8-2020

DETECTION OF FOOD PROTEINS IN HUMAN SERUM USING MASS SPECTROMETRY METHODS

Abigail S. Burrows
University of Nebraska-Lincoln, aburrows@huskers.unl.edu

DETECTION OF FOOD PROTEINS IN HUMAN SERUM USING MASS SPECTROMETRY METHODS

by

Abigail S. Burrows

A DISSERTATION

Presented to the Faculty of

The Graduate College of the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Food Science and Technology

Under the Supervision of Professor Philip E. Johnson

Lincoln, Nebraska

August, 2020

DETECTION OF FOOD PROTEINS IN HUMAN SERUM USING MASS SPECTROMETRY METHODS

Abigail S. Burrows, Ph.D.

University of Nebraska, 2020

Advisor: Philip E. Johnson

Allergenic peanut proteins are highly resistant to digestion and are detectable by immunoassays after gastrointestinal digestion. The application of liquid chromatography-tandem mass spectrometry (LC-MS/MS) methods for *in vivo* detection of peptides originating from allergenic food proteins has not been thoroughly studied. The aim of this work was to develop an *in vivo* detection method for peanut proteins in serum using LC-MS/MS. The method(s) were validated by analyzing subject serum collected after peanut consumption.

Three de-complexation strategies were evaluated including (1) MS acquisition settings (i.e. inclusion, exclusion lists), (2) commercial depletion kits, and (3) organic solvent fractionation by discovery LC-MS/MS. Overall, none of these approaches were successful. No improvements occurred to peanut peptide detection using inclusion and exclusion lists. The commercial depletion kits removed abundant serum proteins, however, they simultaneously depleted peanut proteins. The fractionation method was efficient in reducing sample complexity, but demonstrated variable peanut protein fractionation.

Due to unsuitable de-complexing strategies, we evaluated non-depleted serum by targeted MS, including parallel reaction monitoring (PRM), multiple reaction monitoring (MRM), and MRM cubed (MRM^3). We identified 10 peanut peptides, representing the major peanut allergens. The limit of detection (LOD) of the sera-peanut model matrix (10:1 (w/w)) was similar for PRM and MRM, with detection at 1.0 ppm peanut protein (4.0 ppm peanut). The MRM^3 method did not provide improvements to LOD.

Following development of typical targeted methods, we re-investigated PRM with increased protein loading (600 µg). Peanut peptides were detected in two subject sera (sera 1, 2) at two different time points (60, 120 minutes, respectively). However, robust method development was unsuccessful, requiring further investigations in methodology.

Lastly, the intermolecular arrangements of peanut seed storage proteins were evaluated by offline size-exclusion chromatography (SEC) with discovery LC-MS/MS. Gaussian modeling was used to determine the native MW of proteins, isoforms, and complexes. The combination of Gaussian modeling and discovery LC-MS/MS of SEC fractions was a highly effective separation and identification tool.

ACKNOWLEDGEMENTS

First, I must thank Dr. Phil Johnson for his belief in my ability to pursue a Ph.D. in a new and expanding field, and providing me with abundant guidance and trust to conduct this research. Phil has provided an environment encouraging of independent research and thought, which I will continue to apply in my future research endeavors.

Secondly, if not as equally important, I must thank Dr. Justin Marsh for his continual support, advice, criticisms, and friendship. I am deeply grateful for his guidance and for making me a better scientist, but not a perfect scientist.

I would like to thank my committee members, Dr. Joe Baumert, Dr. Steve Taylor, Dr. Sophie Alvarez, and Dr. Vanessa Garcia Larsen for their continual guidance and imploring questions through the course of my Ph.D. education.

I additionally, would like to sincerely thank Dr. Steve Taylor and Dr. Joe Baumert for the opportunity to pursue, and achieve, a Master's degree in FARRP, which has ultimately led to the pursuance of my Ph.D., and for fostering an excellent research environment. I truly cherish my time in graduate school under your direction in FARRP.

I would also like to acknowledge and thank Dr. Stef Koppelman, for his continued advice, input, and supply of subject serum samples used routinely in this research project. I must thank Julie Nordlee and Dr. Samudra Wijeratne for their friendship and support since the beginning of my graduate school career. A huge thanks to Pat Gergen for her organization and helpfulness throughout my time at FARRP. I will greatly miss stopping by your office for the occasional chat. I would also like to

acknowledge the many other individuals in FARRP who I have interacted with throughout my many years with FARRP.

I would like to thank my many lab mates throughout the years, Dr. Vera Cao, Dr. Kaye Ivens, Dr. Nathan Marsteller, Dr. Bini Ramachandran, Rachel, Shimin, Lee, Jenna, and Tengfei.

An extra special thanks to Dr. Shyamali Jayasena for her many years of friendship throughout both my MS and PhD programs, for being an exceptional lab mate, for answering any lab questions I may have, and her continual book suggestions.

I would also like to thank my family, Mom, Dad, Cody, Amanda, Cort, Libby, Aunt Margie, and Ila, for their unending support throughout this endeavor.

I would equally like to express my sincere gratitude and thankfulness to my best bro-friends, Boone, David, Brianna, and Chip for keeping me grounded, continually making me laugh, and always allowing me to express my frustrations and excitement about life in general. Words cannot describe how grateful I am to have you four in my life.

To Charity and Bubbi for calling every week with mental health check-ins throughout the course of this writing process.

To Ashley, Max, and Julianne, thank you for your constant support, friendship, and unending plates of cheese and glasses of wine.

To Brittany, for continuing to be my best friend since freshman year of college as a result of random roommate assignments. Who knew our lives would end up this way

and our friendship would have grown this much, but you must know you have made me a more empowered woman, advocate, and friend, of which I am truly grateful.

I would also like to acknowledge the many additional friends and family members who have been with me throughout my educational career.

Lastly, I must express my deepest thanks to Colton. Thank you for understanding the many late nights, stress filled weeks, and joys this past year has brought into our lives. Not many people could understand the sole dedication it takes to complete a dissertation, but you have provided me with unending comfort and love in times of stress. You have brought laughter and joy into this enduring process each and every day and I very much look forward to many more joyful days.

Finally, I would like to thank FARRP for providing the funding for this project.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST OF TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

SUPPLEMENTARY

SUPPLEMENTARY TABLES

SUPPLEMENTARY FIGURES

CHAPTER 1: LITERATURE REVIEW

INTRODUCTION

Proteomics is the study of proteins, their interactions within a biological system (e.g. structure, function), and their relative abundances (Cravatt et al., 2007). The field of proteomics research is vast and expanded significantly due to advancements in protein biochemical techniques, and in mass spectrometry (MS) instruments and methods (Gillette and Carr, 2013). Proteins, composed of amino acids, are important macromolecules required for many routine and critical body functions including molecular transport (e.g. oxygen), immune responses, cell growth and repair, and reaction catalysts (e.g. enzymes) (Berg et al., 2002). Foods are an essential source of proteins, as well as other macromolecules (e.g. carbohydrates, lipids, minerals), and the sole source for essential amino acids, which must be consumed through an individual's diet (Shewry, 2007; Tessari et al., 2016). Unfortunately, for individuals diagnosed with a food allergy or sensitivity, consumption of certain food proteins can be detrimental to an individual's health as some allergenic food proteins are highly thermostable and resistant towards digestion, characteristics contributing to their allergenic behavior (Bannon, 2004).

Food allergens have been studied using a suite of molecular biology and protein biochemistry techniques including enzyme-linked immunosorbent assays (ELISA), immunoblotting, polymerase chain reaction (PCR), and digestibility assays (Koppelman and Hefle, 2006). ELISAs have been the primary method used in routine testing of allergenic food proteins, however recently, mass spectrometry methods have been developed for detection of allergenic food proteins (Croote and Quake, 2016; Koppelman

and Hefle, 2006; Monaci et al., 2015). The adoption of mass spectrometry in food allergen detection provides numerous benefits compared to ELISA methods, including multiplexing, identification of specific peptide and protein sequences, elucidation of protein modifications, and individual protein quantification (Croote and Quake, 2016; Monaci et al., 2015). MS offers numerous advantages however, these methods have not yet been routinely established and still require significant time for method development (Croote and Quake, 2016; Johnson et al., 2011).

Food allergies are individualistic reactions to food proteins affecting previously sensitized individuals, and have demonstrated an increasing prevalence over time (Sicherer et al., 2010; Turner et al., 2015). To better understand mechanisms of sensitization and allergic reactions, it is important to understand the uptake and transport of allergenic food proteins across the gut barrier and their subsequent interactions with the immune system (Reitsma et al., 2014). Previous studies have established *in vivo* detection of allergenic food proteins at very low concentrations, primarily by immunoassays (Baumert et al., 2009; Husby et al., 1985, 1986; JanssenDuijghuijsen et al., 2017; Schocker et al., 2016). However, *in vivo* detection has yet to be established using MS methods due to low analyte concentration, analyte conformation after gastric digestion and uptake, complexity of sera proteome, and sensitivity of mass spectrometers (JanssenDuijghuijsen et al., 2017; Reitsma et al., 2014).

Our aim within this section is to establish an understanding of the current literature and knowledge of *in vivo* uptake of allergenic food proteins, particularly peanut proteins, and the application of mass spectrometry in the study of food allergens. Peanut proteins are considered one of the most potent allergenic foods, and several studies have

focused on the *in vivo* detection using a multitude of protein chemistry techniques (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017; Schocker et al., 2016). The advancements achieved in mass spectrometry have expanded the available proteomic tools for allergenic protein studies, and as such, are the primary focus of this work.

FOOD SENSITIVITIES

Food sensitivities, including both food allergies and food intolerances, are collectively referred to as individualistic reactions to foods (Taylor and Baumert, 2012). These sensitivities only affect a small subset of the population, but can have severe impacts on an individual's quality of life (Stensgaard et al., 2017; Warren et al., 2016). Individuals commonly misidentify food intolerances as food allergies, whose mechanism is distinctly different and involves an immunologically-mediated response (Sloan and Powers, 1986; Taylor and Baumert, 2012). When properly diagnosed, the symptoms associated with various food sensitivities are easily distinguishable (Sloan and Powers, 1986; Taylor and Baumert, 2012). Recognition of the characteristic symptoms associated with each sensitivity, immune or non-immune mediated mechanisms, is critical for accurate diagnosis and dietary management since many individuals adopt modified dietary practices to avoid consumption of offending food(s) (Mazzocchi et al., 2017; Taylor and Baumert, 2012).

Food intolerances are non-immunologically mediated reactions and the predominant food sensitivity individuals experience (Guandalini and Newland, 2011; Sampson and Eigenmann, 1999). Individuals with food intolerances are generally able to tolerate larger doses of a causative food before exhibiting severe symptoms (Taylor and

Baumert, 2012). In contrast, low doses of allergenic food proteins can elicit immunologically mediated reactions in previously sensitized individuals (Hefle et al., 2001; Taylor and Baumert, 2012). Generally, individuals experiencing allergic reactions exhibit more severe symptoms, including anaphylaxis, a distinguishing feature of immunologically mediated reactions (Chinthrajah et al., 2015; Taylor and Baumert, 2012). However, for a majority of individuals, these causative foods, classified as allergens, do not cause adverse reactions (i.e. allergy, intolerance) and pose no imminent health threat (Taylor et al., 1992).

I. Food Intolerances

Food intolerances are non-immunologically mediated reactions caused by toxins, metabolic or pharmacological agents, or other unknown agents originating within foods or food additives (Boyce et al., 2010). These types of reactions may include gastrointestinal or absorption disorders, and, are commonly present in pediatric patients due to their under-developed immunological systems, however, these intolerances are most often outgrown (Sampson and Eigenmann, 1999). Food intolerances can be classified into three categories including, (1) anaphylactoid reactions, (2) metabolic food disorders, and (3) food idiosyncrasies (Lemke and Taylor, 1994; Taylor and Baumert, 2012).

A. Anaphylactoid Reactions

Anaphylactoid reactions are clinically identical to true food allergies presenting many of the same symptoms and mediators, but, are non-immunologically (i.e. IgE) mediated responses (Sampson, 2009). In these reactions, eliciting food substances initiate the release of mediators from mast cells and basophils, identical to mediators of an allergic reaction (Lemke and Taylor, 1994; Taylor and Baumert, 2012).

B. Metabolic Food Disorders

Metabolic food disorders are metabolic deficiencies of a particular food substance or food derived chemical, and are genetically inherited sensitivities (Taylor and Hefle, 2002). Lactose intolerance and favism are two well-characterized disorders, with distinct mediating mechanisms which are discussed in detail in subsequent paragraphs (Taylor and Hefle, 2002).

Lactose intolerance is the inability to digest lactose, the primary cow's milk sugar, due to a deficiency in the hydrolytic enzyme β-galactosidase (i.e. lactase) (Suarez and Savaiano, 1997). The disaccharide, lactose, is hydrolyzed by β-galactosidase, to its constituent monosaccharides, glucose and galactose, and transported across the small intestine for metabolic energy (Suarez and Savaiano, 1997). A deficiency in β-galactosidase results in the passage of non-digested lactose to the large intestine where bacteria metabolize the disaccharide substrate into carbon dioxide, water, and dihydrogen (H_2) (Kocian, 1988; Taylor and Hefle, 2002). Individuals suffering from lactose intolerance exhibit mild localized gastrointestinal (GI) symptoms including gastric and abdominal cramping, flatulence, and diarrhea (Bayless et al., 1975; Suarez and Savaiano,

1997). Many lactose intolerant individuals can safely consume small amounts of dairy products without experiencing significant symptoms (Lemke and Taylor, 1994; Taylor and Hefle, 2002).

Favism, another example of a metabolic food disorder, is caused by a lack of erythrocytic glucose-6-phosphate dehydrogenase (G6PDH), the most prevalent worldwide enzymatic deficiency (Taylor, 2014). The consumption of fava beans or inhalation of *Vicia faba* pollen, which contain the oxidants vicine and convicine, results in damage to erythrocytic membranes in individuals lacking G6PDH (Mager et al., 1980; Marquardt et al., 1997). G6PDH maintains the concentrations of glutathione (GSH) and nicotinamide dinucleotide phosphate (NADPH) in erythrocytes, preventing erythrocyte membrane oxidation (Taylor, 2014; Taylor and Hefle, 2002). Symptoms occur within 5 – 24 hours of ingestion causing fatigue, abdominal pain, nausea, fever, chills, hemolytic anemia, or more severe symptoms including hemoglobinuria, jaundice, and renal failure (Taylor, 2014; Taylor and Hefle, 2002).

C. Idiosyncratic Reactions to Foods

Food idiosyncrasies are adverse reactions to food additives involving unknown mechanisms and capable of causing mild to severe, and potentially life threatening symptoms (Guandalini and Newland, 2011; Taylor and Hefle, 2002; Taylor et al., 1989). Sulfite induced asthma is a principal example of an idiosyncratic reaction and has been elucidated by double blind placebo controlled food challenge (DBPCFC) (Taylor et al., 2014).

Sulfites, the causative agent in sulfite induced asthma, are naturally present in foods due to fermentation or as a food additive to inhibit browning (enzymatic and non-enzymatic), prevent microbial growth, inhibit oxidation, or as an aid in bleaching and dough conditioning (Bush and Montalbano, 2014; Taylor and Hefle, 2002; Taylor et al., 2014). Individuals diagnosed with asthma are more likely to develop sensitivities to sulfites, however not all asthmatics suffer from this sensitivity (Taylor et al., 2014). An estimated 5% of adult asthmatics have sulfite sensitivities, and individuals with severe asthma are more likely to develop sulfite sensitivities (Bush and Montalbano, 2014). As a result, sulfites must be labeled in order to protect those sulfite sensitive individuals (Taylor et al., 2014).

II. Food Hypersensitivities

Food hypersensitivity reactions are abnormal immunological responses to food or environmental proteins (e.g. pollen, dust, mold, animal dander) and categorized into four classes (I, II, III, and IV) based on their immunological mechanisms (Gell and Coombs, 1975; Sampson, 1991; Taylor and Hefle, 2002). Food hypersensitivities encompass two different types of immunologically mediated mechanisms, immediate hypersensitivity (type I) and delayed hypersensitivity (type IV) (Taylor and Baumert, 2012). Other hypersensitivity (allergic) reactions include type(s) II: antibody-dependent cytotoxic reactions and III: antigen-antibody complex mediated (Sampson, 1991). The primary focus of this section will be on delayed- and immediate-type hypersensitivity reactions.

Allergic reactions, including food and environmental reactions, are classified as type I, IgE-mediated, immediate hypersensitivity reactions (Gell and Coombs, 1975;

Sampson, 1991; Taylor and Hefle, 2002). Food allergies are caused by naturally occurring proteins in foods, generally present in high abundances (Bush and Hefle, 1996). Only a handful of foods are known to cause over 90% of reported reactions, and, are commonly known in the US as 'the Big 8.' (Bush and Hefle, 1996). 'The Big 8' includes peanuts, tree nuts, fish, shellfish, milk, egg, soy, and wheat (Hefle et al., 1996). Allergy to other fruits and vegetables have been reported, however, these reactions are often milder in elicited symptoms and localized to the oral cavity (Hefle et al., 1996). These food proteins are labile towards heat, processing, and gastric enzymes resulting in their rapid degradation (Amat Par et al., 1990; Hefle et al., 1996). Any protein is capable of eliciting an allergic reaction however, these reactions are rare in occurrence (Hefle et al., 1996).

Factors influencing the development of IgE mediated reactions include genetic pre-disposition, increase in GI tract permeability, premature birth, viral gastroenteritis (Taylor and Hefle, 2006). A study of monozygotic and dizygotic twins revealed the identical twins (i.e. monozygotic) are likely to develop allergy to the same food, illustrating the role of genetic heredity (Lack et al., 1999; Sicherer et al., 2000). Other studies have demonstrated the importance of gastrointestinal barrier permeability in relation to food allergy and other atopic disorders (Chambers et al., 2004; Nunes et al., 2016; Samadi et al., 2018). Several studies have demonstrated an increase in permeability of the gastrointestinal barrier in allergic individuals, allowing an increased molecular transport of intact allergenic food proteins into circulation (Li et al., 2006; Samadi et al., 2018; Yu et al., 2001).

Cell-mediated reactions are classified as type IV, and are delayed hypersensitivity reactions (Gell and Coombs, 1975; Sampson, 1991; Taylor and Hefle, 2002). These reactions generally cause localized inflammation to eliciting foods (Taylor and Hefle, 2002).

A. Delayed-Type Hypersensitivity

Delayed hypersensitivity reactions are non-IgE, cell-mediated immune responses whose mechanism(s) is not well understood (Taylor and Baumert, 2012; Taylor and Hefle, 2006). Cell-mediated reactions are caused by the interaction of food antigens with sensitized, tissue bound T-cells, causing the release of inflammatory mediators (Sampson, 1991; Taylor and Hefle, 2002). Affected individuals may experience symptoms 6 – 24 hour after ingestion which generally involve tissue inflammation localized to the gastrointestinal tract (Jones and Burks, 2008; Lemke and Taylor, 1994). Other resulting symptoms may include weight loss, anemia, bloating, diarrhea, or chronic fatigue (Taylor and Hefle, 2002).

Celiac disease (celiac sprue, non-tropical sprue, or gluten-sensitive enteropathy) is the most commonly well-known cell-mediated immune response (Taylor and Hefle, 2006). The consumption of gluten containing grains (e.g. wheat, barley, rye, spelt, triticale) causes an inflammatory response in the intestine leading to development of a malabsorption syndrome (Rubio-Tapia and Murray, 2008; Taylor and Hefle, 2001). Nutrient malabsorption results from the gluten-initiated damage to the mucosal lining of the small intestine (Rubio-Tapia and Murray, 2008; Taylor and Hefle, 2001). A

multitude of symptoms including inflammation, bloating, diarrhea, anemia, fatigue, and weight loss ensue (Rubio-Tapia and Murray, 2008).

B. Immediate-Type Hypersensitivity and Food Allergies

Food allergies are classified as immediate-type hypersensitivity reactions, and are IgE-mediated responses to proteins naturally present in foods (Taylor and Baumert, 2012; Taylor and Hefle, 2002; Yu et al., 2017). These abnormal immunological responses to foods can cause a multitude of symptoms affecting the cutaneous, gastrointestinal, respiratory, or even systemic reactions (Taylor and Hefle, 2002; Yu et al., 2017). For a majority of individuals, no adverse health effects occur during the regular consumption of allergenic foods however, for individuals diagnosed with a food allergy, consumption of allergenic foods may cause detrimental and potentially life-threatening health consequences (Taylor and Hefle, 2002). The immune system contains five antibody isotypes (IgA, IgD, IgE, IgG, IgM), each with their own specific functions (Murphy, 2012). IgE antibodies are responsible for elicitation of allergic reactions, in both food and environmental (e.g. pollen, dust, animal dander, mold) allergies (Taylor and Hefle, 2002). In non-atopic individuals, interactions with allergenic food proteins are mitigated by IgG and IgA antibodies, and do not elicit an immune response (Valenta et al., 2015). However, for pre-disposed individuals, interactions with allergenic proteins cause antigen-presenting cells (dendritic cells, B-cells), to induce T-helper 2 cells (Th2) which then produce cytokines (IL4, IL13) and initiate class switching to IgE for a particular allergen (Romagnani, 1997; Vercelli and Geha, 1992).

IgG4, a subtype of IgG, may have important roles in tolerance in certain individuals (Chinthrajah et al., 2016). In subjects participating in oral immunotherapy (OIT) or epicutaneous immunotherapy (EPIT), investigators reported an increase in IgG4 binding after completing the OIT in peanut allergic individuals (Koppelman et al., 2019; Vickery et al., 2013).

i. Allergenic Protein Sensitization

IgE mediated immune responses involve two phases, a sensitization phase and an elicitation phase, illustrated in Figure 1-1 (Taylor and Hefle, 2002). The sensitization phase, an asymptomatic process, initiates the production of allergen-specific IgE antibodies by B-cells after initial exposure to an allergenic protein (Mekori, 1996; Taylor and Hefle, 2006). These allergen specific IgE antibodies bind to the surfaces of tissue mast cells and circulating basophils (Mekori, 1996; Taylor and Baumert, 2012). Upon subsequent exposure to an allergenic protein, the allergenic epitopes crosslink surface bound IgE receptors (FcεRI) on mast cells and basophils (Rivera et al., 2008; Stone et al., 2010). Antigenic crosslinking signals effector cell degranulation, histamine release, and other physiological mediators associated with allergic responses (Rivera et al., 2008; Stone et al., 2010).

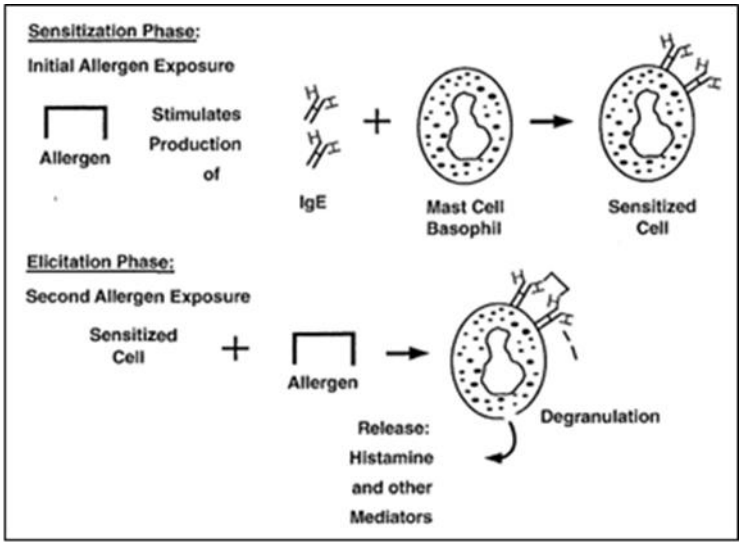


Figure 1-1. Reaction mechanism of an IgE mediated allergic response. Figure from Taylor and Baumert, 2012.

The onset of IgE-mediated symptoms is rapid, occurring within minutes to hours after ingestion of the causative food (Taylor and Hefle, 2006). Pre-formed histamine, the key physiological mediator, is responsible for immediate occurring symptoms including inflammation, pruritus, urticaria, and dermatitis (Stone et al., 2010; Taylor and Hefle, 2006). A myriad of symptoms affecting multiple organ systems may occur, as described in Table 1-1. Anaphylaxis is the most severe symptom associated with food allergies affecting multiple organ systems (respiratory, cardiovascular, cutaneous, and gastrointestinal) (Taylor and Hefle, 2006). Anaphylaxis develops rapidly upon exposure, and may result in fatality if epinephrine is not quickly administered (Taylor and Hefle, 2002). Due to the vast array of potential symptoms, sensitized individuals will often

experience a combination of symptoms rather than all potential symptoms (Taylor and Hefle, 2002). The overall reaction severity is influenced by an individual's sensitization, the amount of protein ingested, and the length of time since last exposure (Asero et al., 2007; Taylor and Hefle, 2002).

System affected	Symptom
Cutaneous	Angioedema
	Dermatitis or eczema
	Pruritus
	Urticaria
Gastrointestinal	Abdominal cramping
	Diarrhea
	Nausea
	Vomiting
Generalized	Anaphylaxis
Respiratory	Asthma
	Laryngeal edema
	Rhinitis

Table 1-1. Symptoms associated with IgE mediated allergic reactions (Adapted from Taylor and Baumert, 2012)

ii. Allergenic Food Proteins

Allergenic food proteins are classified based upon their route of sensitization as either class 1 or class 2 allergens (Sampson, 2003).

Class 1 allergens are oral allergens and sensitize via the gastrointestinal tract (e.g. peanut, egg, milk) (Han et al., 2012; Valenta et al., 2015). Oral allergens are able to elicit systemic reactions, stable towards heat and acid treatments, resistant to proteolytic degradation, and generally water-soluble (Breiteneder and Ebner, 2000; Sampson, 1999,

2003). Class 1 allergens are able to both sensitize and elicit allergic reactions, and considered as 'complete food allergens' (Wang, 2009). Class 2 allergens or inhalant allergens are the result of sensitization via the respiratory tract and the primary cause of oral-allergy syndrome (OAS). (Nowak-Wegrzyn, 2007). Exposure to pollen protein(s) leads to cross-sensitization to class 2 food allergens, due to high homology between pollen and class 2 allergens (e.g. apple, peach, celery) (Han et al., 2012; Nowak-Wegrzyn, 2007; Sampson, 2003). Class 2 food allergens are unable to sensitize, but capable of eliciting allergic reaction due to cross-reactivity (Wang, 2009; Yagami et al., 2000). The class 2 allergens are highly susceptible to heat and proteolysis, degrading conformational epitopes, making these proteins difficult to isolate and study (Nowak-Wegrzyn, 2007; Sampson, 2003). Symptoms of OAS are localized to the oropharyngeal area presenting mild symptoms, due to rapid protein degradation by heat and digestive enzymes in oral mucosa (Nowak-Wegrzyn, 2007; Vieths et al., 2002).

In the United States, an estimated 7.6% of children and 10.8% of adults are diagnosed with a food allergy (Gupta et al., 2018, 2019). In children, the most common allergies are peanut, milk, shellfish, and tree nuts (Gupta et al., 2018). The most common allergies in adults are to shellfish, peanut, milk, tree nut, and fish (Gupta et al., 2019). Peanut proteins are one of the most extensively studied food allergens due to their high prevalence rates in both children and adults. Currently, an estimated 2.2% of children and 1.8% of adults are diagnosed with a peanut allergy, and nearly all individuals are sensitized to multiple peanut proteins (Ara h 1, 2, 3, or 6) (Gupta et al., 2018, 2019).

iii. Diagnosis of Food Allergy

To properly diagnose an IgE mediated food allergy, individuals should seek clinical diagnosis by a physician, as self-diagnosis is often unreliable (Taylor and Hefle, 2002). The combination of rapid symptom onset, symptoms originating from multiple organs, and anaphylaxis strongly suggests an IgE mediated reaction (Chinthrajah et al., 2015). In order to establish a true food allergy, physicians must confirm presence of IgE to the suspected food or ingested material (Taylor and Hefle, 2002).

Food allergies are generally diagnosed using a combination of tools, including patient clinical history, oral food challenges, and laboratory tests (skin, blood tests) (Chinthrajah et al., 2015; Sicherer and Sampson, 2010). Understanding the patient clinical and dietary history is an important component in establishing the implicated foods in relation to the time of ingestion and onset of symptoms (Chinthrajah et al., 2015). Physicians utilize several laboratory diagnostic methods for food allergy including skin prick tests (SPT), analysis of serum, and component resolved diagnostic (CRD) tests (Chinthrajah et al., 2015). Various factors can impact SPT accuracy including the protein extract composition, skin prick site (e.g. arm, back), elapsed time before reading results, variability in result measurements, and individual patient differences (Bernstein et al., 2008; Sicherer and Wood, 2013). Determination of specific IgE by *in vitro* serum immunoassays may also be performed (Hamilton and Franklin Adkinson, 2004). CRD diagnosis, which identifies peptides originating from an allergenic protein, are used to assess the level of IgE binding to peptides or protein fragments, and have recently been adopted as a diagnostic tool in OAS diagnoses

(Chinthrajah et al., 2015). All the diagnostic tests described (SPT, serum analyses, CRD) only demonstrate sensitization (Chinthrajah et al., 2015; Sicherer and Sampson, 2010).

Ideally, an oral food challenge is performed to truly demonstrate allergenic reactivity and confirm results of laboratory tests (Bock et al., 1988). The gold standard for diagnosing true food allergy is the double blind placebo controlled food challenge (DBPCFC) in which both participating parties (e.g. physician and subject) are unbiased to testing materials (Chinthrajah et al., 2015; Sampson, 1988). During a DBPCFC, increasing doses of an allergenic food, most often within a matrix vehicle, is fed to a patient in a blinded manner (Bock et al., 1988; Sampson et al., 2014). If a DBPCFC cannot be administered due to patient concerns, an alternative open food challenge or single-blind food challenge may be performed (Chinthrajah et al., 2015). Additional factors useful in diagnosing food allergy include a thorough understanding of family history, dietary history, suspected causative foods, elapsed time before symptoms began, the type and severity of symptoms, and if other medications or alcohol were ingested prior to elicitation of symptoms (Chinthrajah et al., 2015). Once diagnosed with a food allergy, individuals are advised to practice avoidance diets since no cure has been established (Taylor and Baumert, 2012). Avoidance diets can be difficult to manage due to potential cross-contamination of products during manufacturing, leaving the consumers to rely on adequate labeling (Taylor and Baumert, 2012).

PEANUT PROTEINS

Peanuts (*Arachis hypogaea*) are leguminous plants known for their high protein density and oil contents (Becker and Jappe, 2014; Sebei et al., 2013). Peanuts contain

25% protein by weight, are generally consumed after roasting and found in other prepared food products (e.g. peanut butter, snack products, baked confections) (Koppelman et al., 2016; Oerise et al., 1974; Sebei et al., 2013). Peanuts are the most diagnosed food allergy, with 2.5% of children diagnosed, whereas only 1.8% of adults are diagnosed with peanut allergy (Gupta et al., 2011, 2017, 2019). Furthermore, the prevalence of peanut allergy has increased in children from 0.4% in 1997 to 1.4% in 2010 (Sicherer et al., 2010). Peanuts are considered one of the most important food allergens due to its high prevalence, severity, and potency (Blanc et al., 2009; Klemans et al., 2013; Koppelman et al., 2005). Several varieties of peanuts (Virginia, Spanish, Valencia, and Runner) commonly produced in the US today, exhibit highly similar protein contents and compositions (Koppelman et al., 2016). Individual allergenicity is largely unaffected by the peanut varietal as shown by a comparison of protein profiles and IgE-binding capacities among varietals (Koppelman et al., 2016). For these reasons and public health interest, peanuts have been a major research focus of the allergen community and were chosen for this study.

I. Major Allergenic Peanut Proteins

Peanuts have been cultivated for thousands of years, originating in South America in modern day regions of Bolivia, Brazil, and Paraguay (Becker and Jappe, 2014; Gregory et al., 1980). *A. hypogaea* is the modern day cultivated peanut species expressing an allotetraploid genome and is the resultant product of two diploid ancestral species, *A. duranensis* and *A. ipaensis* (Bertioli et al., 2019; De Carvalho Moretzsohn et al., 2004; Gregory et al., 1980). The major allergenic peanut proteins of *A. hypogaea* are

seed storage proteins, whose function is to provide a reservoir of amino acids during plant growth (Dunwell et al., 2004; Müntz, 1998; Mylne et al., 2014; Shewry et al., 1995). At present, 17 peanut proteins have been identified as allergens by Nomenclature Sub-Committee (World Health Organization/International Union of Immunological Sciences), and of these, four (Ara h 1, 2, 3, and 6) are considered major allergens. Peanut proteins are classified by solubility according to Osborne fractionation (saline soluble globulins, water soluble albumins, or alcohol soluble prolamins) and sedimentation coefficient(s) (e.g. 11S, 7S, 2S) (Branlard and Bancel, 2007; Breiteneder and Ebner, 2000; Osborne, 1907). The major allergenic peanut proteins Ara h 1 and 3 are classified within the cupin superfamily, whereas Ara h 2 and 6 are classified in the prolamin superfamily (Mueller et al., 2014).

A. Cupin superfamily

Ara h 1 (7S vicilin) and Ara h 3 (11S legumin) are both classified as cupins due to their similar conserved β-barrel structures (Dunwell, 1998). The cupins characteristically exhibit a bicupin structure due to the presence of two structural domains (Dunwell et al., 2004; Mills et al., 2002). These structural commonalities confer high thermal stability and resistance to gastric digestion, a trait common among allergenic proteins (van Boxtel et al., 2008; Koppelman et al., 1999, 2010; Maleki et al., 2000). Although Ara h 1 and 3 exhibit similar structural conformations, the two proteins only share 22% sequence identity based on alignments (data not shown). Cupins characteristically denature due to extensive heating (70 – 95°C) leading to the formation of large aggregates (van Boxtel et al., 2008; Koppelman et al., 2003; Yamauch et al., 1991).

i. Ara h 1

Ara h 1, the 7S vicilin, has a monomeric molecular weight of 64.5 kDa, and readily associates into a trimer (~193.5 kDa) stabilized by non-covalent forces (e.g. electrostatic, hydrophobic, disulfide bonds) (van Boxtel et al., 2006; Burks et al., 1991; Schmitt et al., 2010). Ara h 1 is highly susceptible to heating (>80°C) leading to the forming higher order molecular structures (MW ~500–600 kDa), particularly with an increase in secondary structures (e.g. β-sheets) (van Boxtel et al., 2006; Chruszcz et al., 2011; Koppelman et al., 1999). Ara h 1 is glycosylated at residues 521 – 523, with a single glycosylation site (NAS), similar to other 7S vicilins (e.g. pea, soybean), and is readily purified with concanavalin-A by lectin binding chromatography (Van Ree et al., 2000)

Ara h 1 is translated as a pre-pro-protein and undergoes two cleavage events prior to producing mature Ara h 1 (Hurlburt et al., 2014). The signal peptide (amino acids 1-25), responsible for directing the protein to the storage vacuole, is cleaved off after protein transport (Coleman et al., 1985; Hurlburt et al., 2014). Following translocation, the *N*-terminal peptide sequence (amino acids 26-84) undergoes cleavage by proteases in the vacuole, resulting in the mature Ara h 1 protein (Aalberse et al., 2019; Hurlburt et al., 2014). Within the *N*-terminal peptide region, three IgE binding epitopes exist, with two of these determined as major epitopes (Aalberse et al., 2019; Burks et al., 1997; Wichers et al., 2004). The existence of allergenic proteins or epitopes present in other *N*-terminal regions was similarly demonstrated for English walnuts (*Jugulus regia*) (Downs et al., 2014). Aalberse (2019) recently demonstrated the *N*-terminal peptide strongly binds IgE

from allergic patient serum, thus it should be considered as a distinct allergenic molecule and sensitizing protein (Aalberse et al., 2019). The purified Ara h 1 is truncated on its *N*-terminal side, lacking this pro-peptide region (Wichers et al., 2004). Assays utilizing purified Ara h 1 will not include this *N*-terminal peptide and may therefore lead to insufficient results (Wichers et al., 2004).

ii. Ara h 3

Ara h 3 was initially identified as a 14 kDa protein however, through the use of cDNA cloning, it was later determined to exist as a 60 kDa monomer (Eigenmann et al., 1996; Rabjohn et al., 1999). The 14 kDa protein, originally identified as Ara h 3, was characterized as a breakdown product of the *N*-terminal region (Rabjohn et al., 1999). Further analysis of the amino acid sequence revealed Ara h 3 is highly homologous to other 11S globulin proteins found in soy and pea (Rabjohn et al., 1999). Ara h 3 is known to act as a trypsin inhibitor, exhibiting highly homologous *N*-terminal regions to putative trypsin inhibitor proteins (Dodo et al., 2004).

