C., Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology* **2006**, *9* (3), 331-340.

28. Fan, P.; Miller, A. M.; Schilmiller, A. L.; Liu, X.; Ofner, I.; Jones, A. D.; Zamir, D.; Last, R. L., In vitro reconstruction and analysis of evolutionary variation of the tomato acylsucrose metabolic network. *Proceedings of the National Academy of Sciences of the USA* **2016**, *113* (2), E239-E248.

29. Schilmiller, A. L.; Charbonneau, A. L.; Last, R. L., Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proceedings of the National Academy of Sciences of the USA* **2012**, *109* (40), 16377-16382.

30. Schilmiller, A. L.; Moghe, G. D.; Fan, P.; Ghosh, B.; Ning, J.; Jones, A. D.; Last, R. L., Functionally divergent alleles and duplicated loci encoding an acyltransferase contribute to acylsugar metabolite diversity in Solanum trichomes. *The Plant Cell* **2015**, *27* (4), 1002-1017.

31. Kroumova, A. B.; Wagner, G. J., Different elongation pathways in the biosynthesis of acyl groups of trichome exudate sugar esters from various solanaceous plants. *Planta* **2003**, *216* (6), 1013-1021.

32. Slocombe, S. P.; Schauvinhold, I.; McQuinn, R. P.; Besser, K.; Welsby, N. A.; Harper, A.; Aziz, N.; Li, Y.; Larson, T. R.; Giovannoni, J., Transcriptomic and reverse genetic analyses of branched-chain fatty acid and acyl sugar production in *Solanum pennellii* and *Nicotiana benthamiana*. *Plant Physiology* **2008**, *148* (4), 1830-1846.

33. Kandra, L.; Severson, R.; Wagner, G. J., Modified branched- chain amino acid pathways give rise to acyl acids of sucrose esters exuded from tobacco leaf trichomes. *European Journal of Biochemistry* **1990**, *188* (2), 385-391.

34. Kroumova, A. B.; Xie, Z.; Wagner, G. J., A pathway for the biosynthesis of straight and branched, odd-and even-length, medium-chain fatty acids in plants. *Proceedings of the National Academy of Sciences of the USA* **1994**, *91* (24), 11437-11441.

35. Walters, D. S.; Steffens, J. C., Branched chain amino acid metabolism in the biosynthesis of *Lycopersicon pennellii* glucose esters. *Plant Physiology* **1990**, *93* (4), 1544-1551.

36. Kandra, L.; Wagner, G. J., Chlorsulfuron modifies biosynthesis of acyl acid substituents of sucrose esters secreted by tobacco trichomes. *Plant Physiology* **1990**, *94* (3), 906-912.

37. Ning, J.; Moghe, G. D.; Leong, B.; Kim, J.; Ofner, I.; Wang, Z.; Adams, C.; Jones, A. D.; Zamir, D.; Last, R. L., A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant Physiology* **2015**, *169* (3), 1821-1835.

38. Shapiro, J. A.; Steffens, J. C.; Mutschler, M. A., Acylsugars of the wild tomato *Lycopersicon pennellii* in relation to geographic distribution of the species. *Biochemical Systematics and Ecology* **1994**, *22* (6), 545-561.

39. Li, A. X.; Eannetta, N.; Ghangas, G. S.; Steffens, J. C., Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. *Plant physiology* **1999**, *121* (2), 453-460.

40. Ghangas, G. S.; Steffens, J. C., 1-O-Acyl-β-D-glucoses as fatty acid donors in transacylation reactions. *Archives of Biochemistry and Biophysics* **1995**, *316* (1), 370-377.

Chapter Five.

Concluding remarks

The study of genotype difference underlies various phenotypes of a specific plant accession will lead to the discovery of novel genes and enzymes involved in the production of plant specialized metabolites, which are essential towards the understanding of their biosynthesis and accumulation processes in plants [1,2]. One important purpose of this metabolic research is to build a connecting between phenotype and genotype of plants and to generate a suitable analytical platform for rapid profiling of specialized metabolites among plant with various genotypes. Results summarized in this dissertation provide a better understanding of acylsugar diversification, biosynthesis and accumulation within Solanaceae species, and will help to make significant contributions to the scientific control of acylsugar biosynthesis in crop plants that can increase plants resistance towards insects and disease.

The study focused on structure elucidation of acylsugar metabolites identified from petunia species was performed as a complementary research to the previous analysis of acylsugar diversity within *Solanum* species [3,4,5]. As described in Chapter 2, twenty-eight novel acylsucroses, including thirteen tetra-acylsucroses and fifteen penta-malonylated acylsucroses, were purified from three petunia species using SAX-SPE coupled with semi-prep scale HPLC separation, structures of which were proposed according to UHPLC-MS and NMR results. Comparing with *Solanum* acylsugar metabolites characterized from previous studies, all four hydroxyl groups at pyranose ring of sucrose cores of *Petunia* acylsugars were esterified with acyl chains, which indicated new enzymes should be present in petunia trichomes to synthesize these metabolites. The observed structural diversity of acylsucroses in petunia species, especially with respect to the conservations of substitution positions with specific acyl groups, has been served as the reference in studying the bioactive and substrates selectivity of petunia acyltransferases (ASATs). Additionally, malonylated acylsucroses with one malonyl group attached to furanose

ring were identified in this study, which points out one of the future directions of genomic research for identification of genes and enzymes function as acylsugar related malonyl-CoA transferases expressed in petunia trichomes. Moreover, since the lack of standards is one of the big challenges in both quantitative and qualitative analysis of acylsugars in plant species, purified acylsucroses obtained in this study will be a treated as valuable resources to establish a reliable database the for extensive investigation of acylsugar metabolites in the future.