Similar to other 11S seed storage proteins, Ara h 3 associates into hexameric complexes (~360 kDa) composed of two aggregated trimers (Shewry et al., 1995). To achieve its characteristic hexameric conformation, Ara h 3 is initially translated as a 'pre-pro-globulin' and must undergo post-translational processing (Koppelman et al., 2003). The translated Ara h 3 precursor protein is sent to the storage vacuole, where it associates into trimers (Guo et al., 2008; Rabjohn et al., 1999). Endopeptidase cleavage at the flexible loop region produces acidic and basic subunits covalently linked together by a

disulfide bond (Koppelman et al., 2003). Each subunit represents a cupin domain, yielding the mature Ara h 3 protein (Piersma et al., 2005; Scott et al., 1992).

Ara h 3 exists as multiple protein fragments at 14, 25, 42, and 45 kDa under protease inhibited extraction conditions (Koppelman et al., 2003). The *N*-terminal acidic chain is present in two MW bands on a reducing gel at 12 kDa and 42-45 kDa (Piersma et al., 2005). The C-terminal, or basic chain, exists as a 25 kDa fragment (Piersma et al., 2005).

As observed in other 11S globulins, Ara h 3 exists as multiple isoforms encoded by multi-gene families (Mouzo et al., 2018; Piersma et al., 2005; Yan et al., 2005). These genes are located on one or more chromosomes and present in highly homologous gene clusters (Mouzo et al., 2018; Piersma et al., 2005; Yan et al., 2005). The many isoforms of Ara h 3 produced from multiple genes, generating a variety of gene products, and potential truncation(s) at the N- and C-termini generates a larger probability of sequence isoform variants (Bertioli et al., 2019). The contribution that each isoform makes to overall allergenicity is still unknown, due to the challenge of purifying individual protein isoforms.

Rabjohn and co-workers identified four IgE-binding epitopes in Ara h 3, all located within the acidic subunit (Rabjohn et al., 1999). Jin and co-workers demonstrated these epitopes are all solvent exposed owing to their allergenicity (Jin et al., 2009). Many IgE epitopes are not fully elucidated since many of these epitope studies have been performed on recombinant proteins, however the recombinant versions lack the inherent variation present in naturally produced Ara h 3 (Piersma et al., 2005).

B. Prolamins

The 2S albumins, Ara h 2 and 6, are considered the most potent elicitors of allergic reactions in peanuts (Flinterman et al., 2007; Klemans et al., 2013; Peeters et al., 2007). These proteins exhibit strong similarity, sharing 59% sequence identity, and similar tightly bound structures (Koppelman et al., 2005). Together, the 2S albumins were determined as potent allergens by histamine degranulation studies and rat basophil leukemia cell-based immunoassays (Blanc et al., 2009;; Zhuang and Dreskin, 2013). Ara h 2 and Ara h 6 are considered strong clinical predictors of peanut allergy as determined by *in vitro* measurements (Klemans et al., 2013; Koppelman et al., 2005; McDermott et al., 2007). When combined for diagnostic testing, Ara h 6 and Ara h 2 can effectively predict peanut allergy with 85% specificity (Koid et al., 2013).

In peanuts, the 2S albumin proteins are present as monomers, in contrast to most other 2S albumins, which exist as heterodimers (Burks et al., 1992; Shewry et al., 1995; Suhr et al., 2004). Both Ara h 2 and Ara h 6 exhibit similar structures of tightly coiled helices stabilized by four or five disulfide bonds, respectively, which imparts high stability against thermal treatments and gastric digestion (Hazebrouck et al., 2012; Mueller et al., 2011; Sen et al., 2002). Digestibility studies using both gastric and intestinal proteases demonstrate the peanut 2S albumins retain their internal secondary structures (Apostolovic et al., 2016; Astwood et al., 1996; Koppelman et al., 2010; Lehmann et al., 2006; Sen et al., 2002; Suhr et al., 2004). Another digestibility study demonstrated the N- and C-termini of peanut 2S albumins are the most susceptible to proteolysis (Apostolovic et al., 2016). Both Ara h 2 and 6 are capable of generating

large, stable fragments after digestion with proteolytic enzymes pepsin and chymotrypsin (Apostolovic et al., 2016). These digestive resistant peptides (DRP) demonstrated similar IgE binding and secondary structures to the intact protein (Apostolovic et al., 2016).

i. Ara h 2

Ara h 2 exists in two isoforms, Ara h 2.01 (16.3 kDa) and 2.02 (18 kDa) (Chatel et al., 2003). The slightly larger isoform, Ara h 2.02, has a 12 amino acid insert beginning at residue 71 (Apostolovic et al., 2016; Chatel et al., 2003). Comparison of IgE binding between the two isoforms indicated increased IgE binding to Ara h 2.02, which has an additional copy of the peptide sequence DPYSPS (Chatel et al., 2003; Hales et al., 2004). This inserted peptide sequence has been attributed to increased IgE binding (Albrecht et al., 2009; Stanley et al., 1997).

As previously described, Ara h 2 exists as a tightly bound coil, stabilized by four disulfide bonds conferring thermal and digestion resistance (Mueller et al., 2011). Several studies have reported the presence of digestion-resistant peptides (DRP) from Ara h 2 (Koppelman et al., 2010; Sen et al., 2002). Sen and co-workers described native Ara h 2 produces a stable 10-kDa peptide fragment after *in vitro* digestion with three gastroduodenal proteases (pepsin, chymotrypsin, trypsin) (Sen et al., 2002). These larger peptide fragments are capable of eliciting allergic reactions, as demonstrated by IgE binding to DRPs (Apostolovic et al., 2016), and have been detected in human serum and breast milk after consumption (Baumert et al., 2009; Bernard et al., 2014).

Ara h 2 is a strong clinical predictor for peanut allergy, and shows frequent post-translational modifications of proline residues, attributed to its potency (Blanc et al., 2009; Kulis et al., 2012; Li et al., 2010). Bernard and co-workers demonstrated increased IgE binding to hydroxylated prolines in Ara h 2 (Bernard et al., 2015). They reported IgE binding was strongest to hydroxylated prolines of linear Ara h 2 epitopes, whereas IgE binding was noticeably weaker to peptides lacking hydroxyproline residues (Bernard et al., 2015). The presence of hydroxyproline residues with increased IgE binding indicates the relevance of Ara h 2 as a highly potent and severe allergen with consequent usefulness for clinical diagnostics (Bernard et al., 2015).

ii. Ara h 6

Ara h 6 (14.5 kDa) is highly homologous to Ara h 2 both in sequence (59% homology) and structure, with nearly identical α-helical regions (Apostolovic et al., 2013; Koppelman et al., 2005). Five disulfide bridges stabilize the tightly coiled helices (Suhr et al., 2004). Two isoforms exist for Ara h 6, with only minor sequence differences (Bernard et al., 2007). Ara h 6 exhibits similar seroprevalence to Ara h 2, with both thought to elicit severe reactions in sensitized individuals (Codreanu et al., 2011; Flinterman et al., 2007). Both *in vivo* (SPT, basophil degranulation) and *in vitro* (IgE immunoblots) tests were used to establish Ara h 6 as a major allergen, which was not initially considered (Koppelman et al., 2005; Peeters et al., 2007). Presence of a post-translationally cleaved Ara h 6 has demonstrated clinical reactivity to a polyclonal antibody when extracted from four different peanut cultivars (de Jong et al., 2018). Post translational cleavage for Ara h 6 differs compared to other plant species in which most

2S albumins are synthesized as one precursor peptide and then cleaved into two chains and exist as heterodimer (de Jong et al., 2018; Shewry et al., 1995).

II. Minor Peanut Allergens

There are several other peanut proteins which are considered minor peanut allergens, and important clinically, but not as prevalent as the major peanut allergens previously described. In an allergic population, minor allergens are defined as proteins which bind serum IgE in less than 50% of allergic subjects (Matricardi et al., 2016; Mueller et al., 2014).

All other IUIS identified peanut allergens (Ara h 5, Ara h 7-17) are considered minor allergens (Mueller et al., 2014). Ara h 7 and Ara h 9 are classified as prolamins (Kleber-Janke et al., 1999). Ara h 7 shares a similar structure to Ara h 2 and 6 whereas Ara h 9 is an nsLTP (non-specific lipid transfer protein) (Kleber-Janke et al., 1999; Krause et al., 2009). Ara h 5 and 8 are categorized as Class 2 allergens, causing cross reactivity with inhalant allergens (Becker et al., 2018). Ara h 5 is similar to Bet v 2, whereas Ara h 8 is cross-reactive to Bet v 1 (Asarnoj et al., 2012; Mittag et al., 2004). Ara h 8 has hydrophobic regions similar to Bet v 1, known to bind lipids and prevent protein digestion (Petersen et al., 2014).

Ara h 10 and 11 are oleosins existing in association with oil bodies (Pons et al., 2002; Schwager et al., 2015). These proteins are generally present as multimeric conformations, either as dimers or oligomers (Schwager et al., 2015). Oleosins are

generally underrepresented in commercially prepared extracts which generally use aqueous extractions (Schwager et al., 2015).

The defensins, Ara h 12 and 13, are cysteine rich proteins responsible for protection against pathogenic fungi, and exist as dimers (Sagaram et al., 2011).

DIGESTION, ABSORPTION, AND TRANSPORT PROCESSES OF PROTEINS

Unlike most dietary proteins, which are thought to be degraded into individual amino acids, di-, or tripeptide fragments; allergenic food proteins may remain largely intact as large peptide or protein fragments after gastric digestion (Chambers et al., 2004; Wickham et al., 2009). These partially digested macromolecules are transported across the gut epithelium and enter into circulation where they are capable of eliciting an allergic reaction in sensitized individuals (Miner-Williams et al., 2014). However, these specific transport processes for larger allergenic protein fragments are less understood due to a lack of *in vivo* detection methods for allergenic food proteins in human blood.

Evidence of intact, or largely intact, dietary food proteins entering the bloodstream, particularly allergenic proteins, has been reported for ovalbumin, β-lactoglobulin, and the peanut 2S albumin proteins, Ara h 2 and 6 (Baumert et al., 2009; Husby et al., 1985, 1986; JanssenDuijghuijsen et al., 2017; Paganelli and Levinsky, 1980). Initial investigations focusing on the uptake of allergenic proteins indicated some of the ingested protein (ovalbumin) remained intact and detectable by immunoassay following size separation chromatography (Husby et al., 1985). Likewise, recent investigations focused on peanut, a highly prevalent and potent food allergen, have

demonstrated detectable peanut proteins, Ara h 2 and 6, in human serum by immunoassays (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017). Immunoassay detection methods are challenging due to the low abundance of target analytes and potential interactions with other matrix proteins and consequently, may not be the most suitable method for *in vivo* protein detection due to their limited analytical targets and dependence on antibody-antigen binding (JanssenDuijghuijsen et al., 2017). However, improvements to *in vivo* detection methods has been challenging due to the unknown state of allergenic proteins after consumption, gastrointestinal digestion, and transport *in vivo*. Here, we will examine the properties of allergenic proteins after consumption and the process of transportation

I. Digestion of Dietary Proteins

Dietary protein, a vital macromolecule for human health and nutrition, is mostly digested in the gastrointestinal tract for rapid nutrient absorption (Erickson and Kim, 1990; Miner-Williams et al., 2014). Contrastingly, allergenic food proteins, which are highly resistant to digestion, remain intact or partially intact after gastric digestion (Astwood et al., 1996; Bannon, 2004). The purpose of digestion is to reduce food particle size, primarily within the oral cavity and stomach, allowing the release and absorption of vital nutrients into systemic circulation (Kong and Singh, 2008). As a result, digested dietary protein is readily absorbed for nutritional and metabolic processes (Kong and Singh, 2008).

Gastrointestinal digestion is broadly described in three phases, the (1) intraluminal phase, (2) small intestinal phase, and (3) transport and absorption phase

(Erickson and Kim, 1990). The first stage of digestion begins in the mouth, the beginning of the digestive tract, where food is mixed with saliva and reduced in particle size (Kong and Singh, 2008). Saliva, a mixture of water, electrolytes, mucus, and enzymes, is responsible for solubilizing the food bolus for transport to the stomach (Untersmayr and Jensen-Jarolim, 2006). The oral mucosa is the first potential site of antigen uptake during digestion (Dirks et al., 2005; Untersmayr and Jensen-Jarolim, 2006). Dirks et al., 2005 evaluated blood samples by histamine release assays from subjects who were instructed to chew raw peanuts but to not swallow the food material. The collected blood samples triggered mast cell activation and histamine release, supporting protein absorption can occur as early as the oral mucosa in the digestive tract (Dirks et al., 2005).

The food bolus then travels through the esophagus and into the stomach, where it encounters gastric juices composed of mucus, hydrochloric acid (HCl), and other proteases (Untersmayr and Jensen-Jarolim, 2006). The bicarbonate containing mucus layer is responsible for providing protection to the gastric mucosal lining (Allen and Flemström, 2005). HCl reduces the stomach pH, activating pepsinogen autocatalysis producing active pepsin, the primary stomach protease (Erickson and Kim, 1990).

After partial protein digestion in the stomach, the food bolus enters the small intestine where pancreatic digestion occurs (Erickson and Kim, 1990). Here, alkaline pancreatic juices containing proteolytic enzymes (trypsin, chymotrypsin, elastase, and carboxypeptidase) are secreted as inactive zymogens from the pancreas (Erickson and Kim, 1990). The acidic food bolus activates the pancreatic enzymes by decreasing the pH of the alkaline small intestine to pH 6.0 – 6.5, while simultaneously deactivating the

gastric proteases (Erickson and Kim, 1990; Rune and Viskum, 1969). Trypsin is a key digestive enzyme activated by action of enteropeptidase, an enzyme located in the duodenal enterocytes (Erickson and Kim, 1990; Rinderknecht, 1993). Activation of trypsin initiates a signaling cascade activating the remaining pancreatic enzymes with different cleavage sites (Matthews, 1975). The pancreatic digestion phase is fundamental for production of short polypeptides and free amino acids necessary for rapid protein absorption (Erickson and Kim, 1990). During the last portion of gastric digestion, the intestinal brush border peptidases act upon remaining peptides and produce a mixture of di- and tripeptides, as well as free amino acids for transport across the intestinal epithelium (Erickson and Kim, 1990).

Gastric digestion and transit time takes approximately two hours (Untersmayr and Jensen-Jarolim, 2006). The rate of digestion is influenced by several factors including the food matrix, chemical and physical properties of the food(s), and other physiological processes occurring within the GI tract (Kong and Singh, 2008).

II. Dietary Protein Absorption

Protein absorption has been studied using both *in vitro* and *in vivo* assays. Several *in vitro* studies have used Ussing chambers, requiring small sections of intestinal tissue (*ex vivo*), mimicking small intestinal environment and absorption (Reitsma et al., 2014). Cell lines, including Caco-2 and HT-29, have also been utilized in many protein absorption studies (Reitsma et al., 2014). Ideally, *in vivo* studies are the most relevant method for studying protein absorption, however few studies focused on *in vivo* protein uptake (Husby et al., 1985, 1986; Paganelli and Levinsky, 1980). The gastrointestinal

tract, where a majority of absorption occurs consists of several systems working in sequence to absorb dietary molecules including proteins (Samadi et al., 2018). Here, we review the gastrointestinal environment and absorption mechanisms for both sensitized and non-sensitized individuals.

A. Organization of the Gastrointestinal Barrier

In order to understand gastric transport and behavior of allergenic proteins, we must first understand the complex structure of the small intestine, where a majority of digestion occurs (Kong and Singh, 2008). The intestinal epithelium consists of multiple structured, yet differentiated layers working collectively to prevent uptake of harmful antigens whilst allowing absorption of nutritive molecules (Gigante et al., 2011; Heyman, 2005; Reitsma et al., 2014).

In healthy individuals, a tight monolayer of intestinal epithelial cells is formed, preventing passage of undesired molecules (Samadi et al., 2018). The surface of the epithelium is highly folded and organized into villi and crypts resulting in an increased surface area of the intestinal lumen (Peterson and Artis, 2014; Samadi et al., 2018). Four types of intestinal epithelial cells (IECs) are generated from intestinal epithelial stem cells including absorptive enterocytes, mucus producing goblet-cells, anti-microbial secreting Paneth cells, and enteroendocrine cells (Bevins and Salzman, 2011; Peterson and Artis, 2014; Samadi et al., 2018). Upon cell differentiation, the enterocytes form a barrier preventing passage of larger molecules by simultaneous producing tight junction proteins and additional membrane proteins connecting enterocytes on their apical sides (Heyman, 2005; Samadi et al., 2018). Other compounds including secretive IgA, mucus, and α-

defensins, further prevent the absorption of harmful molecules across the intestinal epithelium (Samadi et al., 2018). Figure 1-2 illustrates the organization of intestinal epithelium (Samadi et al., 2018).

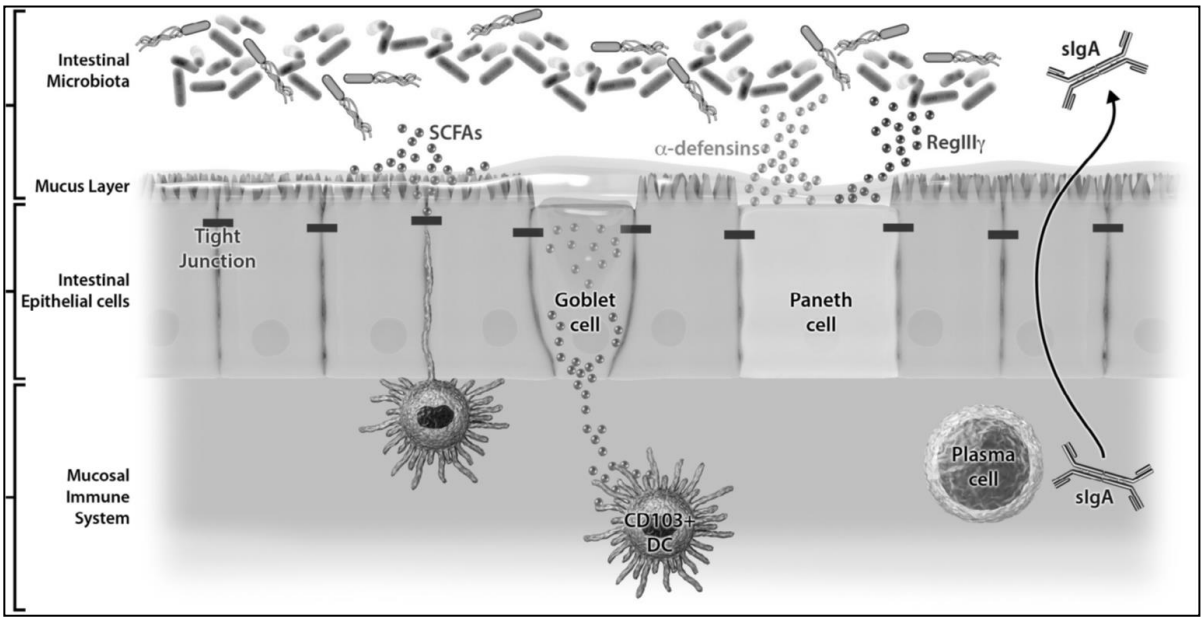


Figure 1-2. Structure of gastrointestinal epithelium. Adapted from Samadi et al., 2018

The mucosal immune system is located immediately below the intestinal epithelium and contains a significant number of immune cells (Samadi et al., 2018). It is the primary organ system involved in oral tolerance acquisition and the development of food allergy (Berin and Shreffler, 2016). The intestinal epithelial cells is responsible for regulating intestinal permeability and contributes significantly to mucosal immune responses (Samadi et al., 2018). The gastrointestinal epithelium functions to protect the mucosal immune system from harmful substances from the external environment (Berin and Sampson, 2013). Within the intestinal mucosca, the antigen-presenting cells and

macrophages respond to microbiota and other molecules signaling secretion of immunoregulatory cytokines (Berin and Sampson, 2013).

The mucosal immune system is made up of several immune cells including, CD4+, CD8+ regulatory and effector T cells, B-cells (antibody secreting), dendritic cells, macrophages, and eosinophils (Berin and Sampson, 2013). A key challenge of the mucosal immune system is differentiating between harmful and harmless molecules during protein uptake (Berin and Sampson, 2013)

B. Transport of Digested Proteins, Peptides, and Amino Acids

The route of allergenic protein transport and introduction into the immune system may play key roles in sensitization. Previous *in vitro* studies have utilized cell culture lines, model digestive systems, or animal models to assess allergenic protein absorption, but the most effective way to study protein absorption is by *in vivo* studies (Reitsma et al., 2014). The sensitization route and initial protein exposure may therefore influence reaction severity, the probability of sustained sensitization, or tolerance acquisition.

Following gastrointestinal digestion, 70% of proteins exist as small oligopeptides, with the remainder present as free amino acids (30%). (Binder and Reuben, 2009; Goodman, 2010). Free amino acids and peptides are transported from the small intestinal to the portal blood via absorptive enterocytes (Erickson and Kim, 1990; Goodman, 2010). Free amino acids are transported via multiple transport systems, with most amino acid uptake utilizing active transport systems (Schultz and Curran, 1970). The amino acid transport systems have broad and shared substrate specificities, allowing for the transport

of multiple amino acids while utilizing several transport systems for individual amino acids (Binder and Reuben, 2009; Goodman, 2010). In instances when amino acid concentrations are high in the small intestine, and not requiring active transport, facilitated or simple diffusion processes are additionally used (Erickson and Kim, 1990; Goodman, 2010).

Oligopeptides are transported independent of the amino acids, occurring primarily through the PEPT1 transporter, the major peptide transporter (Devlin, 2006). Amino acids are more efficiently absorbed when transported in the form of peptides (Erickson and Kim, 1990). PEPT1 is an active H+ coupled transport process, utilizing the electrogenic difference between the luminal brush border (pH 6) and the enterocytic cytoplasm (pH 7) facilitating oligopeptide uptake (Steel and Hediger, 1998). The PEPT1 has broad substrate specificity for di- and tri-peptides capable of transporting ~400 dipeptides and 8000 tripeptides (Daniel, 2004; Goodman, 2010). Di- and tri-peptides absorbed by enterocytes at the brush border membrane, are hydrolyzed by intracellular peptidases (Erickson and Kim, 1990; Goodman, 2010). However, Shimizu et al., 2004 estimated approximately 10% of proteins are able to traverse the epithelial barrier intact using various peptide transport mechanisms (Shimizu, 2004).

C. Protein Transport Routes

In non-sensitized, healthy individuals, the absorption of dietary proteins occurs via the absorptive enterocytes (i.e. transcytosis), where proteins travel from the intestinal lumen to the portal blood (circulation) (Reitsma et al., 2014).

Sensitized individuals, who exhibit immune-mediated responses to allergenic proteins, may demonstrate different protein absorption mechanisms. Para-cellular transport and transport via mast cells (transcellular transport) are the key pathways for protein transport in sensitized individuals, concluded from a thorough literature review conducted by Reitsma et al., 2014. Para-cellular transport is regulated by tight junctions allowing only small, generally hydrophilic compounds to be absorbed (Heyman, 2005). The integrity of tight junctions in sensitized individuals is reduced due to abundance of mast cells, which increases the amount of absorbed intact dietary protein (Berin et al., 1998). Proteins transported via the para-cellular route do not encounter lysosomal degradation in the enterocyte, and remain intact (non-degraded) after transport across the gut epithelium (Shimizu, 2004).

Transcellular transport involves a variety of mechanisms including carrier-mediated transport, endocytosis, or passive diffusion (Reitsma et al., 2014). Carrier-mediated transport for proteins has not been well described in the literature, however carrier-mediated transport of peptides (di- and tri-peptides) and amino acids are present in the small intestine and were previously discussed. Endocytosis, the primary transcellular pathway used, involves several intestinal cell types including enterocytes, M cells, and mast cells (Reitsma et al., 2014). During endocytosis, proteins are endocytosed by a specific cell (e.g. enterocyte) and transported across the intestinal barrier in small vesicles (Reitsma et al., 2014). During transport, proteins are degraded by lysosomes, although a minor amount of protein may remain intact or partially intact (Reitsma et al., 2014; So et al., 2000).

III. Normal Responses to Food Proteins and Oral Tolerance

During the normal course of digestion, nearly all food proteins (90%) are digested into peptides or constituent amino acids and do not cause immunologically mediated responses (Heyman and Desjeux, 1992). The remaining un-degraded proteins (~10%) cross the intestinal epithelium intact, as evidenced by detection of allergenic food proteins in serum (Baumert et al., 2009; Husby et al., 1985; JanssenDuijghuijsen et al., 2017; Shimizu, 2004). In healthy individuals, these intact antigens are sampled by immune cells after digestion, ultimately resulting in a state of acquired immune tolerance (Chehade and Mayer, 2005). This acquired immune tolerance (e.g. oral tolerance) is caused by the production of IgG, IgM, or IgA antibodies that bind intact dietary proteins (Chehade and Mayer, 2005; Ko and Mayer, 2005). Oral tolerance is immunologically defined as clonal anergy, a process resulting in no active response of the immune system (Ko and Mayer, 2005).

A failure of oral tolerance results in development of IgE antibodies and a state of sensitization (Chehade and Mayer, 2005; Pelz and Bryce, 2015; Wambre and Jeong, 2018). In infants, whose immune systems are immature, the exposure to digestively stable allergenic food proteins has been suspected to cause sensitization (Sicherer and Sampson, 2006). However, recent evidence suggests early introduction of allergenic foods could prevent the development of food allergies in children (Du Toit et al., 2015).

Acquiring oral tolerance, or loss of sensitivity to a particular food, mitigates a food allergy, and some individuals may even naturally acquire tolerance over time and eventually outgrow an allergy (Wood, 2003). Tolerance acquisition is dependent on the causative allergen but most often, children outgrow allergies to milk, wheat, soy, and egg

(Wood, 2003). Allergy to peanuts, tree nuts, and shellfish are generally not outgrown and sustained into adulthood (Chinthrajah et al., 2015). IgE antibodies directed against conformational epitopes, such as in milk and egg proteins, are generally outgrown (Sicherer and Sampson, 2010). Whereas, IgE antibodies directed against linear or sequential epitopes, such as in peanuts or tree nuts, are associated with persistent allergy (Järvinen et al., 2001). Several immunological factors are involved in the process of tolerance acquisition, however the complete mechanism is not yet fully understood (Ko and Mayer, 2005; Strobel and Mowat, 1998)

The state of an antigen during absorption, either as a soluble or an intact particulate (insoluble) molecule influences tolerance acquisition (Ko and Mayer, 2005). Soluble antigens are more tolerated than particulate antigens, as particulate antigens have been shown to induce an immune response (Brandtzaeg, 2002; Sampson, 1999). Encapsulated ovalbumin, a particulate antigen, when exposed only at level of the gut associated lymphoid tissue (GALT), induced an immunologically mediated response, whereas soluble ovalbumin protein ingested orally did not induce an immune mediated response (Jain et al., 1996a , 1996b). Soluble antigens are largely absorbed by IECs, which are present in high abundance in the intestinal epithelium (Ko and Mayer, 2005). M cells, located above Peyer's patches, absorb particulate antigens through expressed surface receptors for particulate antigens (Ko and Mayer, 2005). Dendritic cells (DCs), localized to the Peyer's patches, intestinal lamina propria, and mesenteric lymph nodes, act as key antigen presenting cells (APCs) capable of direct antigen sampling by intercalating IEC's, without disrupting the epithelial barrier (Ko and Mayer, 2005; Rescigno et al., 2001). The overall factors contributing to antigenic protein uptake and in

acquiring tolerance between sensitized and non-sensitized individuals is largely unknown.

Food allergic individuals may participate in immunotherapy protocols to re-establish oral tolerance and achieve desensitization (Wang and Sampson, 2011). Several types of immunotherapy protocols have been developed including oral immunotherapy (OIT), sublingual immunotherapy (SCIT), and epicutaneous immunotherapy (EPIT) (Koppelman et al., 2019; Sicherer and Sampson, 2010; Wang and Sampson, 2011). Gaining an understanding of how individuals are sensitized will help the development of immunotherapy protocols.

IV. *In vivo* Detection of Allergenic Food Proteins

In vivo detection of dietary protein, particularly those classified as major allergens, has been reported in serum, saliva, and breast milk, for several allergenic proteins including ovalbumin, β-lactoglobulin, and the peanut 2S albumin proteins (Ara h 2 and 6) (Baumert et al., 2009; Husby et al., 1985; JanssenDuijghuijsen et al., 2017; Paganelli and Levinsky, 1980). Initial investigations evaluated ovalbumin for dietary protein uptake and indicated a portion of the ingested protein remained intact and detectable by immunoassay following size separation chromatography (Husby et al., 1985).

In many cases, detection by ELISA methods is challenging due to the low abundance of target analytes (allergenic protein) and potential interactions with other endogenous serum proteins (JanssenDuijghuijsen et al., 2017). As a result,

immunoassays may not be the most suitable choice for *in vivo* protein detection due to their limited specificity and reliance on antibody-antigen binding. Immunoassays are the most commonly used analytical method for detection of allergenic food proteins and rely on antibody recognition of epitopes present on an allergenic protein (Immer and Lacorn, 2015). However, if antigenic epitopes are unable to be detected, this could lead to under estimations of the true allergenic protein concentration *in vivo*. Recent evidence has described *in vivo* interactions occurring between peanut proteins (e.g. antigen) and endogenous serum immunoglobulins (IgG) preventing recognition of the antigenic epitopes by immunoassay capture antibodies (JanssenDuijghuijsen et al., 2017). There is an imperative need for improved analytical methods, which are not reliant on antigen-antibody binding and offers comparable sensitivity to immunoassays.

PRINCIPLES OF MASS SPECTROMETRY

The growing importance of MS in clinical medicine has enabled investigations of protein biomarkers and other biologically relevant sera proteins (Gillette and Carr, 2013). Recently, MS has been implemented in food allergen detection, providing an alternative method for protein quantification in complex food matrices (Monaci et al., 2018). Food allergens are typically present in low concentrations, but remain readily detectable by LC-MS/MS methods (Monaci et al., 2018). Immunoassays are the most commonly established technique for detection of allergenic food proteins, however recent advancements in serum and food allergen proteomics, as well as mass spectrometers (e.g. sensitivity, resolution, mass accuracy, duty cycle) has enabled clinical investigations of *in*

vivo allergenic proteins (Gillet et al., 2016; Immer and Lacorn, 2015; Mann and Kelleher, 2008).

In bottom-up proteomics, mass spectrometers detect protein-derived peptides from biological samples by measuring the mass-to-charge ratio (m/z) of gas phase ions with high sensitivity (Lane, 2005). Historically, MS methods are used to determine the relative abundance and absolute abundance of proteins, with the former being the most predominantly employed method (Lane, 2005; Monaci et al., 2018). Recently, many quantifiable, relative and absolute, targeted methods (e.g. parallel reaction monitoring (PRM), multiple reaction monitoring (MRM) have been developed and implemented in food allergen detection (Monaci et al., 2018).

I. MS Instrument Design

The overall purpose of a mass spectrometer is to identify the mass-to-charge ratios of ions in the gas phase (Aebersold and Mann, 2003; Savaryn et al., 2016). In principle, mass spectrometers consist of three primary components including the ion source, mass analyzer, and detector (Aebersold and Mann, 2003). The analytes are ionized, entering into the gas phase through various ionization methods including electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and matrix-assisted laser desorption ionization (MALDI) (Aebersold and Mann, 2003; Lane, 2005). The mass analyzer, the key component of any MS platform, measures the mass-to-charge (m/z) ratio of gaseous phase ions (Aebersold and Mann, 2003). Lastly, the detector is responsible for determining the abundance of each m/z identified by the mass analyzer (Aebersold and Mann, 2003).

Most often, the type of mass analyzer is used to describe, or often name, various MS platforms (Savaryn et al., 2016). Several types of mass analyzers have been developed including linear ion traps (LIT), quadrupole mass filters (QMF), quadrupole ion traps (QIT), high-resolution Orbitraps, time-of-flight (TOF), and Fourier transform ion cyclotron resonance (FT-ICR) (Aebersold and Mann, 2003; Savaryn et al., 2016). For the purposes of this review, we will focus on the instrumentation of high-resolution Orbitraps and triple quadrupole (quadrupole mass filters), as these instrument types were utilized in our studies.

A. Orbitrap Instruments

The Orbitrap, the predominant instrument used for PRM experiments, offers major advancements over previous instrument platforms including higher mass accuracy (compared to ion traps), increased sensitivity and dynamic range (compared to time-of-flight), and with a smaller footprint (compared to FT-ICR) (Eliuk and Makarov, 2015). The Orbitrap is based on a Kingdon trap, where an electrostatic field is established in an enclosed can by placing a wire along its axis (Eliuk and Makarov, 2015; Kingdon, 1923). The Orbitrap itself consists of three electrodes including two outer cup shaped electrodes, which face each other and an additional spindle shaped central electrode (Eliuk and Makarov, 2015; Zubarev and Makarov, 2013). Ions enter the Orbitrap and due to the electric fields traveling in a circular manner around the central electrode creating a digitized image current (e.g. in time). The image current is processed by an ‘enhanced Fourier Transformation’ algorithm (e.g. in frequency) and ultimately converted to a mass spectrum (Michalski et al., 2011; Zubarev and Makarov, 2013).

The Q-Exactive (QE) platform is a hybrid design pairing a QMF with an Orbitrap mass analyzer (Michalski et al., 2011). The QMF allows for rapid isolation of selected m/z ions during targeted acquisitions (Michalski et al., 2011). The addition of an S-lens, immediately following injection from the ion source, improves ion filtering (Michalski et al., 2011; Zubarev and Makarov, 2013). The QE platform is equipped with a C-trap and a higher energy collisional dissociation (HCD) cell, where ion fragmentation occurs (Michalski et al., 2011). The C-trap is a large ion storage device, separate to the Orbitrap analyzer, where ions are collected in 'packets' and shuttled between the HCD cell and Orbitrap analyzer (Eliuk and Makarov, 2015; Liebler and Zimmerman, 2013).

B. Triple Quadrupole Instruments (QQQ)

Triple quadrupole (QQQ) instruments, as their name implies, consist of three quadrupoles. The first and third quadrupole act as mass filters for precursor and fragment ions, respectively, with the second quadrupole acting as the collision cell (Croote and Quake, 2016; Yost and Enke, 1978). The first and third quadrupoles are used for ion selection providing two stage mass filtering, and increased sensitivity (Lange et al., 2008). When used in MRM mode, QQQ instruments are non-scanning, and only acquire spectral data for selected transitions, improving overall sensitivity by up to two orders of magnitude compared to scanning type acquisitions (PRM, Full scan data-dependent acquisitions (DDA)) (Domon and Aebersold, 2010; Lange et al., 2008).

II. MS Data Acquisition

Two types of MS acquisitions, top-down or bottom-up, are predominantly used in proteomic studies (Chait, 2006; Monaci and Visconti, 2009). Top-down methods measure the masses of intact proteins, wherein the intact protein mass and corresponding fragment masses are measured (Catherman et al., 2014). Consequently, analytical samples do not require enzymatic digestion (Catherman et al., 2014). For the purposes of our studies, we utilized bottom-up methods, which will be the primary focus of this section. Top-down methods have been thoroughly reviewed elsewhere (Catherman et al., 2014; Toby et al., 2016).

Bottom-up methods measure digested peptide fragments after proteins undergo enzymatic digestion generating smaller peptide fragments (Yates, 1998; Zhang et al., 2013). Protein samples are reduced, alkylated, and enzymatically digested generally by trypsin (Zhang et al., 2013). Trypsin has predictable cleavage sites at C-terminal arginine (R) and lysine (K) amino acids, except after a proline (P), allowing peptide masses to be accurately determined (Aebersold and Mann, 2003; Zhang et al., 2013).

In bottom-up experiments, using tandem MS, instruments are coupled with liquid chromatography system (LC, therefore LC-MS/MS), wherein the digested peptides are partially separated by a stationary chromatographic matrix and organic mobile phase gradient (Croote and Quake, 2016). This additional level of peptide separation provides improvements to overall sensitivity and detection within a single chromatographic run (Croote and Quake, 2016). Following chromatographic separation, peptides are ionized and desorbed into the gas phase when using ESI (Catherman et al., 2014; Croote and Quake, 2016; Zhang et al., 2013).

Peptide sequences are determined by comparing the parent ions detected on the acquired mass spectra to a theoretical enzymatically digested *in silico* mass spectral database generated from protein sequences in the database (Zhang et al., 2013). The amino acid sequence of peptides is determined/confirmed from the fragment ions detected in the mass spectra (MS2) resulting from the collision-induced dissociation (CID; QTRAP) or higher-energy C-trap dissociation (HCD; QExactive) of the isolated parent ion (Gillet et al., 2016). In order for peptides to be identified, the peptide sequences must be present in the protein sequence database (Gillet et al., 2016). Peptides are identified based upon the quality of peptide-spectrum matches and the false discovery rate (Gillet et al., 2016). The false discovery rate (FDR), is calculated by comparing the acquired data against the actual database and the decoy database using same analysis conditions (Gillet et al., 2016). The decoy database may be generated by performing an *in silico* digestion of the reversed amino acid sequences from the sequence database (Gillet et al., 2016).

III. MS Acquisition Methods

MS methods vary depending on the type of experiment an investigator chooses to conduct. All instruments are able to perform discovery mode (DDA) experiments and targeted MS methods (e.g. Orbitraps), however some instruments (e.g. QQQ) are better suited for targeted experiments. Differences between instrument sensitivity, resolution, and scan speed affect overall detection in each given experiment.

A. Discovery Methods

Thousands of peptides and proteins can be identified using discovery MS and often referred to as ‘shotgun’ or ‘bottom-up’ proteomic methods (Domon and Aebersold, 2010). The sheer number of proteins identified by MS easily surpasses the number of proteins identifiable by traditional affinity assays (Domon and Aebersold, 2010). Affinity-binding assays rely on the analyte binding, which are highly specific and not easily multiplexed, presenting a significant challenge in elucidating multiple proteins in a matrix (Monaci et al., 2015).

In discovery acquisitions (DDA), peptides for experimental monitoring do not need to be pre-determined (Domon and Aebersold, 2010). DDA provides a minimally biased peptide sampling strategy wherein ions are selected for fragmentation based on abundance following TopN selection, where N is often between 10 and 20 ions (Eliuk and Makarov, 2015). However, discovery methods generally exhibit poorer overall sensitivity due to the inherent variability among protein samples and stochastic ion sampling (Gillette and Carr, 2013; Lange et al., 2008).

B. Targeted Acquisition Methods

Targeted MS methods, such as PRM and MRM, have been implemented for food allergen detection and in some *in vivo* studies focused on allergenic proteins (Hands et al., 2020; Monaci et al., 2015; Mose et al., 2019). Targeted MS methods can monitor multiple pre-determined peptides and transitions using their specific m/z values and chromatographic retention times (Lange et al., 2008; Picotti and Aebersold, 2012). In comparison to discovery MS, targeted MS provides an alternative for detection of low

abundance proteins, which may not be sampled in abundance driven discovery acquisitions (Liebler and Zimmerman, 2013). In addition, targeted MS methods are highly sensitive due to minimized background interferences (Domon and Aebersold, 2010). A well-developed targeted method encompasses pre-selected transitions, stable peptide elution times, and detectable fragment ion intensities (Domon and Aebersold, 2010). Consequently, targeted MS methods have been quickly adopted in food allergy research (e.g. clinical, industry) (Monaci et al., 2018).

i. Multiple Reaction Monitoring (MRM)

MRM is considered the most robust and reproducible method for targeted peptide detection and, has been routinely implemented for routine protein analytical measurements (James and Jorgensen, 2010). MRM methods reduce background noise due to narrow selective mass windows for transitions (Lange et al., 2008). As a result, MRM methods are quantifiable over a broad dynamic range spanning five orders of magnitude (Lange et al., 2008).

As with any targeted MS experiment, MRM methods rely entirely on suitably selected peptide targets and transitions (Rauniyar, 2015). In traditional MRM, only one fragmentation step is performed, where the parent ion is fragmented producing daughter ions (Figure 1-3) (Lange et al., 2008; Yost and Enke, 1978). During a quantifiable MRM assay 3 – 5 transitions per peptide are monitored for a given protein (Picotti and Aebersold, 2012). Quadrupole mass analyzers are lower-resolution compared to others (time-of-flight, Orbitrap), but in QQQ instruments, they offer high selectivity, sensitivity, and optimized duty cycles (Abbatiello et al., 2010; Sherman et al., 2009). MRM methods

are considered the ‘gold standard’ for quantification in targeted proteomics, and a suitable replacement for quantifiable immunoassays due to their high specificity for selected transitions (Addona et al., 2009; Lange et al., 2008).

ii. MRM Cubed (MRM^3)

In complex matrices, MRM methods may not achieve the desired limit of detection in samples with complex matrices and low analyte concentrations. Co-isolation of ions with similar m/z values and other matrix interferences may still hinder detection of low abundance compounds by using MRM alone (Ronsein et al., 2015). In certain instrument platforms, like a QTRAP instrument, a secondary fragmentation step, MRM^3 (“MRM cubed”), can improve selectivity, sensitivity and the signal-to-noise ratio (Von Bargen et al., 2013). In MRM^3, the previous fragment ion (MS2), is fragmented an additional time resulting in a secondary fragment ions (MS3) (Figure 1-3) (Korte and Brockmeyer, 2016). This additional fragmentation step has demonstrated up to a 30-fold increase in signal intensity, while significantly reducing baseline noise (Hunter, 2010; Korte and Brockmeyer, 2016). MRM^3 has been successful in quantifying proteins and peptide biomarkers from non-depleted serum or plasma (Fortin et al., 2009; Jeudy et al., 2014). The QTRAP 6500+ has the capability to perform MRM^3, wherein the third quadrupole (Q3) acts as a linear ion trap (LIT) which captures and fragments the MS2 ions (Korte and Brockmeyer, 2016). The 3rd mass analyzer has the capability to record a full scan spectrum of fragment ions or capture pre-defined m/z values for MS3 ions (Korte and Brockmeyer, 2016).

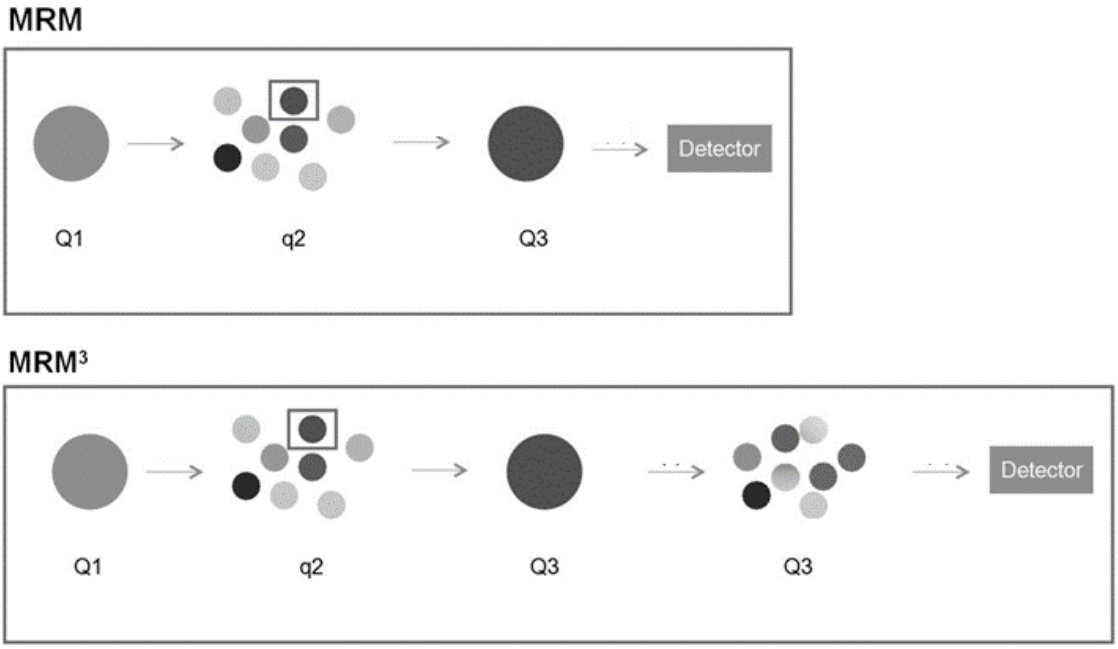


Figure 1-3. Comparison of MRM and MRM3 acquisition methods. Image adapted from Sciex (Plomley).

iii. Parallel Reaction Monitoring (PRM)

Parallel reaction monitoring methods (PRM) are an alternative targeted acquisition method providing high sensitivity, resolution, and accurate quantification with isotopically labeled peptides (e.g. heavy peptides) (Rauniyar, 2015). Due to instrument design improvements (e.g. injection, dynamic range, sensitivity, and resolution), PRM methods are frequently conducted on time-of-flight (TOF) or Orbitrap type instruments (Peterson et al., 2012). In a PRM method, the selected precursor peptides (MS1) are fragmented and all resulting fragment ions (MS2) are recorded in parallel (Figure 1-4) (Ronsein et al., 2015). As a result, pre-selection of transitions (precursor-fragment ion pairs) are not required since all transitions are monitored (Ronsein et al., 2015). PRM methods have a broad dynamic range with quantification spanning four orders of magnitude (Peterson et al., 2012). The benefit of high-resolution selection allows for distinction between isobaric ions and reductions in background noise interferences (Gallien et al., 2013; Rauniyar, 2015).

In some cases, the monitoring of multiple fragment ions from a given precursor in PRM methods are more informative than traditional MRM (Ronsein et al., 2015). By monitoring all fragment ions, the transition ratios can confirm detection of specific peptides (Peterson et al., 2012). PRM methods provide less method development time, less complex data analysis, and statistically similar quantification results (Doerr, 2012; Duncan et al., 2009; Peterson et al., 2012; Sherman et al., 2009).

To improve sensitivity in some cases, PRM methods can be further multiplexed (‘msx’), wherein the fragment ions are co-isolated and fragmented together (Sidoli et al., 2016; Wilson and Vachet, 1996). The simultaneous fragmentation of different m/z ions

reduces overall duty cycle times, increases signal-to-noise ratios and sample throughput (Sidoli et al., 2016; Wilson and Vachet, 1996). The combination of these instrument and method components allows development of various combinations of LC-MS/MS methods producing highly sensitive and robust MS acquisition methods.

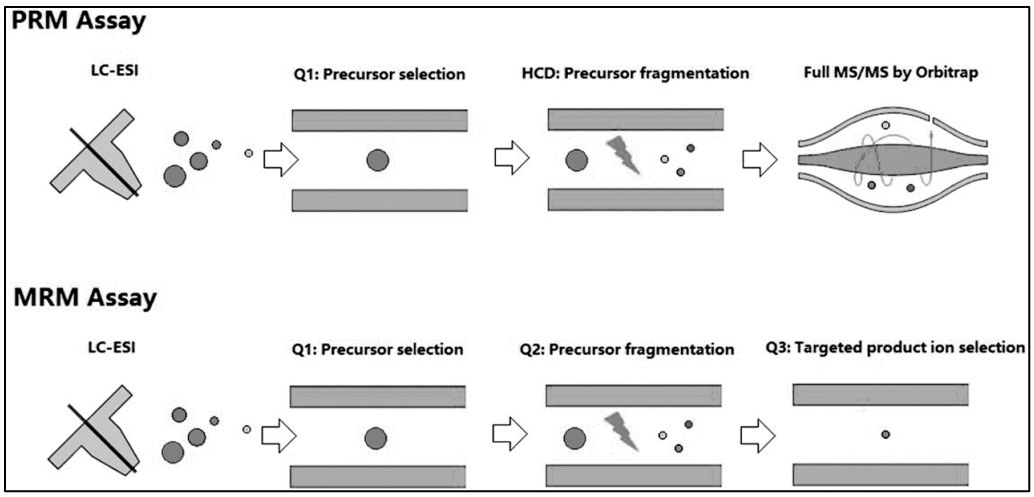


Figure 1-4. Comparison of PRM and MRM acquisition methods. Image adapted from Zhou and Yin, 2016.

IV. Quantifiable MS Methods

The aim of many MS proteomic studies is to characterize and potentially quantify targeted proteins of interest (Domon and Aebersold, 2010). Several methods have been developed for estimating protein abundances, including both relative and absolute quantification methods.

A. Relative Quantification

Relative quantification is achieved by either label-free or labeling methods of peptide or protein analytes (metabolic, chemical) (Liebler and Zimmerman, 2013; Lindemann et al., 2017). Label-free quantification is a very cost effective and high-throughput quantification strategy, allowing for comparison of a suite of proteins (Lindemann et al., 2017). Protein quantification is determined by spectral counting or by comparison of the peak area from a selected precursor-peptide signal intensity (Bantscheff et al., 2007; Lindemann et al., 2017). Alternatively, protein quantification can occur through labeling methods including both metabolic and chemical labeling (Bantscheff et al., 2007). Metabolic labeling techniques include stable isotope labeling of amino acids in cell culture (SILAC) or ^{15}N labeling (Lindemann et al., 2017). Chemical labeling based on *in vitro* chemical reactions between peptides and reagents produce a heavy labeled peptide mixture (Lindemann et al., 2017). Other chemical labeling methods includes isobaric tags for relative and absolute quantification (iTRAQ) or tandem mass tags (TMT), which label peptide or proteins following digestion (Lindemann et al., 2017; Ross et al., 2004).

B. Absolute Quantification

True quantification, the gold standard for peptide quantification, is determined by labeled synthetic peptides, such as AQUA (Absolute Quantification) peptides (Gerber et al., 2003). Synthetically prepared peptides contain isotopically labeled amino acids ^{13}C and ^{15}N resulting in a predictable mass shift, and these labeled heavy peptides behave identically to their light (unlabeled) counterparts (Gerber et al.; 2003; Lindemann et al., 2017). To quantify proteins, the concentrations of synthetically labeled peptides can be directly compared to the signal intensity of its equivalent light peptide (Lindemann et al., 2017).

MRM methods offer an absolute quantification method by the addition of isotopically labeled peptides (Lange et al., 2008).

CONCLUSIONS

Peanut allergies have increased in prevalence and become a health concern for many individuals. Peanuts are potent allergens causing severe reactions to those affected. Due to the extensive interest in peanut proteins and their allergenic properties, numerous studies have been published regarding digestive stability and uptake (Apostolovic et al., 2016; Baumert et al., 2009; JanssenDuijghuijsen et al., 2017; Koppelman et al., 2010). The illustrated digestive and thermal stability are key factors in the severity of allergic reactions to peanuts with the 2S albumins considered the most potent peanut allergens.

Several factors influence the development of food allergies, ultimately causing modification(s) in absorption of these allergenic proteins. We suspect different mechanisms are suspected to be involved in allergenic protein absorption; however, *in vivo* testing has yet to be established to further demonstrate these effects. Furthermore, diagnostic testing still relies on a DBPCFC's.

Overall improvements in both immunoassay detection methods and the understanding of allergenic protein behavior, particularly for peanut, has recently been established. However, the detection of exogenous food proteins in serum remains challenging.

Mass spectrometry is a powerful analytical tool with great sensitivity capable of detecting multiple protein targets simultaneously. Mass spectrometry has been used in other *in vivo* studies for low protein detection, but only been applied for *in vivo* detection of allergens in a handful of studies. Targeted acquisition methods, MRM, MRM^3 and PRM, offer individual advantages to the detection of low abundance proteins due to their increased selectivity and sensitivity.

A thorough understanding of the molecular and structural properties of individual peanut allergens will enhance our understanding of allergenic protein uptake and transport processes. The route(s) of protein uptake, as well the overall rate, may provide better insights into an allergic reaction, and particularly those mechanisms which differ greatly between sensitized and non-sensitized person.

REFERENCES

Aalberse, R.C., Mueller, G.A., Derksen, N.I.L., Aalberse, J.A., Edwards, L.L., Pomés, A., Lidholm, J., Rispens, T., and Briza, P. (2019). Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. Clin. Exp. Allergy cea.13554.

Abbatiello, S.E., Mani, D.R., Keshishian, H., and Carr, S.A. (2010). Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry. Clin. Chem. *56*, 291–305.

Addona, T.A., Abbatiello, S.E., Schilling, B., Skates, S.J., Mani, D.R., Bunk, D.M., Spiegelman, C.H., Zimmerman, L.J., Ham, A.J.L., Keshishian, H., et al. (2009). Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat. Biotechnol. *27*, 633–641.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207.

Albrecht, M., Kühne, Y., Ballmer-Weber, B.K., Becker, W.M., Holzhauser, T., Lauer, I., Reuter, A., Randow, S., Falk, S., Wangorsch, A., et al. (2009). Relevance of IgE binding to short peptides for the allergenic activity of food allergens. J. Allergy Clin. Immunol. *124*.

Allen, A., and Flemström, G. (2005). Gastroduodenal mucus bicarbonate barrier: Protection against acid and pepsin. Am. J. Physiol. - Cell Physiol. *288*.

Amat Par, P., Sanosa Valls, J., Lluch Pérez, M., Malet Casajuana, A., and García Calderón, P.A. (1990). Dried fruit hypersensitivity and its correlation with pollen allergy. Allergol. Immunopathol. (Madr). *18*, 27–34.

Apostolovic, D., Luykx, D., Warmenhoven, H., Verbart, D., Stanic-Vucinic, D., De Jong, G.A.H., Velickovic, T.C., and Koppelman, S.J. (2013). Reduction and alkylation of peanut allergen isoforms Ara h 2 and Ara h 6; Characterization of intermediate- and end products. Biochim. Biophys. Acta - Proteins Proteomics *1834*, 2832–2842.

Apostolovic, D., Stanic-Vucinic, D., De Jongh, H.H.J., De Jong, G.A.H., Mihailovic, J., Radosavljevic, J., Radibratovic, M., Nordlee, J.A., Baumert, J.L., Milcic, M., et al. (2016). Conformational stability of digestion-resistant peptides of peanut conglutins reveals the molecular basis of their allergenicity. Nat. Publ. Gr. *6*, 1–12.

Asarnoj, A., Glaumann, S., Elfström, L., Lilja, G., Lidholm, J., Nilsson, C., and Wickman, M. (2012). Anaphylaxis to Peanut in a Patient Predominantly Sensitized to Ara h 6. Int. Arch. Allergy Immunol. *159*, 209–212.

Asero, R., Ballmer-Weber, B.K., Beyer, K., Conti, A., Dubakiene, R., Fernandez-Rivas, M., Hoffmann-Sommergruber, K., Lidholm, J., Mustakov, T., Oude Elberink, J.N.G., et al. (2007). IgE-Mediated food allergy diagnosis: Current status and new perspectives. Mol. Nutr. Food Res. *51*, 135–147.

Astwood, J.D., Leach, J.N., and Fuchs, R.L. (1996). Stability of Food Allergens to

Digestion in Vitro. Nat. Biotechnol. *14*, 1269–1273.

Bannon, G.A. (2004). What makes a food protein an allergen? Curr. Allergy Asthma Rep. *4*, 43–46.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. Anal. Bioanal. Chem. *389*, 1017–1031.

Von Bargen, C., Rg Dojahn, J., Waidelich, D., Humpf, H.-U., and Brockmeyer, J. (2013). New Sensitive High-Performance Liquid Chromatography−Tandem Mass Spectrometry Method for the Detection of Horse and Pork in Halal Beef. J. Agric. Food Chem. *61*, 11986–11994.

Baumert, J.L., Bush, R.K., Levy, M.B., Koppelman, S.J., Nordlee, J.A., Hefle, S.L., and Taylor, S.L. (2009). Distribution of Intact Peanut Protein and Digestion-Resistant Ara h 2 Peptide in Human Serum and Saliva. J. Allergy Clin. Immunol. *123*, S268–S268.

Bayless, T.M., Rothfeld, B., Massa, C., Wise, L., Paige, D., and Bedine, M.S. (1975). Lactose and Milk Intolerance: Clinical Implications. N. Engl. J. Med. *292*, 1156–1159.

Becker, W.M., and Jappe, U. (2014). Peanut allergens. Chem Immunol Allergy *100*, 256–267.

Becker, W.M., Petersen, A., and Jappe, U. (2018). Peanut allergens: new consolidated findings on structure, characteristics and allergom. Allergol. Sel. *2*, 67–79.

Berg, J.M., Tymoczko, J.L., Stryer, L., and Stryer, L. (2002). Biochemistry (W.H. Freeman).

Berin, M.C., and Sampson, H.A. (2013). Mucosal immunology of food allergy. Curr. Biol. *23*, R389-400.

Berin, M.C., and Shreffler, W.G. (2016). Mechanisms Underlying Induction of Tolerance to Foods. Immunol. Allergy Clin. North Am. *36*, 87–102.

Berin, M.C., Kiliaan, A.J., Yang, P.C., Groot, J.A., Kitamura, Y., and Perdue, M.H. (1998). The influence of mast cells on pathways of transepithelial antigen transport in rat intestine. J. IImunology *161*, 2561–2566.

Bernard, H., Mondoulet, L., Drumare, M.F., Paty, E., Scheinmann, P., Thai, R., and Wal, J.-M. (2007). Identification of a new natural Ara h 6 isoform and of its proteolytic product as major allergens in peanut. J. Agric. Food Chem. *55*, 9663–9669.

Bernard, H., Ah-Leung, S., Drumare, M.-F., Feraudet-Tarisse, C., Verhasselt, V., Wal, J.-M., Créminon, C., and Adel-Patient, K. (2014). Peanut allergens are rapidly transferred in human breast milk and can prevent sensitization in mice. Allergy *69*, 888–897.

Bernard, H., Guillon, B., Drumare, M.-F., Paty, E., Dreskin, S.C., Wal, J.-M., Adel-Patient, K., and Hazebrouck, S. (2015). Allergenicity of peanut component Ara h 2: Contribution of conformational versus linear hydroxyproline-containing epitopes. J. Allergy Clin. Immunol. *135*, 1267-1274.e8.

Bernstein, I.L., Li, J.T., Bernstein, D.I., Hamilton, R., Spector, S.L., Tan, R., Sicherer, S., Golden, B.K., Khan, D.A., Nicklas, R.A., et al. (2008). Allergy Diagnostic Testing: An Updated Practice Parameter. Ann. Allergy, Asthma Immunol. 100, S1–S148.

Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S.C.M., Ren, L., Farmer, A.D., Pandey, M.K., et al. (2019). The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat. Genet. *51*, 877–884.

Bevins, C.L., and Salzman, N.H. (2011). Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. Nat. Rev. Microbiol. *9*, 356–368.

Binder, H., and Reuben, A. (2009). Nutrient digestion and absorption. In Medical Physiology: A Cellular and Molecular Approach, W. Boron, and E. Boulpaep, eds. (Philadelphia, PA: Saunders), pp. 949–979.

Blanc, F., Adel-Patient, K., Drumare, M.-F., Paty, E., Wal, J.-M., Bernard, H., and Bernard, H. (2009). Capacity of purified peanut allergens to induce degranulation in a functional in vitro assay: Ara h 2 and Ara h 6 are the most efficient elicitors. Clin. Exp. Allergy *39*, 1277–1285.

Bock, S.A., Sampson, H.A., Atkins, F.M., Zeiger, R.S., Lehrer, S., Sachs, M., Bush, R.K., and Metcalfe, D.D. (1988). Double-blind, placebo-controlled food challenge (DBPCFC) as an office procedure: a manual. J. Allergy Clin. Immunol. *82*, 986–997.

Boswell, A., Mundo, E.E., Firestein, R., Zhang, C., Mao, W., Gill, H., Young, C., Ljumanovic, N., Stainton, S., Ulufatu, S., et al. (2013). An integrated approach to identify normal tissue expression of targets for antibody-drug conjugates: case study of TENB2. Br. J. Pharmacol. *168*, 445.

van Boxtel, E.L., Van Beers, M.M.C., Koppelman, S.J., van den Broek, L.A.M., and Gruppen, H. (2006). Allergen Ara h 1 occurs in peanuts as a large oligomer rather than as a trimer. J. Agric. Food Chem. *54*, 7180–7186.

van Boxtel, E.L., van Den Broek, L.A.M., Koppelman, S.J., and Gruppen, H. (2008). Legumin allergens from peanuts and soybeans: Effects of denaturation and aggregation on allergenicity. Mol. Nutr. Food Res. *52*, 674–682.

Boyce, J.A., Assa'ad, A., Burks, A.W., Jones, S.M., Sampson, H.A., Wood, R.A., Plaut, M., Cooper, S.F., and Fenton, M.J. (2010). Guidelines for the diagnosis and management of food allergy in the United States: summary of the NIAID-sponsored expert panel report. J. Allergy Clin. Immunol. *126*, 1105–1118.

Brandtzaeg, P. (2002). Current understanding of gastrointestinal immunoregulation and its relation to food allergy. Ann. N. Y. Acad. Sci. *964*, 13–45.

Branlard, G., and Bancel, E. (2007). Protein Extraction from Cereal Seeds. In Plant Proteomics: Methods and Protocols, H. Thiellement, M. Zivy, C. Damerval, and V. Méchin, eds. (Totowa, NJ: Humana Press), pp. 15–25.

Breiteneder, H., and Ebner, C. (2000). Molecular and biochemical classification of plant-derived food allergens. J. Allergy Clin. Immunol. *106*, 27–36.

Burks, A., Williams, L., Connaughton, C., Cockrell, G., O'Brien, T., and Helm, R. (1992). Identification and characterization of a second major peanut allergen, Ara h 2, with use of the sera of patients with atopic dermatitis and positive peanut challenge. J. Allergy Clin. Immunol. *90*, 962–969.

Burks, A., Shin, D., Cockrell, G., Stanley, J., Helm, R., and Bannon, G. (1997). Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity. Eur. J. Biochem. *245*, 334–339.

Burks, A.W., Williams, L.W., Helm, R.M., Connaughton, C., Cockrell, G., and O'Brien, T. (1991). Identification of a major peanut allergen, Ara h 1, in patients with atopic dermatitis and positive peanut challenges. J. Allergy Clin. Immunol. *88*, 172–179.

Bush, R.K., and Hefle, S.L. (1996). Food allergens. Crit. Rev. Food Sci. Nutr. *36*, S119–S163.

Bush, R.K., and Montalbano, M. (2014). Asthma and Food Additives. In Food Allergy: Adverse Reactions to Foods and Food Additives, D. Metcalfe, H. Sampson, R. Simon, and G. Lack, eds. (John Wiley & Sons Ltd), pp. 339–345.

Catherman, A.D., Skinner, O.S., and Kelleher, N.L. (2014). Top Down Proteomics: Facts and Perspectives.

Chait, B.T. (2006). Mass spectrometry: Bottom-up or top-down? Science (80-.). *314*, 65–66.

Chambers, S.J., Wickham, M.S.J., Regoli, M., Bertelli, E., Gunning, P.A., and Nicoletti, C. (2004). Rapid in vivo transport of proteins from digested allergen across pre-sensitized gut. Biochem. Biophys. Res. Commun. *325*, 1258–1263.

Chatel, J.-M., Bernard, H., and Orson, F.M. (2003). Isolation and Characterization of Two Complete Ara h 2 Isoforms cDNA. Int. Arch. Allergy Immunol. *131*, 14–18.

Chehade, M., and Mayer, L. (2005). Oral tolerance and its relation to food hypersensitivities. J. Allergy Clin. Immunol. *115*, 3–12.

Chinthrajah, R.S., Tupa, D., Prince, B.T., Block, W.M., Rosa, J.S., Singh, A.M., and Nadeau, K. (2015). Diagnosis of Food Allergy. Pediatr. Clin. North Am. *62*, 1393–1408.

Chinthrajah, R.S., Hernandez, J.D., Boyd, S.D., Galli, S.J., Nadeau, K.C., and Stanford, À. (2016). Molecular and cellular mechanisms of food allergy and food tolerance. Mech. Allerg. Dis. *137*, 984–997.

Chruszcz, M., Maleki, S.J., Majorek, K.A., Demas, M., Bublin, M., Solberg, R., Hurlburt, B.K., Ruan, S., Mattisohn, C.P., Breiteneder, H., et al. (2011). Structural and immunologic characterization of Ara h 1, a major peanut allergen. J. Biol. Chem. *286*, 39318–39327.

Codreanu, F., Collignon, O., Roitel, O., Thouvenot, B., Sauvage, C., Vilain, A.-C., Cousin, M.-O., Decoster, A., Renaudin, J.-M., Astier, C., et al. (2011). A Novel Immunoassay Using Recombinant Allergens Simplifies Peanut Allergy Diagnosis. Int. Arch. Allergy Immunol. *154*, 216–226.

Coleman, J., Inukai, M., and Inouye, M. (1985). Dual functions of the signal peptide in protein transfer across the membrane. Cell *43*, 351–360.

Cravatt, B.F., Simon, G.M., and Yates, J.R. (2007). The biological impact of mass-spectrometry-based proteomics. Nature *450*, 991–1000.

Croote, D., and Quake, S.R. (2016). Food allergen detection by mass spectrometry: the role of systems biology. Syst. Biol. Appl. *2*, 16022.

Daniel, H. (2004). Molecular and Integrative Physiology of Intestinal Peptide Transport. Annu. Rev. Physiol. *66*, 361–384.

Dirks, C.G., Pedersen, M.H., Platzer, M.H., Bindslev-Jensen, C., Skov, P.S., and Poulsen, L.K. (2005). Does absorption across the buccal mucosa explain early onset of food-induced allergic systemic reactions? J. Allergy Clin. Immunol. *115*, 1321–1323.

Dodo, H.W., Viquez, O.M., Maleki, S.J., and Konan, K.N. (2004). cDNA Clone of a Putative Peanut (Arachis hypogaea L.) Trypsin Inhibitor Has Homology with Peanut Allergens Ara h 3 and Ara h 4 . J. Agric. Food Chem. *52*, 1404–1409.

Doerr, A. (2012). Targeting with PRM. Nat. Methods *9*, 950–950.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. Nat. Biotechnol. *28*, 710–721.

Downs, M.L., Semic-Jusufagic, A., Simpson, A., Bartra, J., Fernandez-Rivas, M., Rigby, N.M., Taylor, S.L., Baumert, J.L., and Mills, E.N.C. (2014). Characterization of Low Molecular Weight Allergens from English Walnut (Juglans regia). J. Agric. Food Chem. *62*, 11767–11775.

Duncan, M.W., Yergey, A.L., and Patterson, S.D. (2009). Quantifying proteins by mass spectrometry: The selectivity of SRM is only part of the problem. Proteomics *9*, 1124–1127.

Dunwell, J.M., Purvis, A., and Khuri, S. (2004). Cupins: The most functionally diverse protein superfamily? Phytochemistry *65*, 7–17.

Eigenmann, P.A., Burks, A.W., Bannon, G.A., and Sampson, H.A. (1996). Identification of unique peanut and soy allergens in sera adsorbed with cross-reacting antibodies. J. Allergy Clin. Immunol. *98*, 969–978.

Eliuk, S., and Makarov, A. (2015). Evolution of Orbitrap Mass Spectrometry Instrumentation. Annu. Rev. Anal. Chem. *8*, 61–80.

Erickson, R.H., and Kim, Y.S. (1990). Digestion and absorption of dietary protein.

Flinterman, A.E., Van Hoffen, E., Den Hartog Jager, C.F., Koppelman, S., Pasmans, S.G., Hoekstra, M.O., Bruijnzeel-Koomen, C.A., Knulst, A.C., and Knol, E.F. (2007). Children with peanut allergy recognize predominantly Ara h 2 and Ara h 6, which remains stable over time. Clin. Exp. Allergy *37*, 1221–1228.

Fortin, T., Salvador, A., Charrier, J.P., Lenz, C., Bettsworth, F., Lacoux, X., Choquet-Kastylevsky, G., and Lemoine, J. (2009). Multiple Reaction Monitoring Cubed for

Protein Quantification at the Low Nanogram/Milliliter Level in Nondepleted Human Serum. Anal. Chem. *81*, 9343–9352.

Gallien, S., Duriez, E., Demeure, K., and Domon, B. (2013). Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. J. Proteomics *81*, 148–158.

Gell, P., and Coombs, R. (1975). Classification of allergic reactions responsible for clinical hypersensitivity and disease. Clin. Asp. Immunol. 575–596.

Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. PNAS *100*, 6940–6945.

Gigante, G., Tortora, A., Ianiro, G., Ojetti, V., Purchiaroni, F., Campanale, M., Cesario, V., Scarpellini, E., and Gasbarrini, A. (2011). Role of gut microbiota in food tolerance and allergies. Dig. Dis. *29*, 540–549.

Gillet, L.C., Leitner, A., and Aebersold, R. (2016). Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. Annu. Rev. Anal. Chem. Is Online *9*, 449–472.

Gillette, M.A., and Carr, S.A. (2013). Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat. Methods *10*, 28–34.

Goodman, B.E. (2010). Insights into digestion and absorption of major nutrients in humans. Adv Physiol Educ *34*, 44–53.

Gregory, W.C., Krapovickas, A., and Gregory, M.P. (1980). Structure, variation, evolution, and classification in Arachis. Struct. Var. Evol. Classif. Arachis. 469–481.

Guandalini, S., and Newland, C. (2011). Differentiating Food Allergies from Food Intolerances. Curr. Gastroenterol. Rep. *13*, 426–434.

Guo, B., Liang, X., Chung, S.-Y., Holbrook, C., and Maleki, S. (2008). Proteomic Analysis of Peanut Seed Storage Proteins and Genetic Variation in a Potential Peanut Allergen. Protein Pept. Lett. *15*, 567–577.

Gupta, R., Blumenstock, J., Warren, C., Mittal, K., Kotowska, J., and Smith, B. (2017). The prevalence of nut and seafood allergies among adults in the United States.