A rapid UHPLC-MS/MS coupled with stable isotopic labeling strategy documented in Chapter 3 could be used as one attempt to overcome the bottleneck of novel gene discovery that is limited by inefficient metabolite identification. [^{13}C]-Branched chain amino acids (BCAAs) were used as isotopic tracers in order to label and distinguish isomeric plant metabolites with diverse structures, particularly with regard to the lengths and branching type of aliphatic chains derived from metabolism of branched chain amino acids. Plant specialized metabolites tagged with [^{13}C]-BCAAs were identified easily according to their corresponded mass shift from unlabeled species using UHPLC-MS with both multiplexed CID and data dependent MS/MS analysis. Acylsugar isomeric metabolites with different acyl chain compositions could be demonstrated using this approach, which provided a fast and simplified strategy to investigate the structural diversity of plant specialized metabolites without purification.

Another important question needs to be answered by the research of plant metabolites is how are these natural products synthesized in plant tissues. To address this question, Chapter 4 describes the application of stable isotope labeling coupled with UHPLC-MS/MS methodology for investigation of acylsugar related acyl chain elongation mechanism *in vivo*. The presence of aiC6 group in *S. pennellii* LA2560 acylsugar metabolites was confirmed by NMR, and one-carbon

biosynthetic mechanism of this particular acyl chain was proposed based on UHPLC-MS/MS analysis of isotopic labeling patterns from [^{13}C]-BCAAs feeding plants. This BCAA derived one-carbon elongation pathway existed in both wild tomato *S. pennellii* LA2560 and cultivated tomato M82, which took place in the synthesis of short branched-acyl chains from C4 to C6. Additionally, in petunia plant *P. exserta*, branching acyl groups with skeleton length up to C8 were elongated from BCAAs through this one-carbon mechanism. In contrast, two-carbon elongation was identified in *S. pennellii* accessions LA0716 and LA2560 for long branching acyl chain biosynthesis (C9-C11), while no evidence for this mechanism was obtained from analysis of cultivated tomato and petunia species. Interestingly, C12 groups from all tomato species and C10, C12 from tomato cannot be labeled by stable isotopic procures, which indicate other mechanisms should be taken place in producing these straight long acyl groups within tomato species. As our understandings of acylsugar related genes and enzymes keep increasing in recent years, this study contributes to filling the missing puzzle regarding the generation of diverse acyl-CoA pool in acylsugar metabolite biosynthesis.

In summary, our research presented in this dissertation focused on questions about which kinds of specialized metabolites were produced in plants, what is the diversity of specialized metabolites within and among species, as well as how do plants produce such metabolites within particular species. Some achievements have been obtained and demonstrated here, but more questions still remain unanswered and need to be deeply studied in the future. For example, limited by the current analytical methodology and instrumental sensitivity, plenty of low abundance metabolites, which contribute to a large portion of chemical diversity, have not been characterized yet. Detailed mechanisms of biosynthesis of specialized metabolites, like malonylated acylsugars in petunia, iC10-contained acylsugar in *S. pennellii* LA2560 and straight

long chain (C10-C12) containing acylsugar in tomato, still cannot be explained by our current established pathways, which indicate novel genes and enzymes, as well as new functions of known enzymes, should be involved in making these compounds. Therefore, great efforts needed to be devoted to the plant specialized metabolism research, especially for comprehensive analysis of involved genes and improving utilized analytical techniques, which is essential to speed up the genetic engineering of plants to making more and more valuable compounds efficiently.

REFERENCES

REFERENCES

1. Schilmiller, A. L.; Charbonneau, A. L.; Last, R. L., Identification of a BAHD acetyltransferase that produces protective acyl sugars in tomato trichomes. *Proceedings of the National Academy of Sciences of the USA* **2012**, *109* (40), 16377-16382.

2. Ning, J.; Moghe, G. D.; Leong, B.; Kim, J.; Ofner, I.; Wang, Z.; Adams, C.; Jones, A. D.; Zamir, D.; Last, R. L., A feedback-insensitive isopropylmalate synthase affects acylsugar composition in cultivated and wild tomato. *Plant Physiology* **2015**, *169* (3), 1821-1835.

3. Schilmiller, A. L.; Moghe, G. D.; Fan, P.; Ghosh, B.; Ning, J.; Jones, A. D.; Last, R. L., Functionally divergent alleles and duplicated loci encoding an acyltransferase contribute to acylsugar metabolite diversity in Solanum trichomes. *The Plant Cell* **2015**, *27* (4), 1002-1017.

4. Ghosh, B.; Westbrook, T. C.; Jones, A. D., Comparative structural profiling of trichome specialized metabolites in tomato (*Solanum lycopersicum*) and *S. habrochaites*: acylsugar profiles revealed by UHPLC/MS and NMR. *Metabolomics* **2014**, *10* (3), 496-507.

5. Liu, X.; Enright, M.; Barry, C. S.; Jones, A. D., Profiling, isolation and structure elucidation of specialized acylsucrose metabolites accumulating in trichomes of Petunia species. *Metabolomics* **2017**, *13* (7), 85.