Gupta, R.S., Springston, E.E., Warrier, M.R., Smith, B., Kumar, R., Pongracic, J., and Holl, J.L. (2011). The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics *128*, e9–e16.

Gupta, R.S., Warren, C.M., Smith, B.M., Blumenstock, J.A., Jiang, J., Davis, M.M., and Nadeau, K.C. (2018). The public health impact of parent-reported childhood food allergies in the United States. Pediatrics *142*.

Gupta, R.S., Warren, C.M., Smith, B.M., Jiang, J., Blumenstock, J.A., Davis, M.M., Schleimer, R.P., and Nadeau, K.C. (2019). Prevalence and Severity of Food Allergies Among US Adults. JAMA Netw. Open *2*, e185630.

Hales, B.J., Bosco, A., Mills, K.L., Hazell, L.A., Loh, R., Holt, P.G., and Thomas, W.R.

(2004). Isoforms of the major peanut allergen Ara h 2: IgE binding in children with peanut allergy. Int. Arch. Allergy Immunol. *135*, 101–107.

Hamilton, R.G., and Franklin Adkinson, N. (2004). In vitro assays for the diagnosis of IgE-mediated disorders. J. Allergy Clin. Immunol. *114*, 213–225.

Han, Y., Kim, J., and Ahn, K. (2012). Food allergy. Korean J. Pediatr. *55*, 153–158.

Hands, C.M., Sayers, R.L., Nitride, C., Gethings, L.A., and Mills, E.N.C. (2020). A multiple reaction monitoring method for determining peanut (Arachis hypogea) allergens in serum using quadrupole and time-of-flight mass spectrometry. Anal. Bioanal. Chem.

Hazebrouck, S., Guillon, B., Drumare, M.-F., Paty, E., Wal, J.-M., and Bernard, H. (2012). Trypsin resistance of the major peanut allergen Ara h 6 and allergenicity of the digestion products are abolished after selective disruption of disulfide bonds. Mol. Nutr. Food Res. *56*, 548–557.

Hefle, S.L., Nordlee, J.A., and Taylor, S.L. (1996). Allergenic foods. Crit. Rev. Food Sci. Nutr. *36*, 69S-89S.

Hefle, S.L., Jeanniton, E., and Taylor, S.L. (2001). Development of a Sandwich Enzyme-Linked Immunosorbent Assay for the Detection of Egg Residues in Processed Foods. J. Food Prot. *64*, 1812–1816.

Heyman, M. (2005). Gut barrier dysfunction in food allergy. Eur. J. Gastroenterol. Hepatol. *17*, 1279–1285.

Heyman, M., and Desjeux, J.F. (1992). Significance of intestinal food protein transport. J. Pediatr. Gastroenterol. Nutr. *15*, 48–57.

Hunter, C. (2010). MRM3 Quantitation for Highest Selectivity of Proteins in Complex Matrices. J. Biomol. Tech. *21*, S34–S35.

Hurlburt, B., McBride, J., Nesbit, J., Ruan, S., and Maleki, S. (2014). Purification of Recombinant Peanut Allergen Ara h 1 and Comparison of IgE Binding to the Natural Protein. Foods *3*, 642–657.

Husby, S., Jensenius, J.C., and Svehag, S.E. (1985). Passage of undegraded dietary antigen into the blood of healthy adults. Quantification, estimation of size distribution, and relation of uptake to levels of specific antibodies. Scand. J. Immunol. *22*, 83–92.

Husby, S., Jensenius, J.C., and Svehag, S.E. (1986). Passage of undegraded dietary antigen into the blood of healthy adults. Further characterization of the kinetics of uptake and the size distribution of the antigen. Scand. J. Immunol. *24*, 447–455.

Immer, U., and Lacorn, M. (2015). 10 - Enzyme-linked immunosorbent assays (ELISAs) for detecting allergens in food. In Handbook of Food Allergen Detection and Control, S. Flanagan, ed. (Woodhead Publishing), pp. 199–217.

Jain, S.L., Barone, K.S., Flanagan, M.P., and Michael, J.G. (1996a). Activation patterns of murine B cells after oral administration of an encapsulated soluble antigen. Vaccine *14*, 42–48.

Jain, S.L., Barone, K.S., and Michael, J.G. (1996b). Activation patterns of murine T cells after oral administration of an enterocoated soluble antigen. Cell. Immunol. *167*, 170–175.

James, A., and Jorgensen, C. (2010). Basic Design of MRM Assays for Peptide Quantification. (Humana Press, Totowa, NJ), pp. 167–185.

JanssenDuijghuijsen, L.M., Wichers, H.J., van Norren, K., Keijer, J., Baumert, J.L., de Jong, G.A., Witkamp, R.F., and Koppelman, S.J. (2017). Detection of peanut allergen in human blood after consumption of peanuts is skewed by endogenous immunoglobulins. J. Immunol. Methods *440*, 52–57.

Järvinen, K.-M., Chatchatee, P., Bardina, L., Beyer, K., and Sampson, H.A. (2001). IgE and IgG Binding Epitopes on α-Lactalbumin and β-Lactoglobulin in Cow's Milk Allergy. Int. Arch. Allergy Immunol. *126*, 111–118.

Jeudy, J., Salvador, A., Simon, R., Jaffuel, A., Fonbonne, C., Léonard, J.-F., Gautier, J.-C., Pasquier, O., and Lemoine, J. (2014). Overcoming biofluid protein complexity during targeted mass spectrometry detection and quantification of protein biomarkers by MRM cubed (MRM 3). Anal. Bioanal. Chem. *406*, 1193–1200.

Jin, T., Guo, F., Chen, Y. wei, Howard, A., and Zhang, Y.Z. (2009). Crystal structure of Ara h 3, a major allergen in peanut. Mol. Immunol. *46*, 1796–1804.

Johnson, P.E., Baumgartner, S., Aldick, T., Bessant, C., Giosafatto, V., Heick, J., Mamone, G., O' Connor, G., Poms, R., Popping, B., et al. (2011). Current perspectives and recommendations for the development of mass spectrometry methods for the determination of allergens in foods. J. AOAC Int. *94*, 1026–1033.

Jones, S.M., and Burks, A.W. (2008). The Spectrum of Allergic Reaction to Foods. In Food Allergy: Adverse Reactions to Foods and Food Additives, D.D. Metcalfe, H.A. Sampson, and R.A. Simon, eds. (Oxford, UK: Blackwell Publishing Ltd.), pp. 101–107.

de Jong, G.A., Jayasena, S., Johnson, P., Marsh, J., Apostolovic, D., van Hage, M., Nordlee, J., Baumert, J., Taylor, S.L., Roucairol, C., et al. (2018). Purification and Characterization of Naturally Occurring Post-Translationally Cleaved Ara h 6, an Allergen That Contributes Substantially to the Allergenic Potency of Peanut.

Kingdon, K.H. (1923). A method for the neutralizatoin of electron space charge by positive ionization at very low gas pressures. Am. Phys. Soc. *21*, 408–418.

Kleber-Janke, T., Crameri, R., Appenzeller, U., Schlaak, M., and Becker, W.M. (1999). Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. Int. Arch. Allergy Immunol. *119*, 265–274.

Klemans, R.J.B., Broekman, H.C.H.P., Knol, E.F., Bruijnzeel-Koomen, C.A.F.M., Otten, H.G., Pasmans, S.G.M.A., and Knulst, A.C. (2013). Ara h 2 is the best predictor for peanut allergy in adults. J. Allergy Clin. Immunol. Pract. *1*, 632-638.e1.

Ko, J., and Mayer, L. (2005). Oral tolerance: lessons on treatment of food allergy. Eur. J. Gastroenterol. Hepatol. *17*, 1299–1303.

Kocian, J. (1988). Lactose intolerance. Int. J. Biochem. *20*, 1–5.

Koid, A.E., Chapman, M.D., Hamilton, R.G., Van Ree, R., Versteeg, S.A., Dreskin, S.C., Koppelman, S.J., and Wü, S. (2013). Ara h 6 Complements Ara h 2 as an Important Marker for IgE Reactivity to Peanut. J. Agric. Food Chem. *62*, 206–213.

Kong, F., and Singh, R.P. (2008). Disintegration of Solid Foods in Human Stomach. J. Food Sci. *73*, R67–R80.

Koppelman, S., and Hefle, S. (2006). Detecting allergens in food (Elsevier Ltd).

Koppelman, S., Bruijnzeel-Koomen, C., Hessing, M., and de Jongh, H. (1999). Heat-induced conformational changes of Ara h 1, a major peanut allergen, do not affect its allergenic properties. J. Biol. Chem. *274*, 4770–4777.

Koppelman, S., Knol, E., Vlooswijk, R., Wensing, M., Knulst, A., Hefle, S., Gruppen, H., and Piersma, S. (2003). Peanut allergen Ara h 3: isolation from peanuts and biochemical characterization. Allergy *58*, 1144–1151.

Koppelman, S., Hefle, S., Taylor, S., and de Jong, G. (2010). Digestion of peanut allergens Ara h 1, Ara h 2, Ara h 3, and Ara h 6: A comparitive in vitro study and partial characterization of digestion-resistant peptides. Mol. Nutr. Food Res. *54*, 1711–1721.

Koppelman, S.J., de Jong, G.A.H., Laaper-Ertmann, M., Peeters, K.A.B.M., Knulst, A.C., Hefle, S.L., and Knol, E.F. (2005). Purification and immunoglobulin E-binding properties of peanut allergen Ara h 6: evidence for cross-reactivity with Ara h 2. Clin. Exp. Allergy *35*, 490–497.

Koppelman, S.J., Jayasena, S., Luykx, D., Schepens, E., Apostolovic, D., de Jong, G.A.H., Isleib, T.G., Nordlee, J., Baumert, J., Taylor, S.L., et al. (2016). Allergenicity attributes of different peanut market types. Food Chem. Toxicol. *91*, 82–90.

Koppelman, S.J., Peillon, A., Agbotounou, W., Sampson, H.A., and Martin, L. (2019). Epicutaneous immunotherapy for peanut allergy modifies IgG4 responses to major peanut allergens. J. Allergy Clin. Immunol. *143*, 1218-1221.e4.

Korte, R., and Brockmeyer, J. (2016). MRM3-based LC-MS multi-method for the detection and quantification of nut allergens. Anal. Bioanal. Chem. *408*, 7845–7855.

Krause, S., Reese, G., Randow, S., Zennaro, D., Quaratino, D., Palazzo, P., Ciardiello, M.A., Petersen, A., Becker, W.M., and Mari, A. (2009). Lipid transfer protein (Ara h 9) as a new peanut allergen relevant for a Mediterranean allergic population. J. Allergy Clin. Immunol. *124*.

Kulis, M., Chen, X., Lew, J., Wang, Q., Patel, O.P., Zhuang, Y., Murray, K.S., Duncan, M.W., Porterfield, H.S., Burks, A.W., et al. (2012). The 2S albumin allergens of Arachis hypogaea, Ara h 2 and Ara h 6, are the major elicitors of anaphylaxis and can effectively desensitize peanut-allergic mice. Clin. Exp. Allergy *42*, 326–336.

Lack, G., Fox, D., and Golding, J. (1999). The Role of the Uterine Environment in the Pathogenesis of Peanut Allergy. J. Allergy Clin. Immunol. *103*.

Lane, C.S. (2005). Mass spectrometry-based proteomics in the life sciences. Cell. Mol. Life Sci. *62*, 848–869.

Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. Mol. Syst. Biol. *4*.

Lehmann, K., Schweimer, K., Reese, G., Randow, S., Suhr, M., Becker, W.M., Vieths, S., and Rösch, P. (2006). Structure and stability of 2S albumin-type peanut allergens: Implications for the severity of peanut allergic reactions. Biochem. J. *395*, 463–472.

Lemke, P.J., and Taylor, S.L. (1994). Allergic reactions and food intolerances (New York: Raven Press, Ltd.).

Li, H., Nowak-Wegrzyn, A., Charlop-Powers, Z., Shreffler, W., Chehade, M., Thomas, S., Roda, G., Dahan, S., Sperber, K., and Berin, M.C. (2006). Transcytosis of IgE-Antigen Complexes by CD23a in Human Intestinal Epithelial Cells and Its Role in Food Allergy. Gastroenterology *131*, 47–58.

Li, J., Shefcheck, K., Callahan, J., and Fenselau, C. (2010). Primary sequence and site-selective hydroxylation of prolines in isoforms of a major peanut allergen protein Ara h 2. Protein Sci. *19*, 174–182.

Liebler, D.C., and Zimmerman, L.J. (2013). Targeted Quantitation of Proteins by Mass Spectrometry. Biochemistry *52*, 3797–3806.

Lindemann, C., Thomanek, N., Hundt, F., Lerari, T., Meyer, H.E., Wolters, D., and Marcus, K. (2017). Strategies in relative and absolute quantitative mass spectrometry based proteomics. Biol. Chem *398*, 687–699.

Mager, J., Chevion, M., and Glaser, G. (1980). Favism (New York: Academic).

Maleki, S.J., Kopper, R.A., Shin, D.S., Park, C.-W., Compadre, C.M., Sampson, H., Burks, A.W., and Bannon, G.A. (2000). Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. J. Immunol. *164*, 5844–5849.

Mann, M., and Kelleher, N.L. (2008). Precision proteomics: The case for high resolution and high mass accuracy. Proc. Natl. Acad. Sci. U. S. A. *105*, 18132–18138.

Marquardt, R.R., Wang, N., and Arbid, M.S. (1997). Pyrimidine glycosides. In Handbook of Plant and Fungal Toxicants, (Boca Raton, FL: CRC Press), pp. 139–155.

Matricardi, P.M., Kleine-Tebbe, J., Hoffmann, H.J., Valenta, R., Hilger, C., Hofmaier, S., Aalberse, R.C., Agache, I., Asero, R., Ballmer-Weber, B., et al. (2016). EAACI Molecular Allergology User's Guide. Pediatr. Allergy Immunol. *27*, 1–250.

Matthews, D.M. (1975). Intestinal absorption of peptides. Physiol. Rev. *55*, 537–608.

Mazzocchi, A., Venter, C., Maslin, K., and Agostoni, C. (2017). The role of nutritional aspects in food allergy: Prevention and management. Nutrients *9*.

McDermott, R.A., Porterfield, H.S., Mezayen, R. El, Burks, A.W., Pons, L., Schlichting, D.G., Solomon, B., Redzic, J.S., Harbeck, R.J., Duncan, M.W., et al. (2007). Contribution of Ara h 2 to peanut-specific, immunoglobulin E-mediated, cell activation.

Clin. Exp. Allergy *37*, 752–763.

Mekori, Y.A. (1996). Introduction to allergic diseases. Crit. Rev. Food Sci. Nutr. *36*, 1–18.

Michalski, A., Damoc, E., Hauschild, J.-P., Lange, O., Wieghaus, A., Makarov, A., Nagaraj, N., Cox, J., Mann, M., and Horning, S. (2011). Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. Mol. Cell. Proteomics *10*, M111.011015.

Mills, E., Jenkins, J., Marigheto, N., Beltont, P., Gunning, A., and Morris, V. (2002). Allergens of the cupin superfamily. Biochem. Soc. Trans. *30*, 925–929.

Miner-Williams, W.M., Stevens, B.R., Moughan, P.J., Wehkamp, J., Schmid, M., Stange, E.F., Cohnheim., O., Abderhalden, E., von Körösy, K., London, E.S., et al. (2014). Are intact peptides absorbed from the healthy gut in the adult human? Nutr. Res. Rev. *27*, 308–329.

Mittag, D., Akkerdaas, J., Ballmer-Weber, B.K., Vogel, L., Wensing, M., Becker, W.M., Koppelman, S.J., Knulst, A.C., Helbling, A., Hefle, S.L., et al. (2004). Ara h 8, a Bet v 1-homologous allergen from peanut, is a major allergen in patients with combined birch pollen and peanut allergy. J. Allergy Clin. Immunol. *114*, 1410–1417.

Monaci, L., and Visconti, A. (2009). Mass spectrometry-based proteomics methods for analysis of food allergens. TRAC Trends Anal. Chem. *28*, 581–591.

Monaci, L., Pilolli, R., De Angelis, E., and Mamone, G. (2015). Mass spectrometry in food allergen research. In Comprehensive Analytical Chemistry, Y. Pico, ed. pp. 359–393.

Monaci, L., De Angelis, E., Montemurro, N., and Pilolli, R. (2018). Comprehensive overview and recent advances in proteomics MS based methods for food allergens analysis. TrAC Trends Anal. Chem. *106*, 21–36.

Mose, A.P., Mortz, E., Skov, P.S., Mortz, C.G., Eller, E., Sprogøe, U., Barington, T., and Bindslev-Jensen, C. (2019). The quest for ingested peanut protein in human serum. Allergy all.14109.

Mouzo, D., Bernal, J., López-Pedrouso, M., Franco, D., and Zapata, C. (2018). Advances in the biology of seed and vegetative storage proteins based on two-dimensional electrophoresis coupled to mass spectrometry. Molecules *23*.

Mueller, G.A., Gosavi, R.A., Pomés, A., Wünschmann, S., Moon, A.F., London, R.E., and Pedersen, L.C. (2011). Ara h 2: crystal structure and IgE binding distinguish two subpopulations of peanut allergic patients by epitope diversity. Allergy *66*, 878–885.

Mueller, G.A., Maleki, S.J., and Pedersen, L.C. (2014). The Molecular Basis of Peanut Allergy. Curr. Allergy Asthma Rep. *14*, 1–9.

Müntz, K. (1998). Deposition of storage proteins.

Murphy, K. (2012). Janeway's immunobiology (Garland Science).

Mylne, J.S., Hara-Nishimura, I., and Rosengren, K.J. (2014). Seed storage albumins: biosynthesis, trafficking and structures. Funct. Plant Biol. 41, 671–677.

Nowak-Wegrzyn, A. (2007). Food Allergy to Proteins.

Nunes, R., Silva, C., and Chaves, L. (2016). Tissue-based in vitro and ex vivo models for intestinal permeability studies. Concepts Model. Drug Permeability Stud. 203–236.

Oerise, N.L., Lau, H.A., Ritchey, S.J., and Murphy, E.W. (1974). Yield proximate composition, and mineral element content of three cultivars of raw and roasted peanuts. J. Food Sci. *39*, 264–266.

Osborne, T. (1907). The proteins of the wheat kernel (Washington, DC: Judd and Detweiler, Inc).

Paganelli, R., and Levinsky, R.J. (1980). Solid phase radioimmunoassay for detection of circulating food protein antigens in human serum. J. Immunol. Methods *37*, 333–341.

Peeters, K.A.B.M., Koppelman, S.J., van Hoffen, E., van der Tas, C.W.H., den Hartog Jager, C.F., Penninks, A.H., Hefle, S.L., Bruijnzeel-Koomen, C.A.F.M., Knol, E.F., and Knulst, A.C. (2007). Does skin prick test reactivity to purified allergens correlate with clinical severity of peanut allergy? Clin. Exp. Allergy *37*, 108–115.

Pelz, B.J., and Bryce, P.J. (2015). Pathophysiology of Food Allergy.

Petersen, A., Rennert, S., Kull, S., Becker, W.M., Notbohm, H., Goldmann, T., and Jappe, U. (2014). Roasting and lipid binding provide allergenic and proteolytic stability to the peanut allergen Ara h 8. Biol. Chem. *395*, 239–250.

Peterson, L.W., and Artis, D. (2014). Intestinal epithelial cells: Regulators of barrier function and immune homeostasis. Nat. Rev. Immunol. *14*, 141–153.

Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., and Coon, J.J. (2012). Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics.

Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods *9*, 555–566.

Piersma, S.R., Gaspari, M., Hefle, S.L., and Koppelman, S.J. (2005). Proteolytic processing of the peanut allergen Ara h 3. Mol. Nutr. Food Res. *49*, 744–755.

Plomley, J. Harnessing the Power of MRM3 for Large Molecule Quantitative Bioanalysis | SCIEX Community.

Pons, L., Chery, C., Romano, A., Namour, F., Artesani, M.C., and Guéant, J.L. (2002). The 18 kDa peanut oleosin is a candidate allergen for IgE-mediated reactions to peanuts. In Allergy: European Journal of Allergy and Clinical Immunology, Supplement, (Blackwell Munksgaard), pp. 88–93.

Rabjohn, P., Helm, E., Stanley, J., West, C., Sampson, H., Burks, A., and Bannon, G. (1999). Molecular cloning and epitope analysis of the peanut allergen Ara h 3. J. Clin. Invest. 103, 535–542.

Rauniyar, N. (2015). Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. Int. J. Mol. Sci. *16*, 28566–28581.

Van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., et al. (2000). β(1,2)-xylose and α(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. J. Biol. Chem. *275*, 11451–11458.

Reitsma, M., Westerhout, J., Wichers, H.J., Wortelboer, H.M., and Verhoeckx, K.C.M. (2014). Protein transport across the small intestine in food allergy. Mol. Nutr. Food Res *58*, 194–205.

Rescigno, M., Urbano, M., Valzasina, B., Francolini, M., Rotta, G., Bonasio, R., Granucci, F., Kraehenbuhl, J.P., and Ricciardi-Castagnoli, P. (2001). Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. Nat. Immunol. *2*, 361–367.

Rinderknecht, H. (1993). Pancreatic secretory enzymes. In The Pancreas: Biology, Pathobiology, and Diseas, pp. 219–251.

Rivera, J., Fierro, N.A., Olivera, A., and Suzuki, R. (2008). Chapter 3 New Insights on Mast Cell Activation via the High Affinity Receptor for IgE1. Adv. Immunol. *98*, 85–120.

Romagnani, S. (1997). The Th1/Th2 paradigm. Immunol. Today *18*, 263–266.

Ronsein, G.E., Pamir, N., von Haller, P.D., Kim, D.S., Oda, M.N., Jarvik, G.P., Vaisar, T., and Heinecke, J.W. (2015). Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. J. Proteomics *113*, 388–399.

Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., et al. (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol. Cell. Proteomics *3*, 1154–1169.

Rubio-Tapia, A., and Murray, J.A. (2008). Gluten-sensitive enteropathy. In Food Allergy: Adverse Reactions to Foods and Food Additives, D.D. Metcalfe, H.A. Sampson, and R.A. Simon, eds. (Malden, MA: Blackwell Publishing), pp. 211–222.

Rune, S.J., and Viskum, K. (1969). Duodenal pH values in normal controls and in patients with duodenal ulcer. Gut *10*, 569–571.

Sagaram, U.S., Pandurangi, R., Kaur, J., Smith, T.J., and Shah, D.M. (2011). Structure-Activity Determinants in Antifungal Plant Defensins MsDef1 and MtDef4 with Different Modes of Action against Fusarium graminearum. PLoS One *6*, e18550.

Samadi, N., Klems, M., and Untersmayr, E. (2018). The role of gastrointestinal permeability in food allergy. Ann. Allergy, Asthma Immunol.

Sampson, H.A. (1988). IgE-mediated food intolerance. J. Allergy Clin. Immunol. *81*, 495–504.

Sampson, H.A. (1991). Immunologic mechanisms in adverse reactions to foods. Immunol. Allergy Clin. North Am. *11*, 701–716.

Sampson, H.A. (1999). Food allergy. Part 1: Immunopathogenesis and clinical disorders. J. Allergy Clin. Immunol. *103*, 717–728.

Sampson, H.A. (2003). 9. Food allergy. J. Allergy Clin. Immunol. *111*.

Sampson, H.A. (2009). Anaphylaxis and Food Allergy. In Food Allergy, D.D. Metcalfe, H.A. Sampson, and R.A. Simon, eds. (Oxford, UK: Blackwell Publishing Ltd.), pp. 157–170.

Sampson, H.A., and Eigenmann, P.A. (1999). Food allergy and intolerance (New York: Marcel Dekker, Inc.).

Sampson, H.A., Aceves, S., Bock, S.A., James, J., Jones, S., Lang, D., Nadeau, K., Nowak-Wegrzyn, A., Oppenheimer, J., Perry, T.T., et al. (2014). Food allergy: A practice parameter update—2014. J. Allergy Clin. Immunol. *134*, 1016-1025.e43.

Savaryn, J.P., Toby, T.K., and Kelleher, N.L. (2016). A researcher's guide to mass spectrometry-based proteomics. Proteomics *0*, 1–9.

Schmitt, D.A., Nesbit, J.B., Hurlburt, B.K., Cheng, H., and Maleki, S.J. (2010). Processing can alter the properties of peanut extract preparations. J. Agric. Food Chem. *58*, 1138–1143.

Schocker, F., Baumert, J., Kull, S., Petersen, A., Becker, W.-M., and Jappe, U. (2016). Prospective investigation on the transfer of Ara h 2, the most potent peanut allergen, in human breast milk. Pediatr. Allergy Immunol. *27*, 348–355.

Schultz, S.G., and Curran, P.F. (1970). Coupled transport of sodium and organic solutes. Physiol. Rev. *50*, 637–718.

Schwager, C., Kull, S., Krause, S., Schocker, F., Petersen, A., Becker, W.-M., and Jappe, U. (2015). Development of a Novel Strategy to Isolate Lipophilic Allergens (Oleosins) from Peanuts.

Scott, M.P., Jung, R., Muntzt, K., and Nielsen, N.C. (1992). A protease responsible for post-translational cleavage of a conserved Asn-Gly linkage in glycinin, the major seed storage protein of soybean (protein maturation/endopeptidase/purification/Glycine max/ulS globulins).

Sebei, K., Gnouma, A., Herchi, W., Sakouhi, F., and Boukhchina, S. (2013). Lipids, proteins, phenolic composition, antioxidant and antibacterial activities of seeds of peanuts (Arachis hypogaea l) cultivated in Tunisia. Biol. Res. *46*, 257–263.

Sen, M., Kopper, R., Pons, L., Abraham, E.C., Burks, A.W., and Bannon, G.A. (2002).

Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes. J. Immunol. *169*, 882–887.

Sherman, J., McKay, M.J., Ashman, K., and Molloy, M.P. (2009). How specific is my SRM?: The issue of precursor and product ion redundancy. Proteomics *9*, 1120–1123.

Shewry, P.R. (2007). Improving the protein content and composition of cereal grain. J. Cereal Sci. *46*, 239–250.

Shewry, P.R., Napier, J.A., and Tatham, A.S. (1995). Seed storage proteins: structures and biosynthesis. Plant Cell *7*, 945–956.

Shimizu, M. (2004). Food-derived peptides and intestinal functions (IOS Press).

Sicherer, S.H., and Sampson, H.A. (2006). 9. Food allergy. J. Allergy Clin. Immunol. *117*, S470–S475.

Sicherer, S.H., and Sampson, H.A. (2010). Food allergy. J. Allergy Clin. Immunol. *125*.

Sicherer, S.H., and Wood, R.A. (2013). Advances in diagnosing peanut allergy. J Allergy Clin Immunol Pr. *1*, 1–13.

Sicherer, S., Muñoz-Furlong, A., Godbold, J., and Sampson, H. (2010). US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immunol. *125*, 1322–1326.

Sicherer, S.H., Furlong, T.J., Maes, H.H., Desnick, R.J., Sampson, H.A., and Gelb, B.D. (2000). Genetics of peanut allergy: A twin study. J. Allergy Clin. Immunol. *106*, 53–56.

Sidoli, S., Fujiwara, R., and Garcia, B.A. (2016). Multiplexed data independent acquisition (MSX-DIA) applied by high resolution mass spectrometry improves quantification quality for the analysis of histone peptides. Proteomics *16*, 2095–2105.

Sloan, E.A., and Powers, M.E. (1986). A perspective on popular perceptions of adverse reactions to foods. J. Allergy Clin. Immunol. *78*, 127–133.

So, A.L., Small, G., Sperber, K., Becker, K., Oei, E., Tyorkin, M., and Mayer, L. (2000). Factors affecting antigen uptake by human intestinal epithelial cell lines. Dig. Dis. Sci. *45*, 1130–1137.

Stanley, J.S., King, N., Burks, A.W., Huang, S.K., Sampson, H., Cockrell, G., Helm, R.M., West, C.M., and Bannon, G.A. (1997). Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. Arch. Biochem. Biophys. *342*, 244–253.

Steel, A., and Hediger, M.A. (1998). The Molecular Physiology of Sodium- and Proton-Coupled Solute Transporters. Physiology *13*, 123–131.

Stensgaard, A., Bindslev-Jensen, C., Nielsen, D., Munch, M., and DunnGalvin, A. (2017). Quality of life in childhood, adolescence and adult food allergy: Patient and parent perspectives. Clin. Exp. Allergy *47*, 530–539.

Stone, K.D., Prussin, C., and Metcalfe, D.D. (2010). IgE, mast cells, basophils, and

eosinophils. J. Allergy Clin. Immunol. *125*.

Strobel, S., and Mowat, A. (1998). Immune responses to dietary antigens: Oral tolerance. Immunol. Today *19*, 173–181.

Suarez, F.L., and Savaiano, D.A. (1997). Diet, genetics, and lactose intolerance. Food Technol. *51*, 74–76.

Suhr, M., Wicklein, D., Lepp, U., and Becker, W.-M. (2004). Isolation and characterization of natural Ara h 6: evidence for a further peanut allergen with putative clinical relevance based on resistance to pepsin digestion and heat. Mol. Nutr. Food Res. *48*, 390–399.

Taylor, S.L. (2014). Food Toxicology. In Food Allergy: Adverse Reactions to Foods and Food Additives, D.D. Metcalfe, H.A. Sampson, R.A. Simon, and G. Lack, eds. (Chichester, UK: John Wiley & Sons Ltd), pp. 507–517.

Taylor, S., and Hefle, S. (2002). Allergic reactions and food intolerances (London: Taylor and Francis).

Taylor, S., and Hefle, S. (2006). Food allergen labeling in the USA and Europe. Curr. Opin. Allergy Clin. Immunol. *6*, 186–190.

Taylor, S.L., and Baumert, J.L. (2012). Food Allergies and Intolerances. In Present Knowledge in Nutrition, (Wiley-Blackwell), pp. 1222–1235.

Taylor, S.L., and Hefle, S.L. (2001). Food allergies and other food sensitivities.

Taylor, S.L., Nordlee, J.A., and Rupnow, J.H. (1989). Food allergies and sensitivities. In Food Toxicology A Perspective on the Relative Risks, S.L. Taylor, and R.A. Scanlan, eds. (New York: Marcel Dekker, Inc.), pp. 255–295.

Taylor, S.L., Nordlee, J.A., and Bush, R.K. (1992). Food Allergies. In Food Safety Assessment, J.W. Finley, S.F. Robinson, and D.J. Armstrong, eds. pp. 316–329.

Tessari, P., Lante, A., and Mosca, G. (2016). Essential amino acids: Master regulators of nutrition and environmental footprint? Sci. Rep. *6*.

Toby, T.K., Fornelli, L., and Kelleher, N.L. (2016). Progress in Top-Down Proteomics and the Analysis of Proteoforms.

Du Toit, G., Roberts, G., Sayre, P.H., Bahnson, H.T., Radulovic, S., Santos, A.F., Brough, H.A., Phippard, D., Basting, M., Feeney, M., et al. (2015). Randomized Trial of Peanut Consumption in Infants at Risk for Peanut Allergy. N. Engl. J. Med. *372*, 803–813.

Turner, P.J., Gowland, M.H., Sharma, V., Ierodiakonou, D., Harper, N., Garcez, T., Pumphrey, R., and Boyle, R.J. (2015). Increase in anaphylaxis-related hospitalizations but no increase in fatalities: An analysis of United Kingdom national anaphylaxis data, 1992-2012. J. Allergy Clin. Immunol. *135*, 956-963.e1.

Untersmayr, E., and Jensen-Jarolim, E. (2006). The effect of gastric digestion on food allergy. Curr. Opin. Allergy Clin. Immunol. *6*, 214–219.

Valenta, R., Hochwallner, H., Linhart, B., and Pahr, S. (2015). Food allergies: The basics. Gastroenterology *148*, 1120-1131.e4.

Vercelli, D., and Geha, R.S. (1992). Regulation of isotype switching. Curr. Opin. Immunol. *4*, 794–797.

Vickery, B.P., Lin, J., Kulis, M., Fu, Z., Steele, P.H., Jones, S.M., Scurlock, A.M., Gimenez, G., Bardina, L., Sampson, H.A., et al. (2013). Peanut oral immunotherapy modifies IgE and IgG4 responses to major peanut allergens. J. Allergy Clin. Immunol. *131*, 128-134.e3.

Vieths, S., Scheurer, S., and Ballmer-Weber, B. (2002). Current Understanding of Cross-Reactivity of Food Allergens and Pollen.

Wambre, E., and Jeong, D. (2018). Oral Tolerance Development and Maintenance. Immunol. Allergy Clin. North Am. *38*, 27–37.

Wang, J. (2009). Oral Allergy Syndrome. In Food Allergy: Adverse Reactions to Foods and Food Additive, D.D. Metcalfe, H.A. Sampson, and R.A. Simon, eds. (Oxford, UK: Blackwell Publishing Ltd.), pp. 133–143.

Wang, J., and Sampson, H.A. (2011). Food allergy. J. Clin. Invest. *121*, 827–835.

Warren, C.M., Otto, A.K., Walkner, M.M., and Gupta, R.S. (2016). Quality of Life Among Food Allergic Patients and Their Caregivers. Curr. Allergy Asthma Rep. *16*.

Wichers, H.J., de Beijer, T., Savelkoul, H.F.J., and van Amerongen, A. (2004). The Major Peanut Allergen Ara h 1 and Its Cleaved-off N-Terminal Peptide; Possible Implications for Peanut Allergen Detection. J. Agric. Food Chem. *52*, 4903–4907.

Wickham, M., Faulks, R., and Mills, C. (2009). In vitro digestion methods for assessing the effect of food structure on allergen breakdown. Mol. Nutr. Food Res. *53*, 952–958.

Wilson, J., and Vachet, R.W. (1996). Multiplexed MS/MS in a Quadrupole Ion Trap Mass Spectrometer. J. Am. Soc. Mass Spectrom *121*, 7346–7353.

Wood, R.A. (2003). The natural history of food allergy. Pediatrics *111*, 1631–1637.

Yagami, T., Haishima, Y., Nakamura, A., Osuna, H., and Ikezawa, Z. (2000). Digestibility of allergens extracted from natural rubber latex and vegetable foods. J. Allergy Clin. Immunol. *106*, 752–762.

Yamauch, F., Yamagishi, T., and Iwabuchi, S. (1991). Molecular understanding of heat-induced phenomena of soybean protein. Food Rev. Int. *7*, 283–322.

Yan, Y.S., Lin, X.D., Zhang, Y.S., Wang, L., Wu, K., and Huang, S.Z. (2005). Isolation of peanut genes encoding arachins and conglutins by expressed sequence tags. Plant Sci. *169*, 439–445.

Yates, J.R. (1998). Mass Spectrometry and the Age of the Proteome. J. Mass Spectrom. *33*, 1–19.

Yost, R.A., and Enke, C.G. (1978). Selected Ion Fragmentation with a Tandem

Quadrupole Mass Spectrometer. J. Am. Chem. Soc. *100*, 2274–2275.

Yu, L.C.H., Yang, P.C., Berin, M.C., Di Leo, V., Conrad, D.H., McKay, D.M., Satoskar, A.R., and Perdue, M.H. (2001). Enhanced transepithelial antigen transport in intestine of allergic mice is mediated by IgE/CD23 and regulated by interleukin-4. Gastroenterology *121*, 370–381.

Yu, W., Hussey Freeland, D.M., and Nadeau, K.C. (2017). Food allergy: immune mechanisms, diagnosis and immunotherapy.

Zhou, J., and Yin, Y. (2016). Strategies for large-scale targeted metabolomics quantification by liquid chromatography-mass spectrometry. Analyst *141*, 6362–6373.

Zhuang, Y., and Dreskin, S.C. (2013). Redefining the major peanut allergens. Immunol. Res. *55*, 125–134.

Zubarev, R.A., and Makarov, A. (2013). Orbitrap Mass Spectrometry. Anal. Chem. *85*, 5288–5296.

CHAPTER 2: EVALUATION OF DE-COMPLEXING STRATEGIES FOR *IN VIVO* DETECTION OF FOOD PROTEINS IN HUMAN SERUM USING MASS SPECTROMETRY

ABSTRACT

Food allergy is a serious and potentially life-threatening condition, caused by the ingestion of allergenic foods. The proteins of allergenic foods are resistant to thermal treatment and gastric digestion, and may enter into circulation as intact or partially intact molecules (Baumert et al., 2009; Dirks et al., 2005; Husby et al., 1985; JanssenDuijghuijsen et al., 2017). Food allergy prevalence rates have increased over the past decade, and in particular, peanut allergy has increased 20% since 2010 (Gupta et al., 2017; Jackson et al., 2013; Kotz et al., 2011; Lieberman et al., 2018; Rinaldi et al., 2012; Sicherer et al., 2010). In efforts to understand this increasing prevalence rate, it is important to understand *in vivo* protein transport and sensitization; however, limited studies exist focusing on allergenic protein uptake and allergenic protein characteristics following uptake.

Detection of peanut proteins in human body fluids (serum, breast milk, saliva) has been observed at very low levels by immunoassay(s) with poor overall recoveries in comparison to the amount of protein consumed (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017). The specific aim of this chapter is to evaluate multiple serum de-complexation strategies, with the overall objective of developing an *in vivo* discovery LC-MS/MS acquisition method for peanut proteins in serum. Shotgun LC-

MS/MS is able to identify multiple peanut proteins and other serum proteins within a single sample, making it a highly advantageous and informative method.

Initial analyses of subject serum collected after peanut consumption were negative for peanut peptides when evaluated by a standard LC-MS/MS discovery acquisition method. These subject serum samples were collected and analyzed in studies conducted by Baumert et al., 2009. These subject sera samples were previously determined positive by immunoassay, indicating our current discovery LC-MS/MS method is not sufficiently sensitive (Baumert et al., 2009). As a result, we began our studies by evaluating various de-complexation strategies for removal or separation of abundant serum proteins.

A variety of de-complexing strategies, including four commercial depletion kits, an organic solvent fractionation method, and modified MS acquisition settings (exclusion, inclusion lists) were evaluated using a model incurred matrix of human serum and raw peanut extract (10:1 (w/w) protein) with detection by discovery LC-MS/MS.

Commercial depletion kits, targeting the abundant serum proteins, co-depleted peanut proteins from the model matrices as determined by SDS-PAGE and LC-MS/MS. The commercial depletion kits demonstrated a lack of specificity and simultaneously removed peanut proteins from the analytical matrices, peanut or sera-peanut. We then evaluated an organic fractionation technique to decrease sample complexity prior to LC-MS/MS analysis. This cost-effective de-complexation strategy was useful in decreasing sample complexity, however, we observed variability in fractionation patterns of immunoglobulins and peanut proteins, dependent on individual sera (e.g. sensitized or non-sensitized). Evaluations of MS acquisition settings, inclusion and exclusion lists, for specified peptides (i.e. m/z events) were largely unsuccessful.

Overall, each de-complexation strategy was successful in their described functionality of depleting, fractionating, or including or excluding of specific mass events but, were largely unsuccessful. We suggest evaluating non-depleted serum for future *in vivo* analyses since de-complexation strategies are unfit for *in vivo* detection of allergenic proteins in serum.

INTRODUCTION

Peanut allergy, is the most prevalent allergy in children and third most prevalent in adults (Gupta et al., 2018, 2019). It is a highly severe and potent allergy, which can be elicited by trace amounts of peanut protein in the most sensitive individuals (Klemans et al., 2013; Sicherer and Sampson, 2007). Peanut allergens have been widely studied due to their severity, prevalence, and ability to retain allergenicity into adulthood as most children do not outgrow peanut allergy (Wood, 2003). Together, peanuts and tree nuts account for a majority of anaphylactic cases reported (Sicherer and Sampson, 2007; Sicherer et al., 2010). Due to its continued increasing prevalence, and limited immunotherapy treatments, peanut allergy is a major public health concern.

It is suspected that proteins must enter into circulation in an immunologically reactive form in order to elicit an allergic reaction (Heyman, 2005). In peanuts, the 2S albumins contribute to overall allergenicity and potency, due to their thermostability and resistance to gastric digestion (Sen et al., 2002; Suhr et al., 2004). The 2S albumins and their tightly coiled helices resist digestion and retain secondary structural conformations (Apostolovic et al., 2016). Consequently, Ara h 2 and 6 are clinically relevant proteins for peanut allergy diagnosis based on effector-cell assays (Blanc et al., 2009; Kulis et al.,

2012). These proteins are capable of desensitization in a mouse model, further demonstrating their clinical relevance (Kulis et al., 2012; Zhuang and Dreskin, 2013). In combination, these peanut allergen characteristics enhances their ability to retain conformation and traverse the gastrointestinal barrier largely intact.

Previous studies have predominantly utilized immunoassay methods, which are highly selective for pre-determined antigenic targets (antigen-antibody capture), for *in vivo* allergenic protein(s) measurement (Koppelman and Hefle, 2006). These *in vivo* studies demonstrated multiple analytical targets including Ara h 2, Ara h 6, or their digestive resistant peptide (DRP) fragments, were present in human serum and breastmilk after consumption (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017). Baumert et al., 2009 described detection of a digestion resistant peptide of Ara h 2 in healthy adult serum and saliva using an inhibition ELISA to the DRP-Ara h 2 (Baumert et al., 2009). Schocker et al., 2016 coupled immunoaffinity capture antibody and LC-MS/MS and inhibition ELISA for detection of Ara h 2 in breastmilk (Schocker et al., 2016). More recently, Ara h 6 was detected in subject serum collected after peanut consumption by commercial ELISA (JanssenDuijghuijsen et al., 2017). One key challenge when using ELISA methods is the unknown nature of allergenic proteins after uptake, which may be modified during digestion and unrecognizable by capture antibodies (Reitsma et al., 2014).

To further complicate matters of detection, the human serum proteome is a dynamic and dense protein matrix (60 – 80 mg protein/ml serum) spanning twelve orders of magnitude (Adkins et al., 2002; Anderson and Anderson, 2002). Only a few dozen proteins account for the majority (99%) of serum protein content whereas thousands of

proteins contribute to the remaining 1% (Anderson and Anderson, 2002; Pieper et al., 2003). The lesser abundant proteins are of significant clinical interest due to their potential impacts on human health and disease (Roche et al., 2009).

Due to the inherent complexity of serum, numerous depletion techniques have been developed to remove the majorly abundant serum proteins (e.g. serum albumin, immunoglobulins). Depletion methods selectively remove protein by immunoaffinity, dye binding, or other physico-chemical methods, and most often targeting multiple proteins (Gianazza and Arnaud, 1982; Leatherbarrow and Dean, 1980). Although effective, there is potential of co-elution of low abundant proteins during depletion of the target proteins (Roche et al., 2009). Removal of the abundant serum proteins inherently decreases sample complexity, improving overall detection of lesser abundant proteins, including biomarkers or, for our purposes, allergenic food proteins. The objective of this study was to evaluate a variety of de-complexing strategies for removal of abundant serum proteins.

The utilization of discovery LC-MS/MS allows for the simultaneous identification of serum proteins involved in IgE-mediated reactions (Gillet et al., 2016). We would be able to identify proteins upregulated during an immunologically mediated response (Gillet et al., 2016), making discovery MS an advantageous method for improving our understanding of the mechanisms associated with allergenic protein uptake.

MATERIALS AND METHODS

I. Reagents

All reagents used were of analytical grade for all experiments. All reagents used for LC-MS/MS sample preparations and analyses were of MS grade.

II. Preparation of Peanut Flours

Raw and roasted peanut extracts were prepared at two extraction buffer concentrations, 0.1 and 0.01 M phosphate buffered saline (PBS). The prepared raw peanut extracts were used as positive controls in subsequent LC-MS/MS method developmental experiments.

Raw red-skin peanuts (Wor-Fung blanched peanuts, distributor K.N.T.C, South El Monte, CA 91733, USA), purchased locally, were used to prepare raw peanut flour. Peanut skins were removed and peanuts were washed five times with distilled water and air-dried. Raw peanuts were ground in a Magic Bullet (Homeland Housewares, LLC) blender. Ground peanuts were defatted (1:5 w/v peanut:hexane) three times, filtered, and air-dried. The prepared, defatted raw peanut flour was ground once more as described previously.

Roasted peanut extracts were prepared from partially defatted (12%) light roasted peanut flour purchased from Golden Peanut Company (Alpharetta, GA, USA).

Raw and roasted peanut flours were extracted 1:50 (w/w) in 0.1 or 0.01 M PBS in a 60°C sonicating water bath for 20 minutes, followed by centrifugation (3,500 x g) at 10°C for 10 minutes (Beckman GS-15R centrifuge). The supernatants (aqueous phase)

were removed and clarified by centrifugation (17,000 x g) for 10 minutes (Thermo Scientific™ Sorvall Legend Micro 17). Clarified extracts were pooled and centrifuged (3,500 x g, 10 minutes, 10°C) (Beckman GS-15R centrifuge) to achieve a homogenous extraction solution. Extracts were dialyzed using 3500 MWCO dialysis cassettes (Thermo Scientific™ Slide-A-Lyzer™ Dialysis Cassettes, Cat. No. 66330, Pierce Biotechnology, Rockford, IL, USA) overnight at room temperature to 0.1 or 0.01 M PBS, respectively, with two buffer exchanges. Dialyzed samples were stored in 1 mL aliquots at -20°C until needed for further analysis.

Peanut extracts were characterized by reducing SDS-PAGE and protein concentration was determined by 2D Quant Kit (GE Healthcare Bio-Sciences, Prod. No. 80648356, Piscataway, NJ). The 2D Quant assay determines the protein concentration by precipitation followed by re-suspension in a copper ion solution (GE Healthcare). Triplicate extracts were evaluated by the 2D Quant kit for protein concentration.

SDS-PAGE was performed under reducing conditions using NuPAGE Bis-Tris Mini Gels 4-12% (1.0 mm, 12 wells) and constant voltage (200V) for 40 minutes in an XCell SureLock Mini Cell Electrophoresis System (Invitrogen Life Technologies). Raw and roasted peanut extracts were diluted (1.5x) in 4 x concentrated Laemmli buffer and 1% β-mercapto-ethanol (BME). Extracts were reduced by heating at 95°C for 5 minutes. Samples (20 µL) were loaded into each gel well. Precision Plus Protein Dual Xtra Standards (Bio-Rad) were used as the molecular weight (MW) standard. Gels were stained overnight in Coomassie Brilliant Blue R-250 (Bio-Rad) stain. Gels were de-stained (Coomassie Brilliant Blue R-250 destaining solution, Bio-Rad), rehydrated, and imaged.

Unless otherwise noted, raw peanut extracts were used for preparation of model matrices in all subsequent experiments.

III. Preparation of Model Matrices for Discovery LC-MS/MS Method Development

Subject serum were collected before (baseline) peanut consumption by Baumert et al., 2009 for immunoassay evaluation. Subject sera was then collected after peanut consumption at specified time points (0.5, 1, 2, 3, 4, 6, 8, and 24 hours). Subjects were instructed to consume 25 g roasted peanut flour, administered in capsules with 0.83 g roasted peanut flour per capsule. According to Baumert et al., 2009, venous blood (10 mL) was collected using a heparin lock inserted into an arm vein. Blood was processed into serum and stored at -20°C. Serum was stored long-term at -80°C for use in later studies.

Baseline serum was used in preparation of model-matrices. Serum and raw peanut extract (1.34 µg/µL) were individually diluted 10-fold in 0.01 M PBS, and combined, 10:1 (w/w) serum:peanut. Model matrices were prepared to achieve the desired protein amount (µg) for the subsequent experiments described below. Control samples (sera, peanut) were individually and equivalently prepared to the model matrices by mixing with 0.01 M PBS to the desired protein amount. Samples were vortexed to mix and incubated on ice (30 minutes) until needed for further analysis.

IV. Discovery LC-MS/MS Acquisition Method Optimization

Using the serum-peanut matrix, and equivalently prepared individual serum and peanut extract controls, we optimized the LC-MS/MS acquisition settings to enhance detection of low abundant peptides.

A. In-Solution Reduction, Alkylation, and Trypsin Digestion for LC-MS/MS Analysis of Model Matrices

An in-solution trypsin digestion from Thermo Scientific™ (Thermo Scientific™ In-Solution Tryptic Digestion and Guanidination Kit, 89895, Pierce Biotechnology, Rockford, IL, USA) was modified for sample preparation. The serum-peanut matrix was diluted to 2 µg/µL, and 8 µL were taken for tryptic digestion. Proteins were diluted with 50 mM ammonium bicarbonate and reduced with 100 mM dithiothreitol at 95°C for 5 minutes. Reduced samples were alkylated with 100 mM iodoacetamide, in the dark at room temperature for 20 minutes. Trypsin (100 ng/µL made in 5 mM acetic acid) was added and incubated for 3 hours at 37°C. A second addition of trypsin was added and continued to digest overnight at 30°C, achieving a final enzyme:protein ratio 1:50 (w/w). Digestion was stopped by freezing samples. Digests were de-salted using C18 spin columns (Pierce C18 spin columns, Thermo Scientific™, Rockford, IL, USA) according to the manufacturer's instructions, and eluted in 70% acetonitrile. De-salted peptides were dried to completion under a vacuum by centrifugal evaporation (Jouan RC-10.10; RCT-90; Winchester, VA, USA). Peptides were re-suspended to 0.2 µg/µL in 0.1% formic acid, 5% acetonitrile. Re-suspended peptides were injected (5 µL) for LC-MS/MS.

B. LC-MS/MS Acquisition using DDA

Peptide digests (5 µL) were chromatographically separated using an UltiMate 3000RSL® liquid chromatography (UPLC) system (Thermo Scientific™) equipped with a Hypersil Gold C18 1.9 µm, 100 x 1 mm reversed phase column (Thermo Scientific™) with a pre-column (20 x 2.1 mm reversed phase, 1.9 µm, Thermo Scientific™) set at 35°C. Mobile phase A contained 0.1% (v/v) formic acid in water and mobile phase B contained 0.1% (v/v) formic acid in acetonitrile. Peptides were separated using a linear gradient of 2 – 40% mobile phase B over 70 minutes at a flow rate of 60 µL/min. Following the gradient elution, the column was washed (60 µL/min) for 5 minutes at 98% mobile phase B, followed by 100% methanol for 5 minutes (60 µL/min). The separation column was re-equilibrated at 2% mobile phase B (180 µL/min) for 15 minutes. The flow rate was reduced to 60 µL/min prior to the next sample injection.

DDA were performed on a Thermo Q Exactive Plus™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Scientific™) operating in positive ion mode. The MS acquisition settings adjusted were dynamic exclusion time, MS2 fill time (Ctrap fill time), Top N acquisition, peptide charge states for MS2 acquisition(s), and automatic gain control (AGC) target. The settings described in Table 2-1 were optimized to improve detection of low abundant proteins.

Survey scan mass spectra (400 – 1400 m/z) were acquired at a nominal resolution of 70,000 FWHM (200 m/z) and an AGC target of 3e6. Fragmentation spectra were acquired at a nominal resolution of 70,000 FWHM, normalized collisional energy (NCE) set at 27, and an AGC target of 1e5. The electrospray ionization settings were as follows:

sheath gas 15 AU, spray voltage 3500 V, capillary temperature 320°C, S-lens RF level 60.

Each method, described in Table 2-1, was evaluated sequentially for each analytical sample (1) sera, (2) peanut, and (3) sera-peanut.

Method	Dynamic exclusion (seconds)	TopN	Charges	MS2 fill time (msec)
1	3	10	1, 2, 3, 4, 5, 6	60
2	30	20	2, 3	60
3	20	20	2, 3, 4	240

Table 2-1. Discovery LC-MS/MS acquisition settings for improvement in detection of low abundance peptides.

C. LC-MS/MS Data Analysis

Data were analyzed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada) software against two publically available databases from UniProt, (1) *Homo sapiens* and (2) *Arachis hypogaea*. Protein identifications were made by searching a (A) sequence database of peanut (*Arachis hypogaea*, taxon identifier 3818), and (B) sequence database of human (*Homo sapiens*, taxon identifier 9606) as available in UniProt on 08/31/2016. Mass spectral data were normalized for the total ion current (TIC) and semi-quantified using label-free quantification. The following criteria

were used for protein identification: no missed tryptic cleavages, fixed modification of carbamidomethylation of cysteine; variable modifications of methionine, oxidation, and hydroxylation of proline; parent mass error tolerance 2 ppm, fragment error tolerance 0.02 Da; and an FDR set to 1%.

V. Initial Evaluation of Subject Serum Samples and Model Matrices of Sera-Peanut Matrix

Initial *in vivo* evaluations for peanut protein in serum were conducted by Baumert et al., 2009 by competitive immunoassay. The 60 minute time point was selected for initial discovery LC-MS/MS evaluations due to the high reported concentration by DRP-specific ELISA (Baumert et al., 2009). Over a 24 hour consumption period, the average absorption reported was 131 ng DRP-Ara h 2/mL serum.

A. In-Solution Reduction, Alkylation, and Trypsin Digestion of Subject Sera Samples

Subject sera (baseline, 60-minute) were prepared for LC-MS/MS analysis using the in-solution trypsin digestion, as described previously (page 80). Alternatively, after de-salting, peptides were eluted in 50% (v/v) acetonitrile, and then dried to completion under a vacuum by centrifugal evaporation (Jouan RC-10.10; RCT-90; Winchester, VA, USA). Peptides were re-solubilized in 0.1% formic acid, 5% acetonitrile to a final concentration of 0.5 µg/µL. Peptide digests were injected (2 µL) in duplicate for LC-MS/MS acquisitions.

B. LC-MS/MS Acquisition Settings

Data dependent acquisitions (DDA) were performed using the optimized discovery method, method 3, as previously described (page 81).

C. LC-MS/MS Data Analysis

Data were analyzed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada) software, as previously described (page 82).

VI. Development of Exclusion and Inclusion Lists

A. Exclusion List Development

Data collected from baseline serum samples used in the preceding experiment were analyzed for use in the exclusion list. The optimized LC-MS/MS chromatographic method and acquisition settings were used to generate the exclusion list previously described. Acquired data were analyzed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada) against the publically available *Homo sapiens* (taxon identifier 9606) database (UniProt) as available on 8/31/2016, using the previously described data analysis settings (page 82). The most abundant identified *H. sapiens* m/z events were selected (top 5,000 events) and collated into an exclusion list included in the LC-MS/MS acquisition methodology. These selected m/z's were cross-referenced

against the generated peanut data to eliminate identical m/z's. Charge states of 1+, 5 to 9+ were excluded from acquisition.

B. Inclusion List Development

We similarly investigated inclusion lists in combination with the LC-MS/MS acquisition workflow. The inclusion lists were developed and applied in the same manner as exclusion lists.

Inclusion lists were generated by obtaining spectral data from duplicate digestions of sera-peanut (prepared with baseline sera) and peanut (alone). Sera-peanut (10:1 (w/w)) and an equivalent volume of peanut were enzymatically digested and prepared for LC-MS/MS analysis. Analytical samples were prepared in a volumetrically equivalent manner, insuring comparability. The same in-solution trypsin digestion protocol and de-salting protocol were followed as described in the previous section, (page 83). Digests were re-suspended in 0.1% formic acid, 5% acetonitrile to a concentration of 0.5 µg/µL. The optimized LC-MS/MS chromatographic method and acquisition settings were used to generate the inclusion list. Digested and re-suspended samples were injected (2 µL) in duplicate.

Inclusion lists were evaluated against a tryptic digest of subject serum (60 minutes), collected after consumption (Baumert et al., 2009). Only those parent ions with 2, 3, or 4+ charge states were included.

C. Exclusion and Inclusion List Evaluation

The prepared exclusion and inclusion lists were evaluated against subject serum samples (Baumert et al., 2009) collected after peanut consumption (60 minutes) and prepared for analysis as described previously (page 83). The standard LC-MS/MS acquisition workflow was used for evaluation of each individual list, exclusion and inclusion, as well as the use of the two lists together (page 81; method 3). Only those parent ions with 2, 3, or 4+ charge states were included. Digested and re-suspended samples were injected (2 µL) in duplicate.

D. LC-MS/MS Data Analysis

All data for exclusion and inclusion lists were analyzed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada) against the appropriate database, as previously described (page 84). For exclusion lists, the top 5,000 serum specific m/z events were used. For inclusion lists, all peanut specific m/z events, up to the top 5,000 identifications, were included in an inclusion list. All data analysis parameters remained unchanged from those previously described (page 82).

VII. Evaluation of Commercial Depletion Kits

The following commercial depletion kits (1) Pierce Top2 Abundant Depletion spin columns (albumin, IgG), (2) PureProteome Albumin magnetic beads (albumin), (3) Pierce albumin depletion kit (cibacron dye binding), and (4) BioVision Protein G Sepharose (IgG depletion), were evaluated for removal of targeted serum proteins. Table

2-2 describes the mechanisms employed and serum proteins targeted by each depletion method.

Depletion kits were evaluated for target protein specificity using two key criteria (1) effective depletion of targeted serum proteins and (2) absence of depletion of peanut proteins (e.g. peanut proteins remain in depleted analytical samples). A model matrix of human serum and (raw) peanut extract (10:1 (w/w) protein) were analyzed for specific protein removal. The serum and peanut extract were individually diluted 10-fold prior to combining. Control samples were prepared using equivalent volumes of (a) serum or (b) peanut extract, substituted with 0.01 M PBS for each removed matrix component.

Depletion kit	Manufacturer	Targeted proteins	Depletion mechanism	Maximum protein load (µg)
Pierce Top2 Abundant Depletion columns	Pierce	Albumin, IgG	Resin, spin columns, anti-HSA, anti-IgG	600 µg sera protein
PureProteome Albumin Magnetic Beads	EMD Millipore	Albumin	anti-HSA	1:30 (v/v) serum:depletion slurry
Pierce Albumin Depletion Kit	Pierce	Albumin	Immobilized Cibacron blue dye agarose resin	2,000 µg (2 mg) human serum albumin
BioVision Protein G Sepharose	BioVision	IgG	Protein G conjugated to Sepharose beads	1:2 (v/v) serum:resin

Table 2-2. Commercial serum depletion kits for evaluation in de-complexing strategies

A. Pierce™ Top2 Abundant Depletion Spin Columns (albumin, IgG)

The Pierce Top2 Abundant Depletion spin columns were evaluated for removal of albumin and IgG from serum. The depletion spin columns contained immobilized anti-human serum albumin and anti-IgG antibodies for protein removal. Samples were depleted according to manufacturer's instructions, described as follows. The prepared analytical samples, serum-peanut, serum, and peanut, were applied to depletion spin columns (10 µL). Spin columns were mixed using an end-over-end mixer for 30 minutes at room temperature. After mixing, depleted serum was eluted into a collection tube and centrifuged for 2 minutes at 1,000 x g (Thermo Scientific™ Sorvall Legend Micro 17).

Depleted samples were prepared for LC-MS/MS analysis as previously described (page 80).

Control, non-depleted samples were simultaneously prepared. During depletion, the original sample volume is diluted after mixing with the depletion resin slurry. Therefore, non-depleted samples were prepared equivalently by diluting non-depleted samples (serum-peanut, serum, peanut) in 0.01 M PBS to the same volume. Non-depleted controls were prepared for LC-MS/MS as described below.

i. In-Solution Reduction, Alkylation, and Trypsin Digestion for LC-MS/MS Analysis

All samples, depleted and non-depleted, were prepared for LC-MS/MS analysis, as previously described (page 80). However, during this experiment, 4 µL of each sample (maximum 10 µg protein) were diluted to a final concentration of (0.5 µg/µL) with 50 mM ammonium bicarbonate (ABC). Following de-salting, peptides were eluted in 50% acetonitrile and dried. Peptides were re-solubilized in 5% acetonitrile, 0.1% (v/v) formic acid to a maximum concentration of 0.25 µg/µL prior to injection on LC-MS/MS. DDA analysis was conducted by injecting samples (2 µL) in duplicate, with a maximum protein load of 500 ng on column, using the previously optimized method (page 84).

ii. Modification of Salt (NaCl) Concentration of Pierce Top2 Albumin Depletion Kit

To reduce potential binding of peanut proteins to the depletion matrix, we evaluated the elution buffer salt concentration (NaCl) for recovery of peanut proteins using the Pierce Top2 Abundant Depletion spin columns. After depletion, samples were eluted in 0.01 M PBS, 0.15 NaCl, 0.02% azide, pH 7.4. To assess column binding specificity, the peanut only matrix was evaluated identically to the previously described methodology (*Pierce™ Top2 Abundant Depletion Spin Columns (albumin, IgG)).* The salt (NaCl) concentration(s) of the kit elution buffer, (0.01 M PBS, 0.15 M NaCl, 0.02% azide, pH 7.4) were prepared at the 0.15, 0.25, 0.50, 0.75, 1.00 M NaCl. Depleted eluates were evaluated for the presence of peanut proteins by discovery LC-MS/MS, as described previously (pages 80). SDS-PAGE was not performed on these samples.

iii. LC-MS/MS Data Analysis

Acquired data were processed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada). Protein identifications were made by searching a (a) sequence database of peanut (*Arachis hypogaea*, taxon identifier 3818), and (b) sequence database of human (*Homo sapiens*, taxon identifier 9606) as available in UniProt on 8/31/2016. Mass spectral were normalized for the total ion current (TIC) and semi-quantified using label-free quantification. The following criteria were used for protein identifications: no missed tryptic cleavages, fixed modification of carbamidomethylation of cysteine; variable modifications of methionine, oxidation, and

hydroxylation of proline; parent mass error tolerance 2 ppm, fragment error tolerance 0.02 Da; FDR 1%; charge states of +2, +3, and +4.

B. PureProteome™ Albumin Magnetic Beads

PureProteome™ Albumin Magnetic Beads (LSKMAGL10, EMD Millipore, Billerica MA, USA), were evaluated in a similar manner to the previously described depletion kit (Pierce Top2 Abundant Depletion Columns). Depleted and non-depleted samples (serum-peanut, serum, and peanut) were evaluated for specific removal of albumin and recovery of peanut proteins by discovery LC-MS/MS. The PureProteome™ Magnetic Beads bind albumin by immunoaffinity capture. The anti-albumin antibodies are coupled to magnetic beads, facilitating efficient separation of bound albumin and capture antibodies. Following incubation and separation, the unbound sample fraction, or depleted fraction, was removed and prepared for LC-MS/MS.

Control (non-depleted) samples were prepared in a volumetrically equivalent manner by diluting samples in 0.01 M PBS to the same volume as depleted samples. The maximum concentration of the non-depleted samples was 2.1 µg/µL.

Tryptic digestion was performed, in a volumetrically equivalent manner using the previously described protocol (page 80) modified in-solution trypsin digestion (Thermo Scientific™, 89895) by diluting a maximum of 10 µg protein in 50 mM ammonium bicarbonate to a maximum concentration of 0.6 µg/µL. All protein samples were reduced, alkylated, trypsin digested, and de-salted, as described previously (page 80). Following de-salting peptides were eluted in 50% acetonitrile and dried (page 83).

Peptides were re-suspended in 0.1% formic acid, 5% acetonitrile. A total of 2 µL of each sample, maximum 600 ng protein, was injected onto the analytical column for DDA analysis using the previously developed acquisition method (page 84). Data were analyzed as previously described (page 90).

C. Pierce™ Albumin Depletion Kit

The Pierce™ Albumin Depletion Kit (85160, Pierce Biotechnology, Rockford, IL, USA) depletes albumin by immobilized Cibacron Blue dye agarose resin. This depletion kit was evaluated using the same model matrices (serum-peanut, serum, peanut) as described previously (Pierce™ Top2 Abundant Depletion spin columns).

Albumin depletion resin, spin columns, and binding/wash buffer (0.025 M Tris, 0.075 M NaCl; pH 7.5) were provided with the kit. Following the manufacturer's instructions, spin columns were equilibrated by applying albumin depletion resin (400 µL) followed by binding/wash buffer (200 µL). After column equilibration, 50 µL of each analytical sample was applied and incubated for two minutes, and centrifuged (12,000 x g, 1 minute; Thermo Scientific™ Sorvall Legend Micro 17). The flow-through was retained and reapplied to the column. This process of sample application and centrifugation was repeated four times on the flow-through fractions to ensure albumin removal. Depleted eluates were retained for discovery LC-MS/MS analysis. Non-depleted (control) samples were prepared equivalently, by diluting control samples in 0.01 M PBS to the same final volume as depleted samples.

i. In-Solution Reduction, Alkylation, and Trypsin Digestion for LC-MS/MS Analysis

Depleted and non-depleted samples were prepared in duplicate using the previously described digestion methodology (page 80), and evaluated by discovery LC-MS/MS (page 84). Samples were prepared in a volumetrically equivalent manner allowing all digests to be prepared identically. In total, 12 µL (maximum 30 µg protein) were used for digestion, and trypsin digestion parameters were adjusted for a 30 µg total protein digest. Following de-salting peptides were eluted in 50% acetonitrile and dried to completion (page 83). Prepared digests were re-suspended in 0.1% formic acid, 5% acetonitrile to a maximum final concentration of 0.5 µg/µL. Discovery LC-MS/MS runs were conducted by injecting (2 µL) in duplicate. Data analysis parameters for all depletion kits are described on page 90.

ii. Modification of Salt (NaCl) Concentration of Pierce™ Albumin Depletion Kit

We additionally evaluated the effect of salt concentration on depletion efficacy and recovery of peanut proteins in eluates. To do so, the salt (NaCl) concentration of the kit elution buffer (0.025 M Tris, 0.075 M NaCl; pH 7.5) was evaluated at 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 2.0 M Tris-NaCl concentrations. For this experiment, serum-peanut and peanut (control) samples were analyzed. The depleted eluates were evaluated for the presence of peanut proteins visually by SDS-PAGE, as described below. LC-MS/MS analysis was not conducted on these samples.

iii. Modification of Binding/Wash Buffer pH of Pierce™ Albumin Depletion Kit

We evaluated the effect of pH of the binding/wash buffer (0.025 M Tris, 0.075 M NaCl; pH 7.5) for depletion efficacy and recovery of peanut proteins in eluates. The pH levels evaluated include 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, and 7.5. The buffer pH was adjusted by preparing in-house solutions of 0.025 M Tris, 0.075 M NaCl at their respective pH's. Depleted samples eluted with their respective pH buffers were evaluated by SDS-PAGE, as described below. LC-MS/MS was not performed on these samples.

iv. SDS-PAGE of Pierce™ Albumin Depletion Kit (NaCl, pH)

Depleted eluates, prepared at modified NaCl concentrations or pH levels, were evaluated for the presence of peanut proteins using 1D-gel electrophoresis. SDS-PAGE was performed using NuPAGE Bis-Tris Mini Gels 4-12% (1.0 mm, 12 wells) under reducing conditions and constant voltage (200V) for 40 minutes in an XCell SureLock Mini Cell Electrophoresis System (Invitrogen Life Technologies).

Fifteen microliters of NaCl sample eluates and 22.5 µL of pH level eluates were diluted in 4 x concentrated Laemmli buffer and 1% β-mercapto-ethanol (BME). All samples were reduced by heating for 5 minutes at 95°C and 20 µL each were loaded into each gel well. Precision Plus Protein Dual Xtra Standards (Bio-Rad) were used as the molecular weight (MW) standard. Gels were stained using Coomassie Brilliant Blue R-250 (#1610436, Bio-Rad) stain, followed by de-staining, rehydration, and imaging (Coomassie Brilliant Blue R-250 Destaining solution, #1610438, Bio-Rad).

D. BioVision Protein G-Sepharose (IgG Depletion)

We evaluated depletion of IgG, in efforts to decrease sample complexity prior to LC-MS/MS analysis for peanut proteins in serum. IgG depletion studies were performed in a similar manner as those described previously. However, for this depletion experiment, we used a commercially available serum (ImmunO, human serum sterile, MP Biomedicals LLC, Solon, OH, USA) for preparation of the matrix samples. Equivalently prepared samples of sera-peanut, sera, and peanut (60 µL) were added to 40 µL Protein G-Sepharose (BioVision 6511) resin and 200 µL binding buffer (0.01 M PBS, pH 7.4). Samples were incubated at 4°C for two hours in an end-over-end rotator. After incubation, depleted sample(s) (80 µL) were removed. The remaining resin was washed three times with binding buffer (200 µL), removing only 150 µL during the final wash in order to keep the Protein-G resin suspended.

Control, non-depleted, samples were prepared similarly by diluting controls with 0.01 M PBS to the same final concentration of IgG depleted samples.

IgG depleted and non-depleted (control) samples were evaluated by SDS-PAGE on an equal weight basis under reducing conditions using NuPAGE Bis-Tris Mini Gels 4-12% (1.0 mm, 12 wells) (Invitrogen Life Technologies). Depleted and non-depleted samples were diluted with 4 x Laemmli buffer with 1% β-ME and reduced at 95°C for 5 minutes. Equal weights of protein(s) were prepared and loaded into each gel well. In summary, 11 µg serum protein, 1 µg peanut protein, and 12 µg combined sera-peanut were loaded for both depleted and non-depleted controls for equivalent comparisons. 1D electrophoresis was conducted as previously described (page 94).

VIII. Organic Solvent Fractionation

We investigated the use of organic solvent fractionation as a method for sample de-complexation. Here, we separated our analytical sample by solubility into three soluble fractions and an insoluble pellet. Using this methodology, we are able to retain the entirety of the sample (serum, serum-peanut) while reducing background interferences.

Our methodology was adapted from a previously published method by Liu et al., 2014. One volume serum was combined with ten volumes isopropanol (IPA) with 1% trichloroacetic acid (TCA) by weight. Commercially prepared serum (ImmunO, human serum sterile, MP Biomedicals LLC, Solon, OH, USA) was used for our model matrix samples. As conducted before, samples of serum, peanut, and serum-peanut samples were prepared in a volumetrically equivalent manner. Samples were fractionated and evaluated by SDS-PAGE and LC-MS/MS using the optimized methodology for detection of low abundance peptides (page 83). Samples were vortexed for 2 minutes, followed by centrifugation (Beckman GS-15R centrifuge) at 1,500 x g at 5°C for 5 minutes. The supernatants were removed and retained for analysis. The remaining pellets were re-suspended in methanol (200 µL) and centrifuged (Thermo Scientific™ Sorvall Legend Micro 17) at 1,500 x g for 2 minutes at room temperature. The supernatant was removed and retained. Due to the volatile and evaporative nature of the two retained supernatants, isopropanol-TCA and methanol, the collected supernatants were dried in a centrifugal evaporator (Jouan RC-10.10; RCT-90; Winchester, VA, USA) and re-suspended prior to

downstream analyses. Control samples were prepared identically and not subjected to organic solvent fractionation.

A. SDS-PAGE of Organic Fractionated Samples

SDS-PAGE (NuPAGE Bis-Tris Mini Gels 4-12% (1.0 mm, 12 wells) (Invitrogen Life Technologies)) was conducted on all fractions for each sample using equal protein loading as determined by 2D quant. Each fraction was diluted with 4 x Laemmli buffer with 1% β-ME and reduced by heating at 95°C for 5 minutes. In summary, 0.6 µg peanut protein, 7 µg serum protein, and 7.6 µg total protein (sera-peanut) in each respective sample were loaded into gel wells. 1D electrophoresis was conducted as previously described (page 94)

B. In-Solution Reduction, Alkylation, and Trypsin Digestion for LC-MS/MS Analysis

Each fraction was analyzed to determine protein (peanut and serum) fractionation patterns by LC-MS/MS on an equal protein weight basis, as determined by 2D quant. Control samples (sera-peanut, sera, peanut) were equivalently prepared (10:1 w/w). In total, 20.3 µL of each sample were fractionated as described above. The supernatants were dried under vacuum by centrifugal evaporation (Jouan RC-10.10; RCT-90; Winchester, VA, USA), followed by re-suspension in 5% acetonitrile to a final volume of 31.5 µL. Control, non-fractionated liquid samples, were also prepared to a final concentration of 5% acetonitrile in 31.5 µL total sample volume (0.9 µg/µL for sera-

peanut, the most concentrated sample). Tryptic digestion was performed based on 30 µg total protein in digestion. Following re-suspension and dilution the samples were diluted, reduced, alkylated, and trypsin digested, and de-salted (C18) as described previously (page 80). Following de-salting peptides were eluted in 50% acetonitrile and dried to completion. Samples were re-suspended in 0.1% formic acid, 5% acetonitrile to a maximum peptide concentration of 1 µg/µL. Re-suspended peptides were injected (1 µL) in duplicate for discovery LC-MS/MS analysis using the previously described LC-MS/MS method (page 84). Data were analyzed as previously described (page 82).

IX. Evaluation of Multiple De-Complexation Methods (Organic Solvent Fractionation and IgG Depletion)

We investigated the combination of two de-complexation techniques, organic solvent fractionation and IgG depletion. Evidence of IgG interactions with peanut proteins was recently published (JanssenDuijghuijsen et al., 2017). Therefore, our aim was to incorporate IgG depletion in combination with organic solvent fractionation to decrease sample complexity, improving likelihood of peanut protein detection by discovery LC-MS/MS analysis.

For this study, we prepared two replicates of the same model samples (sera-peanut, sera, peanut). The first set were used as controls and not de-complexed (e.g. no IgG depletion or fractionated). The second set of samples were IgG depleted followed by fractionation. SDS-PAGE was not performed on these samples. Samples were evaluated by LC-MS/MS using the previously described detection method (page 84).

A. Preparation of Matrix Samples

A set of volumetrically equivalent control samples (sera-peanut, sera, peanut) were prepared as described previously (page 86). These samples were not IgG depleted or fractionated.

A set of samples for de-complexation were prepared in a volumetrically equivalent manner (sera-peanut, sera, peanut; 10:1 (w/w)), for IgG depletion and organic solvent fractionation.

The aim of these evaluations was to assess the impact of IgG on peanut detection, by determining recovery of peanut. To do so, we prepared an IgG depleted serum sample using the same methodology as previously described (page 95). Following serum IgG depletion, the peanut extract was added in a volumetrically equivalent manner to controls. IgG depletion and fractionation were performed as previously described in their appropriate methodology sections (95, 96).

B. IgG Depletion and Organic Solvent Fractionation

Samples (sera, sera-peanut) were IgG depleted as previously described. We evaluated serum (non-depleted), IgG depleted serum with peanut extract, and serum-peanut. We analyzed commercial serum and serum collected from peanut allergic subjects. The subject clinical allergenicity characteristics are described in Supplementary Table 2-1. All sample preparations were subjected to IgG depletion prior to organic solvent fractionation.

Following IgG depletion, samples were fractionated using the previously described organic solvent fractionation protocol, (1) serum, (2) IgG-depleted serum-peanut, and (3) serum-peanut. The same fractions were collected as previously described (page 96), (1) isopropanol-TCA supernatant, (2) methanol supernatant, and (3) pellet.

Following sample preparation, all control (e.g. liquid) samples and dried supernatants were diluted or re-solubilized to 5% acetonitrile to a final volume of 31.5 µL. Tryptic digestion (maximum 30 µg protein in digestion) and de-salting (C18) were performed on all samples as previously described (page 80). Following de-salting peptides were eluted in 50% acetonitrile and dried to completion. Digested samples were re-suspended (v/v) with 0.1% formic acid, 5% acetonitrile to a concentration of 0.5 µg/µL. Samples were injected (2 µL) in duplicate for LC-MS/MS analysis.

RESULTS AND DISCUSSION

I. Evaluation of the Peanut Flours Preparation

Raw and roasted peanuts were extracted in two PBS buffer concentrations, 0.01 and 0.1 M. Protein concentrations of each extract were determined by 2D Quant assay. The results of the protein concentration determination reflect the observations from the SDS-PAGE (Table 2-3). For both peanut preparations, raw and roasted, extraction in 0.01 M PBS yielded the most protein, which will be used in subsequent experiments of this chapter. Visualization of the SDS-PAGE gel indicated more protein was extracted under 0.01 M PBS in both raw and roasted preparations (Figure 2-1).

Roasted peanut extracts were also prepared however, thermal processing negatively affects the solubility of peanut proteins, particularly Ara h 1, which aggregates at high temperatures (>85°C) (Koppelman et al., 1999). Although Ara h 1 is less extractable in roasted peanuts, it is still present in the peanut seed and exposed to the immune system upon consumption. In order to represent all major allergenic proteins, raw peanuts were prepared and used in method evaluations.

Peanut preparation	Concentration PBS (M)	Average µg/µL	Standard deviation
Raw	0.01	2.07	±0.13
	0.1	1.10	±0.44
Roasted	0.01	0.54	±0.27
	0.1	0.39	±0.08

Table 2-3. Protein concentrations of peanut extracts prepared in 0.1 or 0.01 M PBS determined by 2-D Quant assay (GE Healthcare).

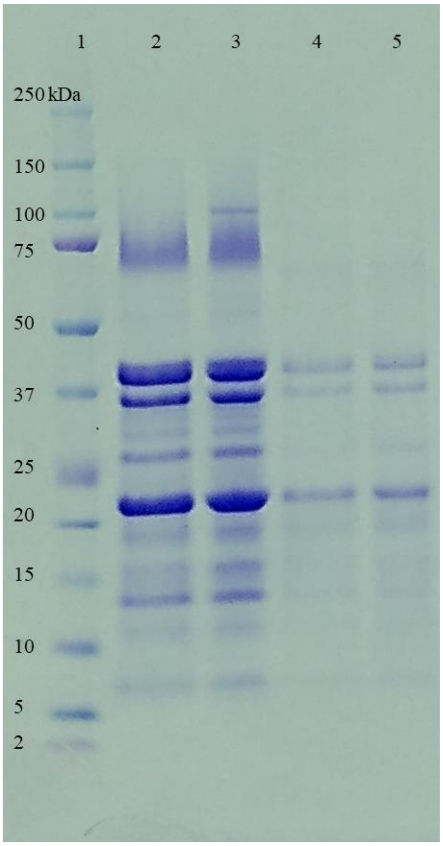


Figure 2-1. Evaluation of peanut protein extracts by reducing SDS-PAGE. Lane (1) Protein marker standard, (2) raw peanut extracted in 0.1 M PBS, (3) raw peanut extracted in 0.01 M PBS, (4) roasted peanut extracted in 0.1 M PBS, (5) roasted peanut extracted in 0.01 M PBS

II. Discovery LC-MS/MS Acquisition Method Optimization

A trypsin digested serum sample, prepared identically to the model matrix sample (sera-peanut), was used for evaluation of LC-MS/MS acquisition methods. During LC-MS/MS method development, a number of MS acquisition settings can be optimized to improve sensitivity and detection of low abundance target peptides. A series of instrument acquisition settings including dynamic exclusion time (seconds), MS2-fill time (msec), TopN acquisition, and MS1 peptide charge states (Table 2-1) were evaluated. All other instrument parameters were kept the same.

.

Method	#MS2	# peptide spectra	# peptides	# proteins
1	9392	1706	385	133
2	4238	731	498	206
3	9276	1569	909	323

Table 2-4. Results of discovery LC-MS/MS acquisition settings for improvement of detection of low abundance peptides and proteins.

Figure 2-2 illustrates the relative abnundance of peanut peptides detection by each method. Table 2-4 contains the number of MS2 spectra, peptide spectra, individual peptides, and individual proteins identified. Method 1, included the Top10 MS 1 ion selection method, 3 s dynamic exclusion times 60 ms MS2 fill time and parent ions (MS1) with +1 to +5 charges, does not provide adequate ion filtering prior to detection by

the MS (Table 2-4). Method 1 reported the highest number of MS2 spectra and peptide spectra; and reported the least amount of peptides and proteins identified. In Method 2, parent ion charges of 2, 3+ were monitored, and dynamic exclusion time was increased. This resulted in more peptides and proteins identifications, while also reporting lower numbers of MS2 and peptide spectra (Table 2-4). Lastly, Method 3, monitored parent ions, 2, 3, 4+, had a decreased dynamic exclusion time (20 msec), and an increased MS2 fill time (120 msec). This method identified the most peptides and proteins, while reporting nearly similar figures of MS2 and peptide spectra than Method 1 (Table 2-4). For these reasons, we chose Method 3 for use in all subsequent discovery experiments.

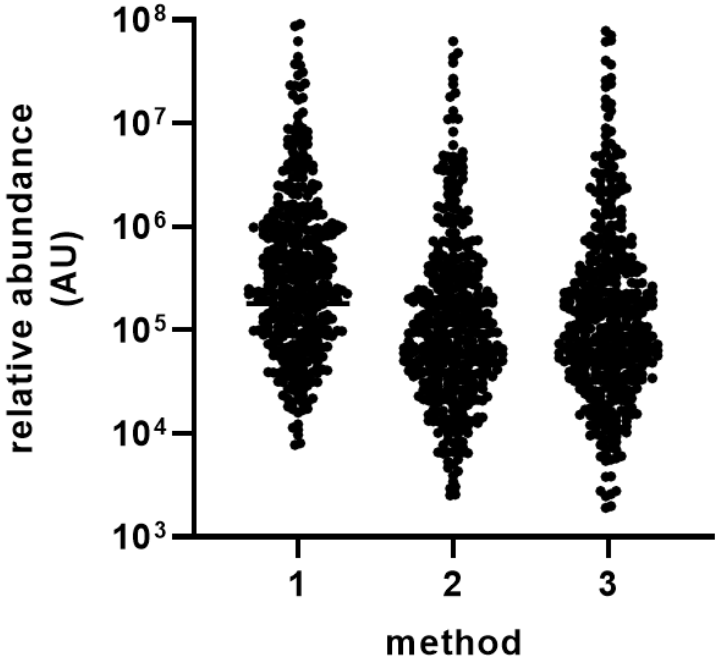


Figure 2-2. Optimization of LC-MS/MS discovery methods. For each method (x-axis), the relative abundance of individual peanut peptides are shown.

III. Evaluation of Subject Serum Samples and Model Matrices of Sera-Peanut Matrix

Initial analyses of subject serum collected after peanut consumption (Baumert et al., 2009) were negative for peanut using an initial discovery LC-MS/MS method (data not shown). The lack of detection of peanut proteins is unsurprising since the expected *in vivo* concentration of peanut after consumption is estimated at extremely low concentrations.

The subject serum used in our study were previously determined positive by competitive immunoassay developed to detect DRP-Ara h 2 (Baumert et al., 2009). They reported an average detection of 131 ng/mL of DRP-Ara h 2 over a 24-hour period after consumption (Baumert et al., 2009). A second study conducted by JanssenDuijghuijsen et al., 2017 analyzed collected serum after feeding trials using an Ara h 6 sandwich ELISA. Here, they reported an average of 0.3 ng/mL Ara h 6, equivalent to 6 ng total peanut protein/mL (JanssenDuijghuijsen et al., 2017). Based on these two studies, and our results, the current LC-MS/MS method and sample preparation protocols are not sufficiently sensitive for *in vivo* peanut detection.

These extremely low (ng/mL) concentrations are similar to the concentration of cytokines which have been detected at very low levels ranging from 1-100 pg/mL (0.001 – 0.1 ng/mL) by the use of antibody capture and subsequent LC-MS/MS (Anderson and Anderson, 2002; Schweitzer et al., 2002). The use of antibody capture methods are beneficial in decreasing background noise, and would likely produce methods with similar sensitivities to current ELISAs. However, we would not be able to assess changes in protein structure, conformation, or free antigen. Due to the anticipated low

concentrations of our target analytes (e.g. peanut proteins), we employed de-complexation strategies to improve overall detection.

IV. Inclusion and Exclusion Lists

The MS instrument used within this chapter was the Thermo Q Exactive Plus™ Hybrid Quadrupole-Orbitrap™ MS (Thermo Scientific™) which has the ability to exclude selected m/z events by using an exclusion list. An exclusion list is a useful tool capable of excluding specified m/z events. Since serum is a protein rich matrix, it generates a substantial number of serum specific m/z events that are recorded by the MS. If a parent m/z (MS1) ion is defined in the exclusion list; this identified ion is excluded from fragmentation (Koelmel et al., 2017). We evaluated the use of exclusion lists in combination with discovery LC-MS/MS acquisitions. As such, we applied this methodology to our model serum-peanut sample matrix and subject serum samples (baseline, 60 minute).

During the initial DDA acquisitions of baseline serum, we used a 'Top20' data acquisition method in which the twenty most abundant parent ions from MS1 are collected for MS2 fragmentation, and, consequently the less abundant ions are not recorded during a particular acquisition scan. By excluding the most abundant m/z events identified in the standard LC-MS/MS acquisition workflow, we anticipate improved detection of lower abundance peptides, and in particular peanut peptides in subject serum.

A. Exclusion List Evaluation

An exclusion list of serum specific m/z values was generated from LC-MS/MS data of baseline (0 hour) serum samples. All peptides included in the exclusion lists met the following criteria:

Exclusion list peptide criteria:

1. Charge states of 2, 3, or 4+

2. The top 5,000 serum identified m/z's based on signal intensity

3. Identified within the chromatographic gradient

4. Serum specific m/z values

The top 5,000 serum specific m/z values were included in the exclusion list. Serum specific m/z values were identified using PEAKS from the UniProt *Homo sapiens* database. Subject serum collected after peanut consumption (60 minutes), or 'active' serum, were analyzed with the generated exclusion list. Acquired data were compared to a peanut database (UniProt, *Arachis hypogaea*). No positive peanut peptides were detected in subject serum when analyzed with an exclusion list (data not shown).

B. Inclusion List Evaluation

The inclusion lists for peanut specific peptides were generated in the same manner as exclusion lists using prepared digested samples of peanut and sera-peanut. We identified 23 peanut proteins and 143 unique peanut peptides. All identified peanut

specific m/z values were compiled into an inclusion list included during acquisition. No positive peanut identifications were made when the inclusion list was included in the analysis of 'active' subject serum (60 minutes post consumption) (data not shown).

C. Evaluation of Combined Exclusion and Inclusion Lists

The overall aim of combining exclusion and inclusion lists was to reduce detection of the highly abundant serum proteins, which largely saturate the MS signal detector. We hypothesized using both ion selection techniques together would improve instrument sensitivity.

No positive peanut protein identifications were made when using the combination of inclusion and exclusion lists, paired with LC-MS/MS (data not shown). This suggests other de-complexation strategies are necessary to detect peanut proteins in serum. In many instances the use of inclusion or exclusion lists yields positive results with other complex matrices (e.g. human embryonic stem cells, characterization of human plasma lipidome), however, we found these tools to be unsuccessful, likely due to the complexity of the serum matrix (Bendall et al., 2008; Koelmel et al., 2017). Although we are able to exclude the most abundant (top 5,000) parent m/z events (MS1), this was not sufficient to identify peanut peptides in serum. Secondly, the pairing of a serum m/z exclusion list with a peanut m/z inclusion list was still unable to detect peanut peptides.

Serum is a protein rich matrix exhibiting vast diversity among its constituent proteins (Anderson and Anderson, 2002). Due to this extreme diversity, it is likely that the low abundance serum proteins share similar sequence(s) and charge(s) to peanut

specific peptides. In order to obtain a comprehensive serum proteome, a mass spectrometer must be able to detect both the highly abundant proteins (e.g. albumin, immunoglobulins) down the to the lowest abundance serum proteins (e.g. cytokines, interleukins) (Anderson and Anderson, 2002). Serum proteins span ten orders of magnitude (Anderson and Anderson, 2002), whereas the Q Exactive mass spectrometer has a detectable range spanning only four orders of magnitude (Eliuk and Makarov, 2015). Therefore, in order to achieve detection at concentrations beyond the dynamic range of the MS instrument, an analytical method which enhances detection of low abundance peptides is needed.

For the first part of this chapter, we chose to first optimize the instrument scan settings of discovery methods. Detection by DDA methods is desirable because it allows other human proteins to be monitored during allergenic protein uptake. Another advantage of using shotgun methods is the ability to track overall changes in protein profiles using the acquired full-scan data, allowing the potential identification of biomarkers associated with consumption of allergenic proteins.

Overall, adjusting MS acquisition settings was not sufficient for detection of peanut proteins in serum. Due to the complexity of serum, and limitations of the instrument's dynamic range, further de-complexation methods are needed.

During this set of studies, we routinely injected approximately 500 ng of protein (10:1 sera:peanut) on a micro-flow HPLC column (100 mm x 1.0 mm, 1.9 um). This total amount of protein has been suitably detected in both discovery and targeted assays of complex food matrixes, as determined by preliminary studies performed prior to the beginning of this dissertation. We will use the optimized discovery method, Method 3,

for our evaluations of detection. Consequently, we began our studies by evaluating various de-complexation strategies, with an overall objective to reduce sample complexity prior to LC-MS/MS analysis in order to detect peanut proteins in serum.

V. Evaluation of Commercial Depletion Kits

To evaluate the utility of various de-complexation strategies, we prepared a model matrix of baseline subject serum (0 hour), collected by Baumert (2009) spiked with a raw peanut extract. Since the anticipated *in vivo* concentration(s) of peanut protein are expected to be extremely low, we opted to use a high concentrated spiked model matrix (10:1 (w/w) serum:peanut) to robustly evaluate these de-complexation strategies. This model matrix, sera-peanut, was prepared to mimic the final analytical sample matrix (i.e. serum) collected from individuals after consuming peanut. The sera-peanut matrix was used for evaluations throughout this chapter.

Commercial depletion kits were evaluated on a case-by-case basis, and we summarize the results of all evaluated depletion kits here (Figures 2-3 to 2-6). Each kit selected for this study utilized different depletion mechanisms. All selected kits successfully removed their specified targeted serum proteins (Figures 2-4; 2-6). However, three out of the four commercial depletion kits (Pierce™ Top2, PureProteome™ Albumin, Pierce™ Albumin) either, partially or entirely removed peanut proteins in peanut containing samples (sera-peanut, peanut), as detected by loss of peanut peptides identified by PEAKS label-free quant. The IgG depletion kit (BioVision) was the only kit that exhibited high specificity and did not remove peanut proteins. For selected kits (Pierce™ Top2, Pierce™ Albumin) whose properties were adjusted (i.e. salt,

pH), no improvement in the recovery of peanut proteins (in serum matrix background or absence of sera matrix background). For this section, we will primarily discuss the results of the PureProteome™ Albumin depletion kit since many extraneous studies and evaluations were performed using this kit. Depletion figures for Pierce™ Albumin (Figures 2-3, 2-4) and PureProteome™ Albumin (Figures 2-5, 2-6) are shown below. No figures are shown for Pierce™ Top2, due to poor label-free quantification (LFQ) results, and lack of robust peptide detection. Data for the Pierce™ Top2 kits were evaluated visually. LFQ requires three unique peptides for quantification, however, three unique peptides could not be determined among evaluated matrix samples. Additionally, 3-fold less protein was used for Pierce Top2 digestions in comparison to other kits evaluated, another potential factor for poor peptide detection.

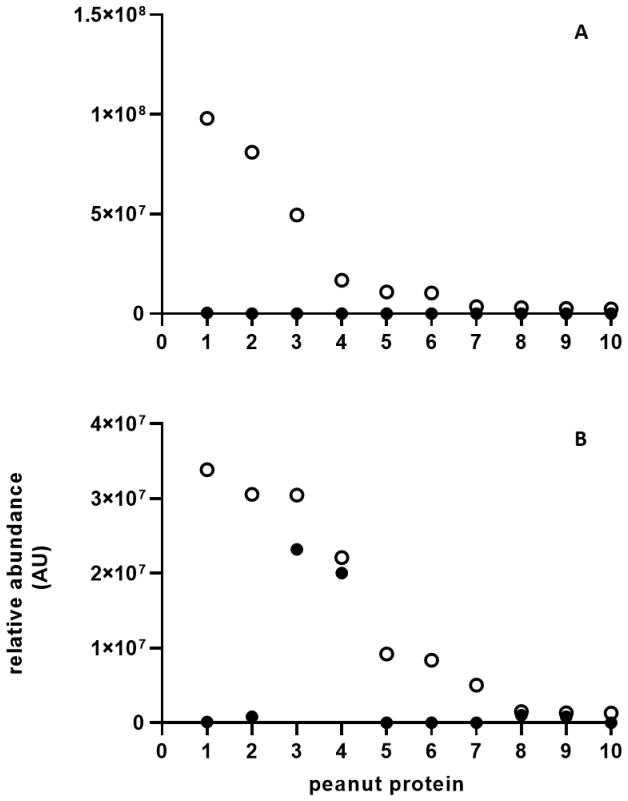


Figure 2-3: Depletion of peanut proteins in (A) peanut only and (B) sera-peanut matrices by the Pierce™ Albumin Depletion kit. The top 10 unique peanut proteins are represented in order of relative abundance. Open symbols represent non-depleted samples; closed symbols represent depleted samples.

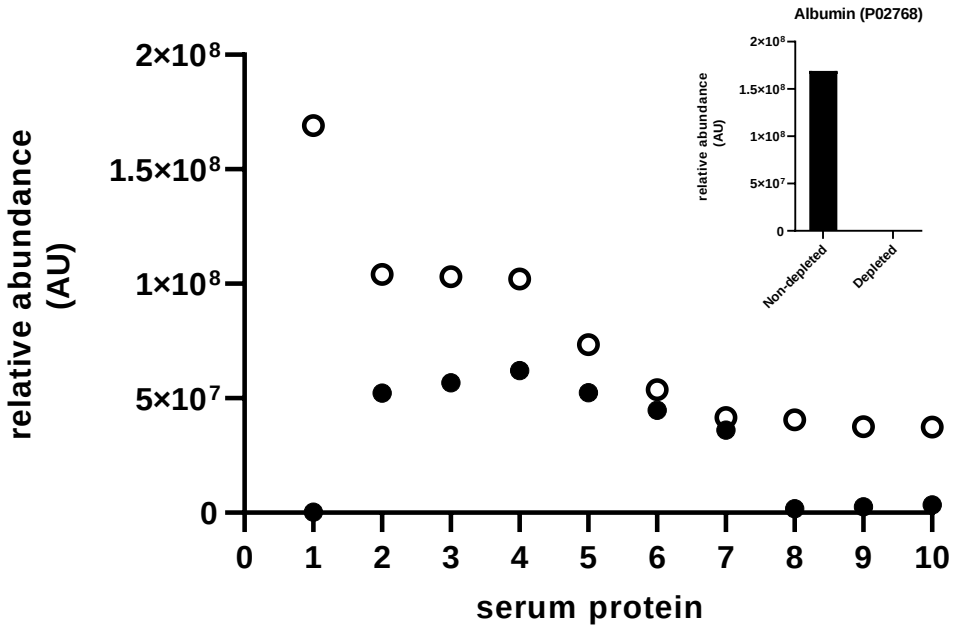


Figure 2-4: Depletion of serum proteins in sera-peanut matrix by Pierce Albumin Depletion kit. The top 10 peanut proteins are represented in relative abundance, determined by label-free quant. Open symbols represent non-depleted samples; closed symbols represent depleted samples. Inset demonstrates depletion of albumin only within this matrix using the same depletion kit.

The Pierce™ Top2 and PureProteome™ Albumin kits are antibody based depletion kits (anti-human serum albumin; anti-IgG). Antibody based methods are highly specific; however we observed removal of peanut proteins (Figure 2-5), strongly suggesting cross-reactivity between kit antibodies and peanut proteins. Commercially developed depletion kits are largely designed to increase detection of endogenous proteins in human serum such as cytokines, peptide biomarkers, hormone peptide, and lipoproteins (Pisanu et al., 2018). Therefore, we cannot insure specificity against exogenous food proteins.

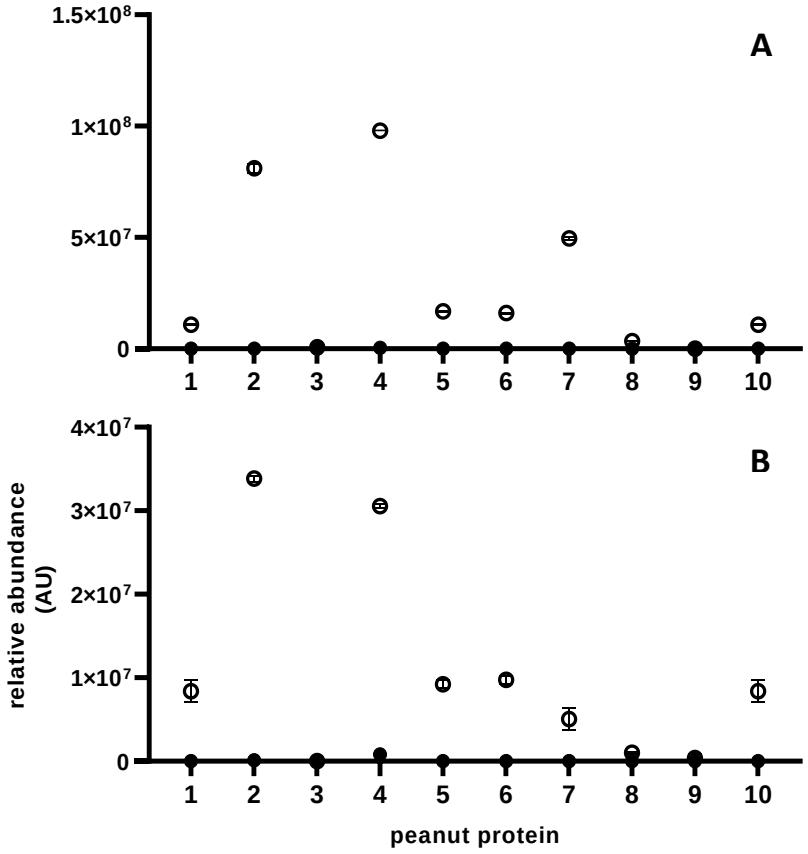


Figure 2-5: Depletion of peanut proteins in (A) peanut only and (B) sera-peanut matrices by the PureProteome Albumin Depletion kit. The top 10 peanut proteins are represented, determined by label-free quant. Open symbols represent non-depleted samples; closed symbols represent depleted samples.

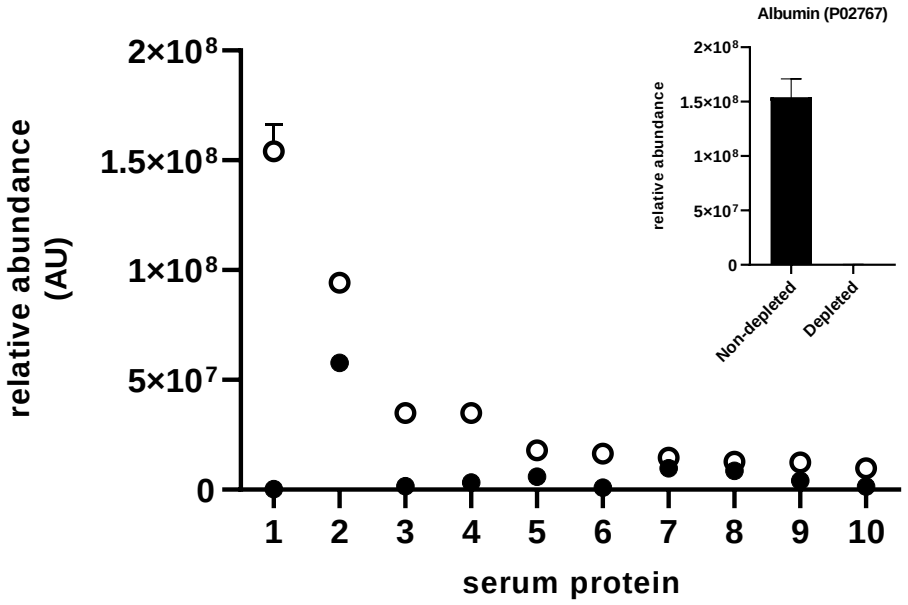


Figure 2-6: Depletion of serum proteins sera-peanut matrix by PureProteome Albumin Depletion kit. The top 10 peanut proteins are represented by relative abundance, determined by label-free quant. Open symbols represent non-depleted samples; closed symbols represent depleted samples. Inset demonstrates depletion of albumin only within this matrix using the same depletion kit.

The Pierce™ Albumin depletion kit utilizes immobilized Cibacron blue dye resin, and binds albumin through electrostatic or hydrophobic interactions (Gianazza and Arnaud, 1982; Travis and Pannell, 1973). Since peanut proteins were simultaneously removed during depletion, we modified the properties of selected depletion kits to reduce peanut protein removal by the matrix depletion. The salt concentrations of the wash/elution buffers of two kits Pierce™ Top2 Abundant Depletion spin columns and Pierce™ Albumin depletion were modified to assess recovery of peanut proteins. For both commercial kits, we observed no improvements in the recovery of peanut proteins at varying concentrations of salt in the wash and elution buffers. In fact, when peanut alone was applied to each respective depletion column, no peanut protein was recovered at the various salt concentrations evaluated (Figure 2-7). We then evaluated the proteins bound to the column matrix by heating columns (95°C, 5 minutes) after adding a solution of 4 x concentrated Laemmli buffer and 5% β-mercapto-ethanol. SDS-PAGE analysis of the aqueous phase revealed peanut proteins were primarily bound to the depletion column matrix (data not shown). This binding pattern was observed in both peanut and sera-peanut matrices. This suggests a high rate of non-specific binding to abundant proteins through antibody and dye-binding mechanisms.

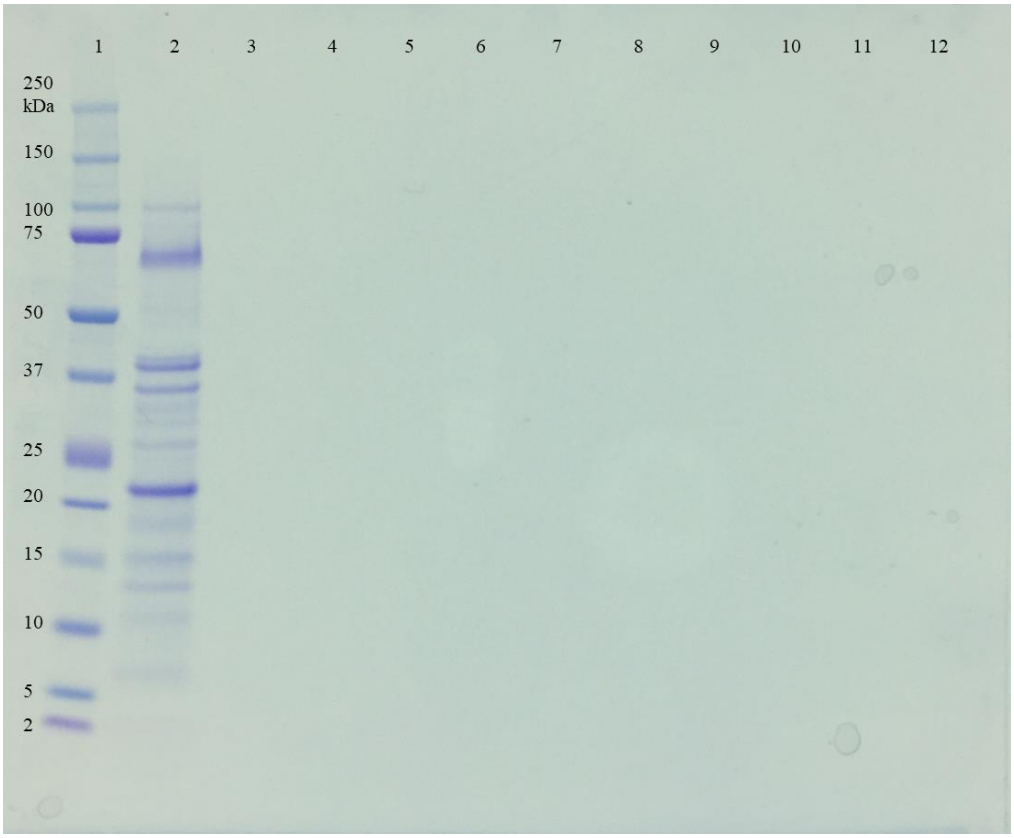


Figure 2-7. Evaluation of Pierce™ Albumin Depletion kit wash and elution buffer with modified NaCl concentrations by reducing SDS-PAGE. Lane (1) Protein marker standard, (2) raw peanut extract (0.01 M PBS), (3) column flow through, lanes (4-12) column flow through with wash buffer concentrations 0.075, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40 M NaCl, respectively

The protein binding interactions, both antibody binding and Cibacron dye-binding (Pierce™ Albumin Depletion), could not be disrupted by modifying the salt concentrations (Figure 2-7). The concentration of salt influences protein-surface or protein-protein binding properties by altering the binding affinities of proteins (Tsumoto et al., 2007). Cibacron dye-binding is an affinity based method, but is relatively non-

specific and has been described to bind other serum proteins including lipoprotein, antithrombin III, interferon, and other serum polypeptides (Gianazza and Arnaud, 1982; Thompson et al., 1975).

We further evaluated the Pierce™ Albumin depletion kit (Cibacron dye binding) by adjusting the pH and assessing recovery of peanut proteins by SDS-PAGE (Figure 2-8). No changes in elution of peanut proteins were observed at any pH level. This further suggests non-specific binding by the depletion matrix. By modifying the pH, we altered the ionic strength, which can influence protein-binding affinities and therefore protein elution from the depletion matrix.

Our aim was to disrupt interactions occurring between peanut proteins and capture antibodies. Decreasing the pH creates an increasingly acidic environment causing changes to the antigen-antibody binding properties, protein conformation, and partial denaturation (Hinderling and Hartmann, 2005; Kochansky et al., 2008). In the biological system evaluated, reducing the pH alone was not sufficient to disrupt the protein-binding interactions occurring between peanut, serum, and the depletion matrix.

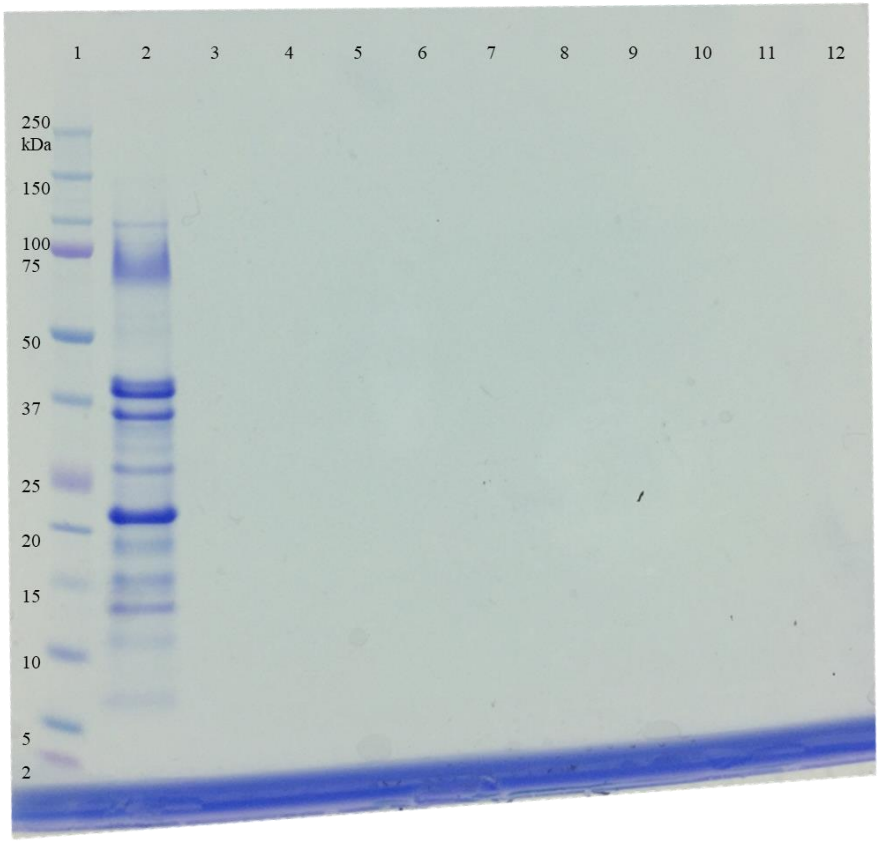


Figure 2-8. Evaluation of Pierce Albumin Depletion kit wash and elution buffer with modified pH by reducing SDS-PAGE. Lane (1) Protein marker standard, (2) raw peanut extract (0.01 M PBS), (3) column flow through, lanes (4-12) column flow through with wash buffer pH 7.5, 7.0, 6.5, 6.0, 5.5, 5.0, 4.5, 4.0, 3.5, respectively.

BioVision IgG depletion kits were successful in removing IgG from serum, while not co-depleting peanut proteins. Protein G, is a bacterial produced protein, and binds the Fc region of immunoglobulins with high affinity, and has been widely used for IgG purification, with kits for IgG depletion readily available (Björck and Kronvall, 1984).

Due to the specificity of protein-G, we hypothesized protein-G would remove IgG-peanut complexes, free IgG, and other IgG-serum complexes. JanssenDuijghssen et al., 2017 demonstrated *in vivo* interactions occurring between exogenous peanut proteins and endogenous IgG antibodies. We investigated removal of IgG using Protein-G Sepharose as a means to co-isolate IgG-peanut complexes from non-allergic patient serum. No LC-MS/MS data were acquired for IgG depletion experiments. SDS-PAGE was used to evaluate fractions from Protein-G Sepharose IgG depletion (Figure 2-9). In each panel, the non-depleted matrix is depicted in the first lane following the protein marker (lanes 2, 6, 10). The unbound fractions (lanes 3, 7, 11) represent the analyzed matrix after IgG depletion (e.g. serum post IgG depletion). The bound fractions (lanes 4, 8, 12) represent those proteins bound by Protein G Sepharose (e.g. depleted proteins from matrix). IgG was successfully removed from serum containing matrices (sera-peanut and serum alone) (lanes 3, 7). Peanut proteins were not depleted from the peanut containing samples (sera-peanut and peanut alone) (lanes 8, 12). Here, we observed the suspected IgG-peanut complexes were not depleted by Protein G Sepharose, as demonstrated by the absence of peanut proteins in lane 4, the bound fraction from the serum-peanut matrix. This suggests these complexes are unable to be removed from the serum matrix. Oda et al., 2003 reported a change in IgG conformation occurs upon antigen binding, resulting in the inability of Protein G Sepharose to bind complexed IgG, similar to our results (Oda et

al., 2003). Additional interactions with albumin or other endogenous proteins could also be influencing peanut behavior and overall detection. Due to the lack of peanut proteins present in the protein-G sepharose fraction following depletion (Lanes 4, 8, 12), it is evident IgG-peanut complexes were not effectively depleted. We suspect protein G is only able to remove free IgG.

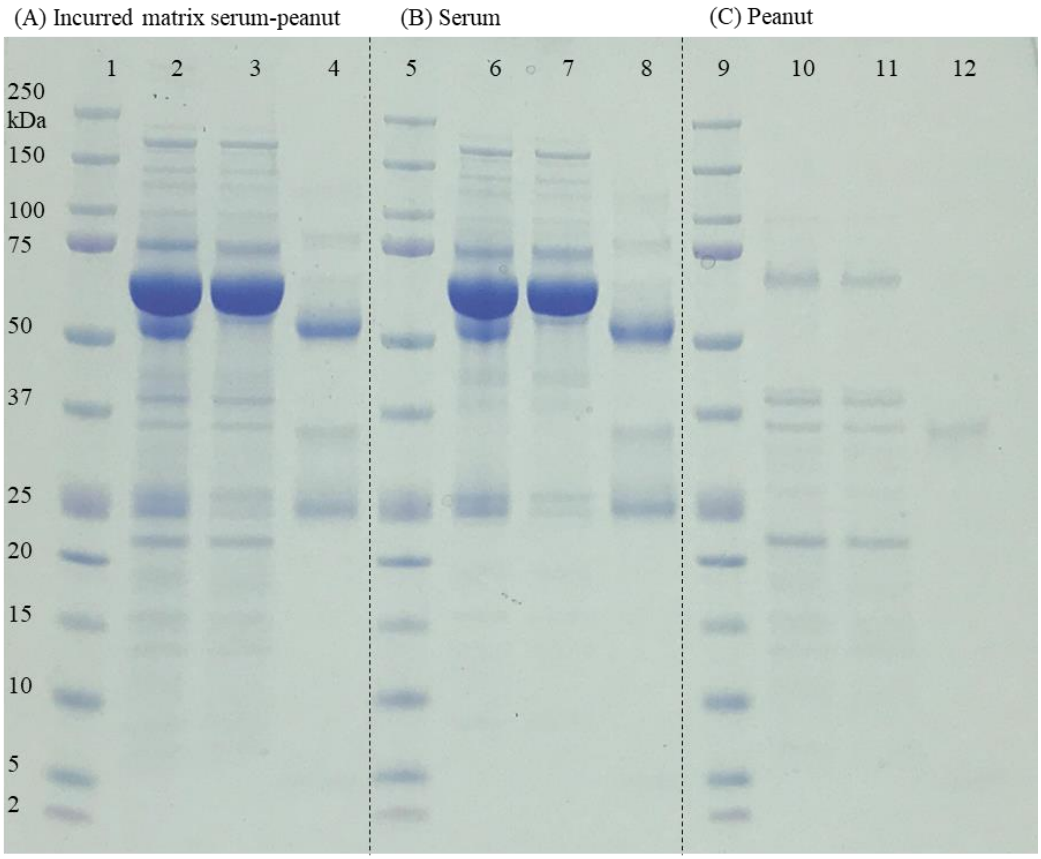


Figure 2-9. Evaluation of BioVision Protein G Sepharose IgG depletion by SDS-PAGE of prepared matrices (A) serum incurred with raw peanut, (B) serum, (C) raw peanut. Lanes (1, 5, 9) Protein marker standard (kDa), (2) sera-peanut matrix (non-depleted), (3) Unbound sera-peanut after IgG depletion, (4) Bound depleted fraction from serum-peanut, (6) serum (non-depleted), (7) Unbound serum after IgG depletion, (8) Bound depleted fraction from serum (10) peanut (non-depleted) (11) Bound peanut after IgG depletion (12) Unbound depleted fraction from peanut.

Allergenic proteins induce specific antibodies including IgG1, IgG4, and IgE (Aalberse et al., 2009; Jutel and Akdis, 2011). The antibody isotype IgG4 is induced after extended and repeated exposures to allergenic proteins, typically in low doses. These repeated, low-dose exposures, as performed in immunotherapy treatments, do not generally elicit a full immunological response (Aalberse et al., 2009; Jutel and Akdis, 2011). Interactions between allergenic food proteins and IgG antibodies, the predominant antibody class, have been reported in both allergic and non-allergic patients (JanssenDuijghuijsen et al., 2017; Murphy, 2012; Platts-Mills et al., 2001; Wachholz and Durham, 2004). For allergic patients undergoing immunotherapy, an increase in prevalence of IgG4 antibodies has been attributed to the acquiring of tolerance (Platts-Mills et al., 2001; Wachholz and Durham, 2004). This suggests non-symptomatic IgG-food protein interactions preclude IgE-food protein binding events (Platts-Mills et al., 2001; Wachholz and Durham, 2004).

Reported interactions between endogenous IgG antibodies and allergenic food proteins (peanut, Ara h 6), in non-allergic individuals, resulted in reduced detection by immunoassay (JanssenDuijghuijsen et al., 2017). To understand the overall role of IgG concentration on tolerance, more studies need to be conducted. We show similar findings to reports by Oda et al., 2003 and JanssenDuijghuijsen et al., 2017. Furthermore, interactions between allergenic food proteins and endogenous antibodies (IgG) prevents adequate detection of exogenous peanut proteins, a key problem for *in vivo* analysis of allergenic food proteins using ELISA. However, during LC-MS/MS sample preparation, these complexes are disrupted due to reduction and trypsin digestion generating peptides. Using this approach, IgG depletion is not advantageous for LC-MS/MS since IgG-peanut

complexes are unable to be removed. Therefore, our serum matrix remains relatively complex, but only by removing a small amount (i.e. free IgG) of protein.

VI. Organic Solvent Fractionation

Organic solvent fractionation methods were investigated as an alternative to depletion methods in order to retain serum composition. The most abundant serum protein, albumin, is soluble in organic solvents (e.g. isopropanol, ethanol, methanol) and is readily fractionated from serum (Liu et al., 2014; Michael, 1962). Samples for evaluation were fractionated into three fractions, (1) isopropanol-TCA, (2) methanol, or (3) pellet precipitate and were analyzed by LC-MS/MS and SDS-PAGE. The primary goal of TCA organic fractionation is to enrich peanut protein relative to all other proteins in the sample.

The organic solvent fractionation patterns of serum, peanut, and sera-peanut, and their non-fractionated controls (stocks) are shown in Figure 2-10. A comparison of fractionation patterns, shows a majority of the albumin protein is removed (MW band 66.5 kDa) into the isopropanol-TCA fraction. The albumin is not entirely removed, with some remaining in the precipitated fraction (Figure 2-10).

We analyzed the fractionation patterns in more detail based on relative abundance. Depending on the matrix background, peanut proteins demonstrated different fractionation patterns (Figure 2-11). In the absence of a serum background, the peanut 2S albumins, Ara h 2 and Ara h 6, primarily fractionated into fraction 1, the isopropanol-TCA fraction. Whereas in the presence of a serum background (serum-peanut), these

proteins primarily fractionated into fraction 2, the methanol fraction. The cupin proteins, Ara h 1 and 3 exhibited similar fractionation patterns (Figure 2-11). In the absence of a serum background, the cupins fractionated largely into the fraction 2, whereas in the presence of a serum background, Ara h 1 fractionated largely into the precipitate.

The organic solvent fractionation method works well to remove albumin from the serum matrix and reduce its complexity. However, this de-complexation strategy is largely non-specific, and dependent upon protein solubility.

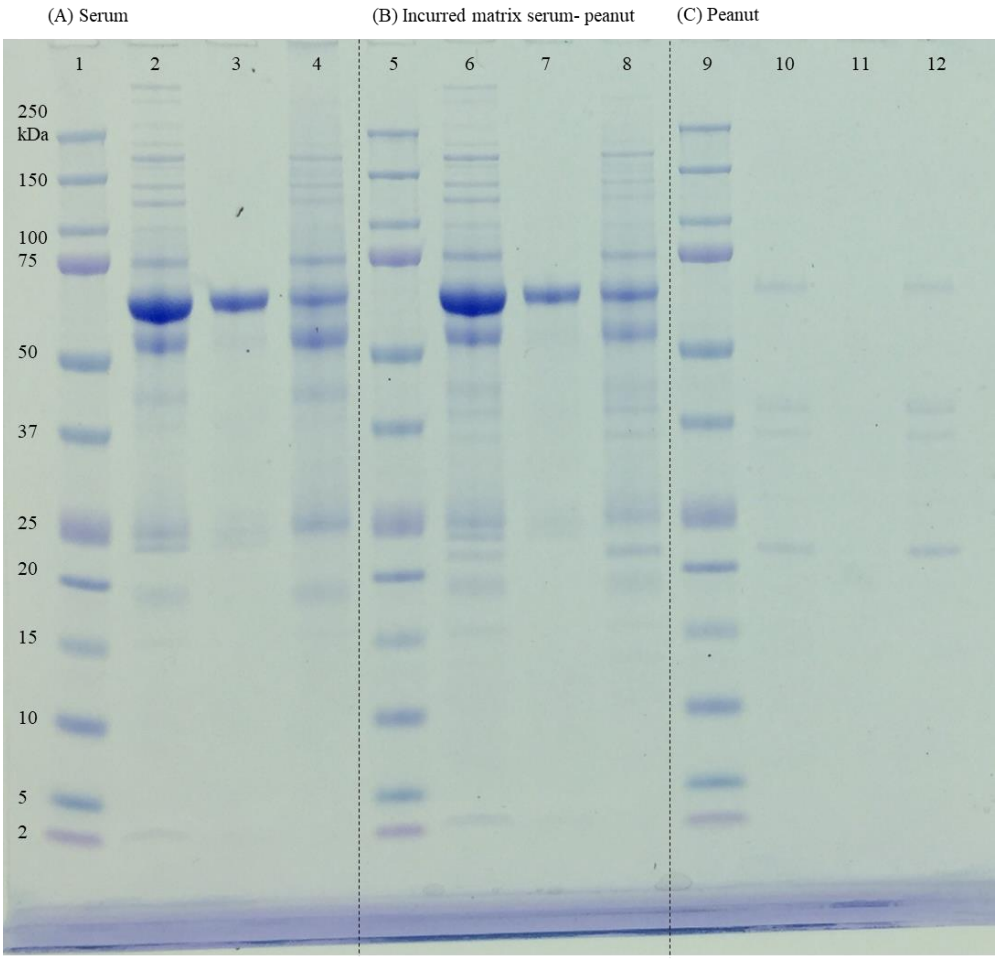


Figure 2-10. Evaluation of organic fractionation patterns by reducing SDS-PAGE of prepared matrices (A) serum, (B) incurred matrix serum-peanut, (C) raw peanut. Lanes (1, 5, 9) Protein marker standard (kDa), (2) unfractionated serum, (3) serum in isopropanol-TCA supernatant, (4) serum in pellet precipitate, (6) unfractionated serum-peanut, (7) serum-peanut in isopropanol-TCA supernatant, (8) serum-peanut in pellet precipitate, (10) unfractionated peanut, (11) peanut in isopropanol-TCA supernatant, (12) peanut in pellet precipitate.

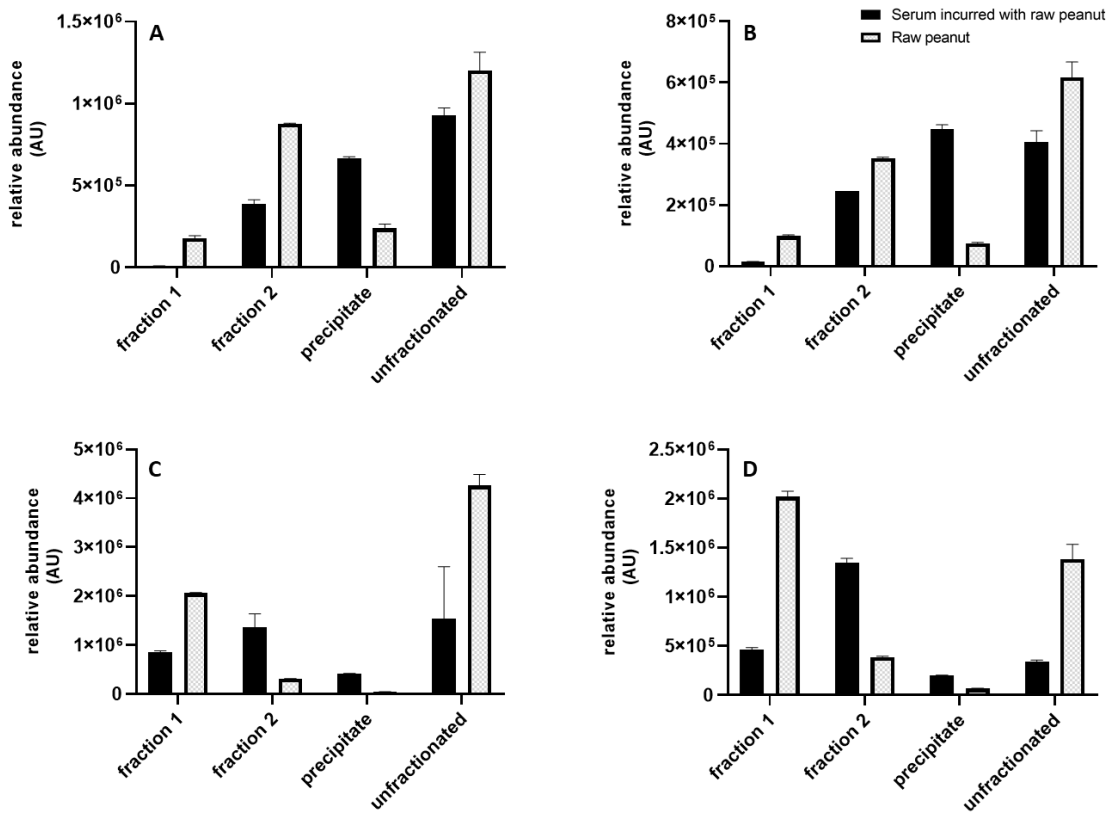


Figure 2-11. Evaluation of organic solvent fractionation by discovery LC-MS/MS in (A) Ara h 1 (Q6PSU3), (B) Ara h 3 (A1DZF0), (C) Ara h 2 (Q6PSU2-4), (D) Ara h 6 (Q647G9) in model samples serum-peanut or peanut.

VII. Evaluation of Multiple De-Complexation Methods (Organic Solvent Fractionation and IgG Depletion)

Our aim in combining two de-complexing strategies, IgG depletion and organic solvent fractionation, was to improve detection of peanut proteins in serum. We hypothesized the removal of IgG followed by organic solvent fractionation will greatly reduce the background and potential interferences.

For this set of experimental results, we will focus on the fractionation patterns of (1) the experimental control model matrix (sera-peanut) compared to (2) IgG depleted sera followed by the addition of raw peanut extract (IgG depleted sera with peanut spike), prepared with commercially prepared human serum (ImmunO). Based on our previous results, IgG-peanut complexes are not removed during IgG depletion. Therefore, we would expect only free IgG to be removed from our sera-peanut sample. To further assess the effect of IgG depletion on our model samples, we depleted serum of IgG followed by adding an equivalent volume peanut extract to the IgG depleted serum. Prepared samples were then fractionated using our organic solvent method and analyzed as before by discovery LC-MS/MS.

Comparison of the sera-peanut matrix before and after IgG depletion indicate no noticeable differences in fractionation patterns in an allergic individual's serum for all major allergenic peanut proteins monitored. (Figure 2-12).

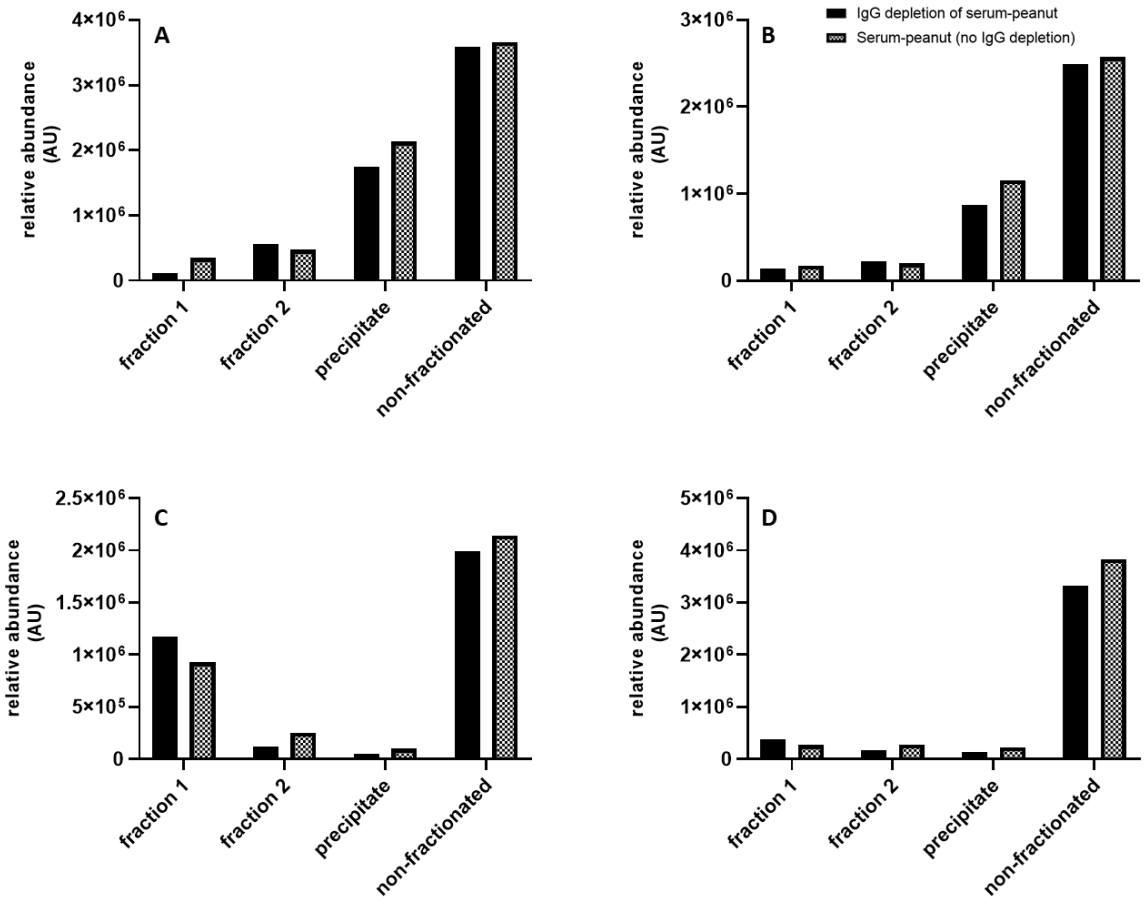


Figure 2-12. Evaluation of organic solvent fractionation of IgG depletion of serum-peanut matrix and no depletion of serum incurred with peanut by discovery LC-MS/MS

We observed differences in fractionation patterns based on the presence of absence of IgG. In the control sera-peanut samples with commercially prepared serum, the comparison of the relative abundance profiles of each peanut protein in each fraction is presented in Figure 2-13. In the sera-peanut matrix, Ara h 2 fractionates primarily in fraction 2, and is second most abundant in fraction 1. Ara h 6 was predominantly abundant in fraction 2, similar to Ara h 2, however, it was second most abundant in the precipitate. The cupin proteins (Ara h 1, Ara h 3), fractionated predominantly into fraction 2 and secondarily fractionated into the precipitate.

After IgG depletion and the addition of peanut extract to the depleted sera, we observed shifts in the fractionation patterns of peanut proteins (Figure 2-13). Ara h 2 fractionated largely into fraction 1, the isopropanol-TCA fraction, and then into fraction 2, the methanol fraction. This pattern differs from the sera-peanut matrix, where Ara h 2 abundance was highest in fraction 2. Ara h 6 was most abundant in fraction 2 and the precipitate. This differs slightly compared to the sera-peanut control which was predominantly present in fraction 2 and second most abundant in fraction 1, illustrating a shift in fractionation of Ara h 6 due to IgG (Figure 2-13D). For the cupin proteins, Ara h 1 was present largely in fraction 2 and the precipitate, exhibiting no shift in fractionation compared to the sera-peanut matrix. We observed Ara h 3 was present predominantly in fraction 2 and second most abundant in fraction 1. This pattern is slightly different than the sera-peanut pattern, in which case Ara h 3 was detected largely in fraction 2 and the precipitate (Figure 2-13B).

Overall, depletion of IgG from serum followed by the addition of peanut extract, demonstrates a fractionation pattern more similar to peanut (alone) and less similar to sera-peanut, the model matrix.

We then performed this same experiment of IgG depletion and organic solvent fractionation using allergic subject serum. Evidence of different IgG profiles between allergic and non-allergic individuals has been published in the literature (Aalberse et al., 1983; Platts-Mills et al., 2001; Rowntree et al., 1987).

We analyzed peanut allergic subject serum of (1) control, sera-peanut matrix and (2) IgG depleted sera followed by the addition of peanut (IgG depleted serum-peanut spike) (Figure 2-14). We compared the fractionation profiles of peanut allergens from the UniProt database. Data are represented by normalized abundance (e.g. the total identified peak area). The IgG depleted allergic sera with peanut extract addition, demonstrated a noticeably different fractionation pattern compared to the experimental control, sera-peanut.

In the sera-peanut controls, the prolamins, Ara h 2 and Ara h 6 exhibited identical fractionation patterns. Ara h 2 fractionated predominantly into fractions 1 and 2. Ara h 6 fractionated equally into fractions 1 and 2. Whereas Ara h 1 and Ara h 3 fractionated largely into the precipitate.

In the IgG depleted-peanut spiked samples, Ara h 2 fractionated largely into fraction 1 and secondarily into fraction 2. Ara h 6 fractionated equally into fraction 1 and fraction 2. Overall, we observed no differences in fractionation patterns of the 2S albumins dependent upon the presence or absence of IgG. Both, Ara h 1 and Ara h 3

fractionated largely into fraction 2. The fractionation pattern differs greatly from the sera-peanut control matrix.

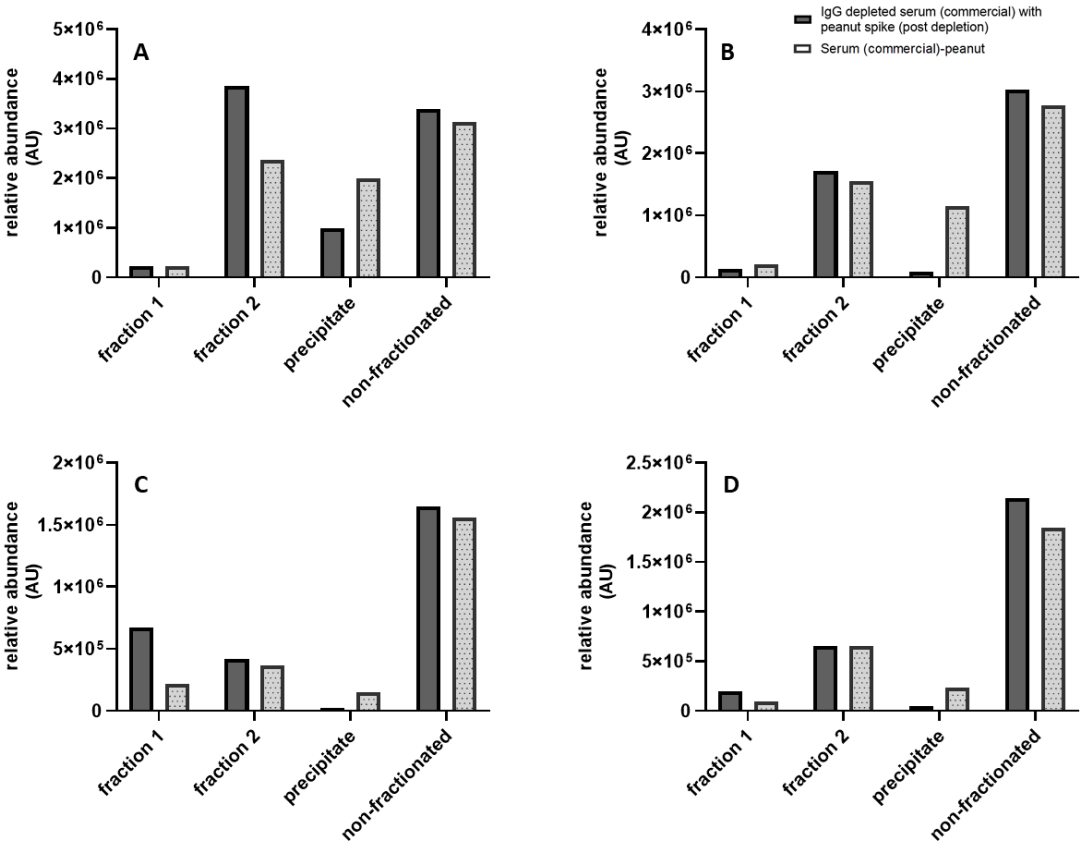


Figure 2-13. Evaluation of organic solvent fractionation of IgG depleted serum (commercial) with peanut spike (post depletion) and serum (commercial)-peanut extract by discovery LC-MS/MS in (A) Ara h 1 (P43237), (B) Ara h 3 (Q0GM57), (C) Ara h 2 (Q6PSU2), (D) Ara h 6 (A5Z1R0) in model samples serum-peanut or peanut.

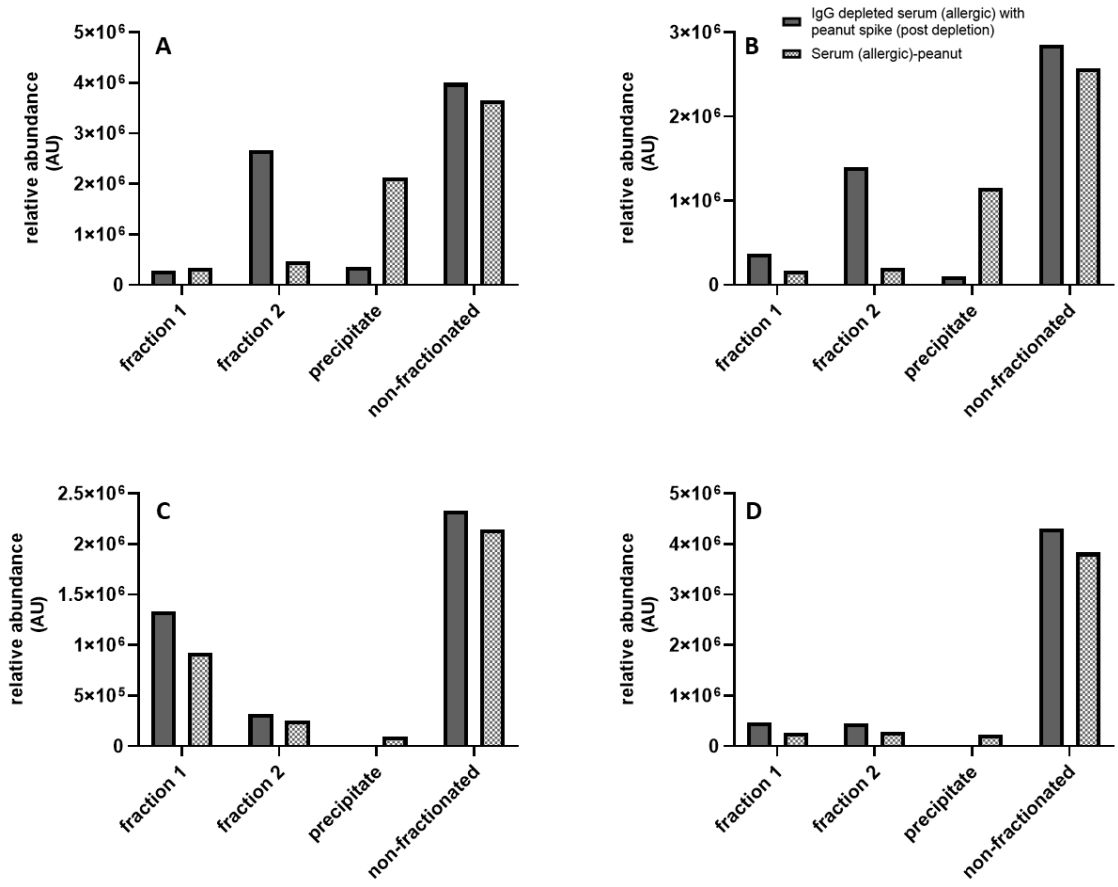


Figure 2-14. Evaluation of organic fractionation of IgG depleted serum (allergic) with addition of peanut extract (post depletion) and allergic serum(allergic)-peanut matrix by discovery LC-MS/MS in (A) Ara h 1 (P43237), (B) Ara h 3 (Q0GM57), (C) Ara h 2 (Q6PSU2), (D) Ara h 6 (A5Z1R0)

The fractionation patterns between allergic and commercially prepared serum, differ due to the presence or absence of IgG. The commercially prepared serum is a pool of serum, which has undergone filtration prior to purchase. The commercial serum is not guaranteed to be selected from non-allergic individuals.

We did not analyze baseline serum using this experiment, however upon review, this would be a suitable comparison for evaluation of these two de-complexation strategies. However, we can illustrate the differences in fractionation in allergic subjects compared to the commercial sera, indicating IgG may have a prominent role in response to allergenic food proteins (Chinthrajah et al., 2016; Koppelman et al., 2019; Vickery et al., 2013).

Initial evaluations of the sera-peanut matrix before and after IgG depletion showed no difference in fractionation patterns. We anticipate this is due to the functionality of the IgG depletion not able to remove IgG-peanut complexes and/or only removing free IgG. Therefore, we chose to not include the fractionation patterns of IgG depleted sera-peanut.

We analyzed several patient serum samples (allergic and non-allergic) incurred with peanut using TCA fractionation (data not shown), which indicated differences in fractionation patterns of peanut proteins in various sera samples. In order to understand differences in fractionation patterns and the role of IgG in individual sera, we would need to profile many patients, if one exists. Due to the individualistic variability of patient serum, and its influence on protein fractionation, we determined this method is unsuitable for monitoring low abundance peanut proteins in serum.

CONCLUSIONS

Discovery LC-MS/MS methods lack sufficient sensitivity to detect consumed allergenic peanut proteins at low concentrations *in vivo*. In these methods, the most abundant (e.g. top 10, top 20) precursor ions are selected and fragmented for MS/MS analysis in stochastic manner (Gillette and Carr, 2013). Consequently, protein abundance significantly influences data-dependent acquisitions and low abundance proteins may not be routinely sampled and remain undetected (Domon and Aebersold, 2010; Gillette and Carr, 2013). MS instruments often have necessary levels of sensitivity required for low detection, however, a high signal-to-noise ratio from a complex sample (e.g. serum) affects overall detection (Gillette and Carr, 2013).

We investigated multiple strategies to de-complex serum prior to LC-MS/MS analysis, including optimizing MS acquisition settings, as previously discussed, and physical serum de-complexation strategies. The use of exclusion and inclusion lists did not yield positive results for peanut detection. The commercial depletion kits successfully reduced sample complexity, however, were not successful for our experiment due to the co-depletion of exogenous peanut proteins in serum. IgG depletion was highly specific, however, provided little utility to achieving our overall aim. IgG complexes (IgG-peanut) are unable to be depleted and therefore, we are only able to deplete free, unbound IgG. One of the main limitations of commercial depletion columns is their limited binding capacity. The maximum load of protein for the Pierce™ Top2 Abundant Depletion spin columns is 600 µg, equivalent to 7.5 – 10 µl non-depleted serum. Working within this limitation and the known concentration of peanut after consumption (131 ng DRP-Ara h 2/ml serum), we would need to achieve detection of

13.1 ng Ara h 2 (262 ng total peanut protein) by discovery LC-MS/MS, assuming complete absorption and no experimental losses. Lastly, the organic solvent fractionation yielded variable results between sera types (commercial, peanut-allergic), and was further complicated by the addition of IgG depletion. Due to inter-subject variability in fractionation, these methods were not fit for the purposes of our study.

Each of the de-complexation strategies evaluated reduced the sample complexity and total protein by selective protein removal, improving the signal-to-noise ratios (S/N) and low abundance protein detection. Ideally, the employed depletion strategy is specific to serum proteins only, however this was not the case based on the results presented.

Based on our results, we suggest development of a targeted MS method for detection of peanut proteins in non-depleted serum. Due to the sampling limitations of discovery LC-MS/MS methods, targeted methods are a suitable alternative for achieving detection. Targeted methods, including MRM and PRM, have been successfully employed for specific analyte detection, overcoming shortcomings of shotgun methods. Targeted methods provide good sensitivity, reproducibility, and capable of quantification (Lange et al., 2008).

In supplement to targeted LC-MS/MS methods, we suggest using non-depleted serum for further investigations due to the co-isolation of peanut proteins by commercial depletion methods and variability of peanut protein fractionation by organic solvent methods. The use of non-depleted serum ensures peanut proteins remain in the subject serum sample. However, detection of low abundance proteins in serum remains challenging due to the depth and complexity of the human serum proteome, which will be addressed by implementing targeted methods.

REFERENCES

Aalberse, R.C., Mueller, G.A., Derksen, N.I.L., Aalberse, J.A., Edwards, L.L., Pomés, A., Lidholm, J., Rispens, T., and Briza, P. (2019). Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. Clin. Exp. Allergy cea.13554.

Adkins, J.N., Varnum, S.M., Auberry, K.J., Moore, R.J., Angell, N.H., Smith, R.D., Springer, D.L., and Pounds, J.G. (2002). Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. Mol. Cell. Proteomics *1*, 947–955.

Anderson, N., and Anderson, N. (2002). The Human Plasma Proteome History, Character, and Diagnostic Prospects. Mol. Cell. Proteomics *1*, 845–867.

Apostolovic, D., Stanic-Vucinic, D., De Jongh, H.H.J., De Jong, G.A.H., Mihailovic, J., Radosavljevic, J., Radibratovic, M., Nordlee, J.A., Baumert, J.L., Milcic, M., et al. (2016). Conformational stability of digestion-resistant peptides of peanut conglutins reveals the molecular basis of their allergenicity. Nat. Publ. Gr. *6*, 1–12.

Baumert, J.L., Bush, R.K., Levy, M.B., Koppelman, S.J., Nordlee, J.A., Hefle, S.L., and Taylor, S.L. (2009). Distribution of Intact Peanut Protein and Digestion-Resistant Ara h 2 Peptide in Human Serum and Saliva. J. Allergy Clin. Immunol. *123*, S268–S268.

Bendall, S.C., Hughes, C., Larry Campbell, J., Stewart, M.H., Pittock, P., Liu, S., Bonneil, E., Thibault, P., Bhatia, M., and Lajoie, G.A. (2008). An Enhanced Mass Spectrometry Approach Reveals Human Embryonic Stem Cell Growth Factors in Culture. Mol. Cell. Proteomics 8.3, 421–432.

Björck, L., and Kronvall, G. (1984). Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. J. Immunol. *133*, 969–974.

Blanc, F., Adel-Patient, K., Drumare, M.-F., Paty, E., Wal, J.-M., Bernard, H., and Bernard, H. (2009). Capacity of purified peanut allergens to induce degranulation in a functional in vitro assay: Ara h 2 and Ara h 6 are the most efficient elicitors. Clin. Exp. Allergy *39*, 1277–1285.

Chinthrajah, R.S., Hernandez, J.D., Boyd, S.D., Galli, S.J., Nadeau, K.C., and Stanford, À. (2016). Molecular and cellular mechanisms of food allergy and food tolerance. Mech. Allerg. Dis. *137*, 984–997.

Dirks, C.G., Pedersen, M.H., Platzer, M.H., Bindslev-Jensen, C., Skov, P.S., and Poulsen, L.K. (2005). Does absorption across the buccal mucosa explain early onset of food-induced allergic systemic reactions? J. Allergy Clin. Immunol. *115*, 1321–1323.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. Nat. Biotechnol. *28*, 710–721.

Eliuk, S., and Makarov, A. (2015). Evolution of Orbitrap Mass Spectrometry Instrumentation. Annu. Rev. Anal. Chem. *8*, 61–80.

Gianazza, E., and Arnaud, P. (1982). Chromatography of plasma proteins on immobilized Cibacron Blue F3-GA. Mechanism of the molecular interaction. Biochem. J. *203*, 637–641.

Gillet, L.C., Leitner, A., and Aebersold, R. (2016). Mass Spectrometry Applied to Bottom-Up Proteomics: Entering the High-Throughput Era for Hypothesis Testing. Annu. Rev. Anal. Chem. Is Online *9*, 449–472.

Gillette, M.A., and Carr, S.A. (2013). Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nat. Methods *10*, 28–34.

Gupta, R., Blumenstock, J., Warren, C., Mittal, K., Kotowska, J., and Smith, B. (2017). The prevalence of nut and seafood allergies among adults in the United States.

Gupta, R.S., Warren, C.M., Smith, B.M., Blumenstock, J.A., Jiang, J., Davis, M.M., and Nadeau, K.C. (2018). The public health impact of parent-reported childhood food allergies in the United States. Pediatrics *142*.

Gupta, R.S., Warren, C.M., Smith, B.M., Jiang, J., Blumenstock, J.A., Davis, M.M., Schleimer, R.P., and Nadeau, K.C. (2019). Prevalence and Severity of Food Allergies Among US Adults. JAMA Netw. Open *2*, e185630.

Heyman, M. (2005). Gut barrier dysfunction in food allergy. Eur. J. Gastroenterol. Hepatol. *17*, 1279–1285.

Hinderling, P.H., and Hartmann, D. (2005). The pH Dependency of the Binding of Drugs to Plasma Proteins in Man. Ther. Drug Monit. *27*, 71–85.

Husby, S., Jensenius, J.C., and Svehag, S.E. (1985). Passage of undegraded dietary antigen into the blood of healthy adults. Quantification, estimation of size distribution, and relation of uptake to levels of specific antibodies. Scand. J. Immunol. *22*, 83–92.

Jackson, K.D., Howie, L.J.D., and Akinbami, L.J. (2013). Trends in allergic conditions among children: United States, 1997-2011. NCHS Data Brief 1–8.

JanssenDuijghuijsen, L.M., Wichers, H.J., van Norren, K., Keijer, J., Baumert, J.L., de Jong, G.A., Witkamp, R.F., and Koppelman, S.J. (2017). Detection of peanut allergen in human blood after consumption of peanuts is skewed by endogenous immunoglobulins. J. Immunol. Methods *440*, 52–57.

Jutel, M., and Akdis, C.A. (2011). Immunological mechanisms of allergen-specific immunotherapy. Allergy Eur. J. Allergy Clin. Immunol. *66*, 725–732.

Klemans, R.J.B., Broekman, H.C.H.P., Knol, E.F., Bruijnzeel-Koomen, C.A.F.M., Otten, H.G., Pasmans, S.G.M.A., and Knulst, A.C. (2013). Ara h 2 is the best predictor for peanut allergy in adults. J. Allergy Clin. Immunol. Pract. *1*, 632-638.e1.

Kochansky, C.J., McMasters, D.R., Lu, P., Koeplinger, K.A., Kerr, H.H., Shou, M., and Korzekwa, K.R. (2008). Impact of pH on plasma protein binding in equilibrium dialysis. Mol. Pharm. *5*, 438–448.

Koelmel, J.P., Kroeger, N.M., Gill, E.L., Ulmer, C.Z., Bowden, J.A., Patterson, R.E.,

Yost, R.A., and Garrett, T.J. (2017). Expanding Lipidome Coverage Using LC-MS/MS Data-Dependent Acquisition with Automated Exclusion List Generation. J. Am. Soc. Mass Spectrom. *28*, 908–917.

Koppelman, S., and Hefle, S. (2006). Detecting allergens in food (Elsevier Ltd).

Koppelman, S.J., Peillon, A., Agbotounou, W., Sampson, H.A., and Martin, L. (2019). Epicutaneous immunotherapy for peanut allergy modifies IgG4 responses to major peanut allergens. J. Allergy Clin. Immunol. *143*, 1218-1221.e4.

Kotz, D., Simpson, C.R., and Sheikh, A. (2011). Incidence, prevalence, and trends of general practitioner-recorded diagnosis of peanut allergy in England, 2001 to 2005. J. Allergy Clin. Immunol. *127*, 623-630.e1.

Kulis, M., Chen, X., Lew, J., Wang, Q., Patel, O.P., Zhuang, Y., Murray, K.S., Duncan, M.W., Porterfield, H.S., Burks, A.W., et al. (2012). The 2S albumin allergens of Arachis hypogaea, Ara h 2 and Ara h 6, are the major elicitors of anaphylaxis and can effectively desensitize peanut-allergic mice. Clin. Exp. Allergy *42*, 326–336.

Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. Mol. Syst. Biol. *4*.

Leatherbarrow, R.J., and Dean, P.D. (1980). Studies on the mechanism of binding of serum albumins to immobilized cibacron blue F3G A. Biochem. J. *189*, 27–34.

Lieberman, J., Sublett, J., Ali, Y., Haselkorn, T., Damle, V., Chidambaram, A., Rosen, K., and Mahr, T. (2018). Increased incidence and prevalence of peanut allergy in children and adolescents in the United States. Ann. Allergy, Asthma Immunol. *121*, S13.

Liu, G., Zhao, Y., Angeles, A., Hamuro, L.L., Arnold, M.E., and Shen, J.X. (2014). A Novel and Cost Effective Method of Removing Excess Albumin from Plasma/Serum Samples and Its Impacts on LC-MS/MS Bioanalysis of Therapeutic Proteins. Anal. Chem. *86*, 8336–8343.

Michael, S. (1962). The isolation of albumin from blood serum or plasma by means of organic solvents. Biochem. J. *82*, 212–218.

Murphy, K. (2012). Janeway's immunobiology (Garland Science).

Oda, M., Kozono, H., Morii, H., and Azuma, T. (2003). Evidence of allosteric conformational changes in the antibody constant region upon antigen binding. Int. Immunol. *15*, 417–426.

Pieper, R., Su, Q., Gatlin, C.L., Huang, S.-T., Anderson, N.L., and Steiner, S. (2003). Multi-component immunoaffinity subtraction chromatography: An innovative step towards a comprehensive survey of the human plasma proteome. Proteomics *3*, 422–432.

Pisanu, S., Biosa, G., Carcangiu, L., Uzzau, S., and Pagnozzi, D. (2018). Comparative evaluation of seven commercial products for human serum enrichment/depletion by shotgun proteomics. Talanta *185*, 213–220.

Platts-Mills, T., Vaughan, J., Squillace, S., Woodfolk, J., and Sporik, R. (2001).

Sensitisation, asthma, and a modified Th2 response in children exposed to cat allergen: a population-based cross-sectional study. Lancet *357*, 752–756.

Reitsma, M., Westerhout, J., Wichers, H.J., Wortelboer, H.M., and Verhoeckx, K.C.M. (2014). Protein transport across the small intestine in food allergy. Mol. Nutr. Food Res *58*, 194–205.

Rinaldi, M., Harnack, L., Oberg, C., Schreiner, P., Sauver, J.S., and Travis, L.L. (2012). Peanut allergy diagnoses among children residing in Olmsted County, Minnesota. J. Allergy Clin. Immunol. *130*, 945–950.

Roche, S., Tiers, L., Provansal, M., Seveno, M., Piva, M.-T., Jouin, P., and Lehmann, S. (2009). Depletion of one, six, twelve or twenty major blood proteins before proteomic analysis: The more the better? J. Proteomics *72*, 945–951.

Rowntree, S., Platts-Mills, T.A.E., Cogswell, J.J., and Mitchell, E.B. (1987). A subclass IgG4-specific antigen-binding radioimmunoassay (RIA): Comparison between IgG and IgG4 antibodies to food and inhaled antigens in adult atopic dermatitis after desensitization treatment and during development of antibody responses in children. J. Allergy Clin. Immunol. *80*, 622–630.

Schocker, F., Baumert, J., Kull, S., Petersen, A., Becker, W.-M., and Jappe, U. (2016). Prospective investigation on the transfer of Ara h 2, the most potent peanut allergen, in human breast milk. Pediatr. Allergy Immunol. *27*, 348–355.

Schweitzer, B., Roberts, S., Grimwade, B., Shao, W., Wang, M., Fu, Q., Shu, Q., Laroche, I., Zhou, Z., Tchernev, V.T., et al. (2002). Multiplexed protein profiling on microarrays by rolling-circle amplification. Nat. Biotechnol. *20*, 359–365.

Sen, M., Kopper, R., Pons, L., Abraham, E.C., Burks, A.W., and Bannon, G.A. (2002). Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes. J. Immunol. *169*, 882–887.

Sicherer, S.H., and Sampson, H.A. (2007). Peanut allergy: emerging concepts and approaches for an apparent epidemic. J. Allergy Clin. Immunol. *120*, 491–503.

Sicherer, S., Muñoz-Furlong, A., Godbold, J., and Sampson, H. (2010). US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immunol. *125*, 1322–1326.

Suhr, M., Wicklein, D., Lepp, U., and Becker, W.-M. (2004). Isolation and characterization of natural Ara h 6: evidence for a further peanut allergen with putative clinical relevance based on resistance to pepsin digestion and heat. Mol. Nutr. Food Res. *48*, 390–399.

Thompson, S.T., Cass, K.H., and Stellwagen, E. (1975). Blue dextran sepharose: An affinity column for the dinucleotide fold in proteins. Proc. Natl. Acad. Sci. U. S. A. *72*, 669–672.

Travis, J., and Pannell, R. (1973). Selective removal of albumin from plasma by affinity chromatography. Clin. Chim. Acta *49*, 49–52.

Tsumoto, K., Ejima, D., Senczuk, A.M., Kita, Y., and Arakawa, T. (2007). Effects of salts on protein–surface interactions: applications for column chromatography. J. Pharm. Sci. *96*, 1677–1690.

Vickery, B.P., Lin, J., Kulis, M., Fu, Z., Steele, P.H., Jones, S.M., Scurlock, A.M., Gimenez, G., Bardina, L., Sampson, H.A., et al. (2013). Peanut oral immunotherapy modifies IgE and IgG4 responses to major peanut allergens. J. Allergy Clin. Immunol. *131*, 128-134.e3.

Wachholz, P.A., and Durham, S.R. (2004). Mechanisms of immunotherapy: IgG revisited. Curr. Opin. Allergy Clin. Immunol. *4*, 313–318.

Wood, R.A. (2003). The natural history of food allergy. Pediatrics *111*, 1631–1637.

Zhuang, Y., and Dreskin, S.C. (2013). Redefining the major peanut allergens. Immunol. Res. *55*, 125–134.

CHAPTER 3: DEVELOPMENT OF TARGETED MASS SPECTROMETRY METHODS FOR *IN VIVO* DETECTION OF PEANUT PROTEINS IN HUMAN SERUM

ABSTRACT

Due to the complexity of human serum and the substantial number of peptides generated after trypsin digestion, discovery MS is not ideal. Therefore, targeted methods (PRM, MRM), will be used instead. The aim of this chapter was to develop a targeted MS method to detect allergenic peanut proteins in non-depleted serum. To maintain sample integrity, non-depleted serum will be analyzed to mitigate any losses of critically relevant protein targets.

The model samples described in Chapter 2 (page 79) were used for targeted LC-MS/MS method development. Peanut protein targets, representing the major allergenic proteins, were selected using discovery driven peptide selection. Targeted LC-MS/MS methods, PRM, MRM, and MRM^3, were developed sequentially based on the required level of sensitivity using our model samples (sera-peanut, sera, peanut). Sample preparation, method chromatography, MS acquisition settings were further modified to improve detection in subject sera.

Peptide targets were selected using discovery MS for the development of PRM and MRM methods. We evaluated a sera-peanut matrix and observed similar LODs using both PRM and MRM methods. The LOD for PRM in the sera-peanut matrix was 800 pg peanut protein (1.0 ppm peanut protein, 4.0 ppm peanut). The MRM LOD was peptide dependent. For Ara h 1, Ara h 2, and Ara h 6 the LOD was 1240 pg peanut (1.53

ppm peanut protein, 6.11 ppm peanut). For Ara h 3, the LOD was 3730 pg peanut protein (4.58 ppm peanut protein, 18.3 ppm peanut). The LOD for the MRM^3 method was variable and inconsistent due to the lack of specific secondary fragment ions for detection.

We further optimized the PRM method by injecting more protein for analysis. Overall, the PRM methods were the optimal methods to use, and were able to detect peanut peptides in serum on multiple occasions. However, due to a lack of robust detection, these methods are not yet optimal. Further work is required to develop a robust targeted detection method for *in vivo* detection of allergenic proteins.

The targeted methods were successful in detecting peanut proteins in the model matrices. Evaluation of subject sera samples determined positive by Ara h 6 ELISA, were evaluated using each of the targeted MS methods, but no reproducible positive detections were observed. However, we were able to confidently detect multiple peptides in several instances from different subject sera and at different time points, suggesting our methodology is a probable candidate for routine evaluation.

INTRODUCTION

The prevalence of food allergies has been increasing in Western countries (Nwaru et al., 2014; Sicherer et al., 2010). Consequently, there has been a steady increase of peanut allergy prevalence, the most prevalent food allergy being in children (Nwaru et al., 2014; Sicherer et al., 2010; Venter et al., 2010). Peanuts are highly potent allergens, and can elicit reactions with ingestion of low doses in sensitized individuals (Clarke et al., 1998; Hourihane et al., 2017; Koppelman et al., 2004). There is further evidence

these ingested peanut proteins are present in serum, breast milk, and saliva, however overall detection was very low (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017b; Schocker et al., 2016). The method of protein uptake, particularly intact proteins, has been described in several reviews but has not been fully elucidated (Chehade and Mayer, 2005; Reitsma et al., 2014;). The role of allergenic protein absorption likely influences sensitization mechanisms. In order for an allergic reaction to occur, the allergenic food proteins must cross-link IgE specific antibodies in an immunologically intact form (Stone et al., 2010; Taylor and Baumert, 2012;). It has been suggested the transport and absorptive mechanisms employed differs between allergic (atopic) and non-allergic (healthy, non-atopic) subjects (Reitsma et al., 2014). In sensitized individuals, there is a decrease is the tight junction barrier function due to the presence of mast cells, which consequently increases the transport of non-degraded or partially degraded allergenic protein (Berin et al., 1998). Therefore, much focus has been placed on development of an *in vivo* detection method to gain understanding in uptake and sensitization.

Several studies have focused on detection of peanut proteins *in vivo*, primarily by ELISA methods and more recently by LC-MS/MS (Baumert et al., 2009; Hands et al., 2020; JanssenDuijghuijsen et al., 2017b). ELISAs have been the primary methods of choice for studying *in vivo* uptake, however these methods need additional confirmatory methods to demonstrate antibody-binding specificity (Hands et al., 2020).

LC-MS/MS methods, which are used as an alternative to ELISAs, remain challenging due to complexity of human serum. Therefore, the comprehensive elucidation of low-abundance proteins by LC-MS/MS remains challenging to do the broad dynamic range and few proteins accounting for a majority of abundance (Kumar

Dey et al., 2019). There is limited literature published regarding *in vivo* detection of allergenic food proteins, further, a majority of these studies utilize ELISA methods (Baumert et al., 2009; JanssenDuijghuijsen et al., 2017a; JanssenDuijghuijsen et al., 2017b; Schocker et al., 2016). Recently, two studies have employed the use of LC-MS/MS methods to monitor the uptake of peanut proteins (Hands et al., 2020; Mose et al., 2019). Hands et al., (2019) developed an MRM detection method for peanut proteins spiked into commercial human serum. Additionally, Hands et al., (2020) used depletion (MARS Hu-6) to remove majorly abundant serum proteins prior to MRM analysis. However, in this study, no subject serum collected after peanut consumption were analyzed using the developed method. Mose et al., 2019 evaluated multiple MS methods (MRM, GeLC MS/MS using discovery MS, and SWATH) with spiked model matrix (serum with Ara h 2) and subject serum collected after peanut consumptions. They reported positive detection using MS methods for the spiked matrix samples (serum-Ara h 2), but were unable to achieve detection with subject samples collected after consumption (Mose et al., 2019). However, subject sera demonstrated reactivity through basophil-histamine release assays of subject serum after ingestion (Mose et al., 2019). This is suggestive that peanut proteins are present *in vivo* (i.e. serum) in an immunologically reactive form, but not readily detectable by the bottom-up LC-MS/MS methods.

Due to the ineffectiveness of depletion kits, as presented in Chapter 2, investigators have adopted methods utilizing non-depleted serum and targeted MS methods (Fortin et al., 2009; Kumar Dey et al., 2019; Percy et al., 2013). Fortin et al., 2009 demonstrated detection of protein biomarkers at ppm concentration levels in non-

depleted serum using MRM^3. Kumar Dey et al., 2019, investigated profiling non-depleted serum to aid in determining biomarker targets for Alzheimer's disease using an LC-MS3 (TOMAHAQ) method. Targeted MS methods are beneficial for their sensitivity to pre-selected peptides, improving an instrument's dynamic range, reducing interfering background noise, and ability to easily multiplex (Picotti and Aebersold, 2012; Ronsein et al., 2015). Targeted MS methods are able to detect individual allergenic protein sequences provided within a database. For targeted methods, it is critical to have a robust peptide selection method, especially in complex matrices to prevent false positive identifications (Sherman et al., 2009).

In our investigations, we utilized three different targeted MS methods to increase sensitivity, for detection of allergenic proteins *in vivo.* PRM based mass spectrometry utilizes high resolution-accurate mass (HR-AM) instruments, and is highly specific since all product fragment ions are analyzed in parallel. The HR-AM analyzer enables in PRM scans the capability to distinguish between co-isolated ions, by determining each ion's fragmentation patterns (Duncan et al., 2009; Peterson et al., 2012; Sherman et al., 2009). MRM methods have been the gold-standard quantification method due to their high selectivity, sensitivity, reproducibility, and quantifiable accuracy for peptide detection (Addona et al., 2009; Lange et al., 2008). MRM^3 methods provide an additional level of sensitivity, and are able to quantify low abundance peptides in non-depleted serum (Fortin et al., 2009).

The objective of this chapter is to develop a series of targeted MS methods, in order of increasing sensitivity, PRM, MRM, and MRM^3, for unique peanut peptide targets selected by DDA selection process. We anticipate needing high levels of

sensitivity with corresponding low levels of detection based on the currently available *in vivo* protein absorption data. In this chapter, subject serum collected by JanssenDuijghuijsen et al., 2017b, which was determined positive by an Ara h 6 ELISA, were evaluated by the targeted methods developed. Ara h 6, a preferential immunoassay target due to its digestive stability, was detected at extremely low concentrations (0.3 ng/mL Ara h 6) in human serum (JanssenDuijghuijsen et al., 2017b). Ara h 6 only accounts for ~6% of the total peanut protein; therefore the estimated protein concentration of peanut in serum is expected to be greater (6 ng peanut protein/mL) (JanssenDuijghuijsen et al., 2017b; Koppelman et al., 2016). Given the unknown concentrations of peanut proteins in serum, the level of detection required by MS methods is largely unknown. We developed the targeted methods in the following order, PRM, MRM, and MRM^3, with each method ideally providing improvements to sensitivity.

MATERIALS AND METHODS

I. Reagents

All reagents used were of analytical grade for all experiments. All reagents used for LC-MS/MS sample preparations and analyses were of MS grade.

II. Discovery Identification of Peanut Peptide Targets

For initial data-dependent acquisitions, samples were prepared on an equal protein basis. All samples prepared will have the same amount protein present in the sera-peanut

matrix sample as is present in their individual component samples (peanut or serum alone); non-depleted samples were prepared as described in Chapter 2 (page 79).

III. Sample Preparation and In-Solution Reduction, Alkylation, and Trypsin Digestion for Discovery LC-MS/MS

Initial identifications of unique peanut peptides were determined experimentally by discovery LC-MS/MS of (1) non-depleted serum, (2) raw peanut extract, and (3) an incurred matrix of non-depleted serum and raw peanut extract (10:1 (w/w)), as described previously in Chapter 2 (page 79).

However, for model matrices prepared in this chapter, we used commercially available human serum (human serum sterile, ImmunO, MP Biomedicals) and raw peanut extract (0.01 M PBS), combined at a 10:1 w/w ratio, with a maximum of 30 µg protein for trypsin digestion. This incurred serum-peanut matrix will serve as our model system for all subsequent stages of MS method development, unless otherwise noted. Control samples were prepared using equivalently.

Samples were evaluated by discovery LC-MS/MS using a standard in-solution reduction, alkylation, and trypsin digestion protocol modified from Thermo™ In-Solution Trypsin Digestion and Guanidination, as previously described (page 80).

A. Discovery LC-MS/MS Acquisition Method

We used the optimized discovery acquisition developed in Chapter 2, using the same instrument platform and equipment (page 84). All instrument parameters remained

the same, except, the UPLC column was equipped with a Javelin™ Direct-Connection Column Filter, 2.1 mm (Thermo Scientific™), a Hypersil Gold aQ C18 1.9 μm, 20 × 2.1 mm pre-column (Thermo Scientific™) and a Hypersil Gold C18 1.9 μm, 100 × 1 mm analytical reversed phase column (Thermo Scientific™).

B. Data Analysis

Acquired discovery data were evaluated using a peanut database (Arachis hypogaea, taxon identifier 3818) downloaded from UniProt (8/31/2016) and PEAKS version 8.5 software (Bioinformatics Solutions; Cheriton et al., 2019) for identification of unique proteotypic peanut peptides, as previously described (page 82), however the parent mass tolerance was set to 5 ppm. Label-free quantification was performed to quantify peptide abundance and data were normalized to TIC.

C. Peptide Selection

Peptides identified using discovery MS were selected for evaluation as candidate PRM targets. The following criteria were used to evaluate peptides: (1) selected peptide sequences must be unique and attributable to peanut protein(s); (2) selected peptides must be absent from sera only data; (3) candidate peptides must originate from a multiply charged precursor ion (2+, 3+, or 4+); (4) contain 5 – 20 amino acid residues; (6) have tryptic cleavage sites; (7) no miscleavage events; and (8) absent of any post-translational modifications (PTMs) other than carbamidomethylation, a product of alkylation (Korte et al., 2016; Rauniyar, 2015). Peptides must be present in relatively high abundances in

both the peanut and serum-peanut matrices. If peptides were detected at multiple charge states both charge states were included in preliminary target selection. Once all candidate peptides were identified by the above listed criteria, the inclusion list was further refined. Of the peptides identified, only those detected within the top 10% of all identified candidate peptides by signal abundance were included for further evaluation in PRM method development.

IV. PRM Method Development

The PRM method was developed by evaluating peptides selected by discovery LC-MS/MS. Due to the large number of peptides included, we divided the peptides into two inclusion lists based on retention times; target list A and target list B. Peptides were ranked according to retention times, and split between the two lists, by selecting every other peptide based on retention time.

A. MS Method Settings for PRM Evaluation

Once peptides were designated to the appropriate inclusion list, we evaluated peptide performance against the same set of samples, sera-peanut matrix, peanut, and serum. Samples were prepared identically as those previously described using the same instrument platform and chromatography in the discovery LC-MS/MS experiment. In total, 52 peptides from the four major allergenic peanut proteins (Ara h 1, Ara h 2, Ara h 3, Ara h 6) were identified (described in detail in results section, page 171). These peptides divided into two inclusion lists, with a maximum of 27 and 26 peptides in each

list, respectively. Peptides in the initial inclusion list(s) were not scheduled. Scheduling reduces background interferences by only scanning for specific m/z events during a defined time window (e.g. peptide retention time). LC method run time was 113 minutes, with PRM scan time of 73 minutes (2 – 75 minutes). MS acquisition settings were set as follows: MS2 resolution was set to 140,000; AGC target 1e6; maximum IT 500 ms; loop count 30; isolation window 1.6 m/z with an isolation offset set to 0.0 m/z; NCE set to 27. Instrument gas and source settings for PRM acquisitions were identical to those previously described (page 81).

B. PRM Optimization and Evaluation of Selected Target Peptides

The PRM peptide list was refined by adjusting the chromatographic gradient to 5 – 50% mobile phase B (0.1% formic acid in acetonitrile). All instrument and gas settings were kept the same as previously described (page 151).

The same set of prepared matrix samples was injected using PRM acquisition. The target list was reduced to 19 target peptides after the last series of peptide evaluations. Peptides were then combined into a single inclusion list, with the loop count set to 20. We included peptide retention times with +/- 1.5 minutes of the mean calculated peptide retention time.

The method sensitivity (limit of detection) was evaluated by preparing serum incurred with serially diluted peanut protein. The volume of serum used in each model sample remained constant while the peanut extract was serially diluted. The concentrations of each peanut extract are described in Table 3-1. An equivalent volume

from each peanut extract was added to sera, to maintain equal sample volume(s). Samples were prepared for LC-MS/MS by digestion, de-salting (C18), and re-suspension, as previously described for model samples (page 79). Samples were injected in duplicate. We evaluated both (1) sera-peanut matrices and (2) peanut, without the matrix background, at decreasing concentrations of peanut protein.

ratio of sera:peanut	ng peanut protein injected	ppm peanut protein in original sample
10:1	80	110
100:1	8	11.0
500:1	1.6	2.20
1,000:1	1.1	1.10
2,000:1	0.40	0.56
5,000:1	0.16	0.22
10,000:1	0.08	0.11

Table 3-1. Dilution series of peanut in serum for evaluation of PRM method sensitivity.

C. Data Analysis of PRM Methodology

i. Peptide Selection

PRM data were analyzed using Skyline v20 software (Maclean et al., 2010). Parent ions were selected by the PRM method and fragmented, therefore we evaluated the transitions of each parent (MS1) ions. Peptides with six or more detectable transitions were retained for further evaluation as target peptides. We evaluated y- or b-ions only and removed other ion products from analyses.

The settings used to analyze PRM data in Skyline were set to match the acquisitions method parameters with an ion match tolerance set to 0.7 m/z, and a method match tolerance set to 0.055 m/z

Transitions, y1, y2, b1, and b2, were removed, as were, transitions which were not detected. Transitions which were positive in sera (alone) were removed. The transition ratio pattern in sera-peanut samples must be identical to peanut (alone), with similar recorded retention times. Peptides were ranked by summed peptide peak area. The top 5

peptides, for each major allergenic peanut protein, which meet the above described criteria were included further evaluation.

ii. Peptide Selection and Refinement

Peptides were evaluated using Skyline, as previously described in the previous section (page 154). The retention time stability was evaluated and the scanning windows were adjusted for their mean observed retention time, if vastly different from previous analyses. Peptides must have three robust transitions, which meet the above criteria in order to remain included as target candidate peptide. The dilution series was evaluated for robust peptide detection at each level. The LOD was assessed in a qualitative manner by observing lack of transition detection, shifts in retention time, and poor peak shape.

V. MRM

A. Instrument Settings and Method Optimization

Since we anticipated needing high levels of sensitivity and low-level peptide detection, we developed an MRM method using the optimized targets and chromatographic settings determined during PRM method development.

MRM analysis was performed on a SCIEX QTRAP 6500+, a triple quadrupole mass spectrometer coupled with a Shimadzu Nexera II UHPLC (binary pump) with a Hypersil Gold uHPLC column (100 x 2.1 mm, 1.9 µm) equipped with a filter (Javelin™ Direct-Connection Column Filter, 2.1 mm (Thermo Scientific™). The PRM chromatographic method was directly transferred. Peptide scheduling was re-evaluated on the QTRAP instrument platform. Mobile phases were as follows: A was 0.1% formic acid in water, B was 0.1% formic acid in acetonitrile. To accommodate the larger column, the flow rate was increased to 300 µL/min and column oven temperature was set to 40°C. Low mass (<1000 m/z) acquisition setting was used.

The following instrument settings were optimized for MRM peptide detection in order they are listed: vertical probe position set to 8, horizontal probe position set to 6, curtain gas (CUR) set to 30, ion spray voltage (IS) of 5500, temperature (TEM) 300°C, ion source gas 1 (GS1) at 40, ion source gas (GS2) at 50, and collision gas (CAD) set to high.

For optimization, we infused mobile phase A (0.1% formic acid in water) buffer and observed the responses for each gas setting. Once all acquisition settings were optimized, a peanut sample was injected for validation.

B. MRM Peptide Target Selection and Evaluation

To evaluate selected peptide targets, we injected (2 µL) the model samples, sera, peanut, and sera-peanut, (0.5 µg/µL) as performed in previous experiments. Initial peptide targets determined by PRM were included for evaluation as MRM targets. We included the top six transitions for each selected peptide for evaluation by MRM.

The peptide dependent instrument settings, collision energy (CE), dwell time, and declustering potential (DP) were optimized to maximize sensitivity for the MRM method. Peptide dependent settings are generally determined by direct infusion of synthesized isotopically labeled peptide(s). However, the instrument settings were optimized by a direct infusion of tryptically digested raw peanut extract, prepared in the same manner as previously described to a final peptide concentration of (0.5 µg/µL), similarly to von Bargen et al., 2013.

Peptide retention times were determined on the QTRAP platform using an Enhanced MS scan. Replicate injections (2 µL) of digested peanut extract (0.5 µg/µL) were performed to determine mean retention time. The mean retention time ± 45 seconds were included.

C. MRM Method Sensitivity Determination

We similarly analyzed the same two prepared dilution series, (1) sera-peanut and (2) peanut, to determine MRM method LOD (Table 3-2). MRM data were analyzed using Skyline software following the same criteria described for PRM analyses (Maclean et al., 2010).

The settings used to analyze MRM data in Skyline were set to match the acquisition method parameters with an ion match tolerance set to 0.5 m/z, and a method match tolerance set to 0.6 m/z. The LOD was assessed in a qualitative manner by observing lack of transition detection, shifts in retention time, and poor peak shape.

ratio of sera:peanut	ng peanut protein injected	ppm peanut protein in original sample
10:1	93.3	115
50:1	18.7	22.9
250:1	3.7	4.58
750:1	1.2	1.53
1,250:1	0.75	0.92
3,750:1	0.25	0.31
6,250:1	0.15	0.18
31,250:1	0.03	0.04

Table 3-2. Dilution series of peanut in serum for evaluation of MRM and MRM^3 method sensitivity.

VI. MRM^3

A. Instrument Setting Optimizations

MRM^3/MS3 ions were determined using direct infusion of a tryptically digested peanut extract prepared as described previously (page 80). A total amount of 30 µg peanut protein were digested, and re-suspended to a concentration of 0.5 µg/µL prior to infusion. The digested peanut sample was infused at a rate of 20 µL/min. This flow rate provided suitable detection of all MRM transitions.

Selection of MRM^3 product ions, or tertiary fragment ions, was determined by evaluating parent ions (MS1) used in MRM experiments. The determination of MRM^3 transitions is an iterative process and must be determined experimentally using the MRM^3 Optimization Script within the instrument software, Analyst.

For each individual peptide, the parent (MS1) ion m/z was input into the MS3 Optimization Script. MS1 ion selection settings were set with an ion mass tolerance of ±2 Da, a scan rate of 1,000 Da/sec, and the MS1 must not be below 10% of the total ion current, using an Enhanced Resolution (ER) scan. DP is determined during the Q1 Multiple Ion scan using a ramp from 0 – 100 V, with increment measurements of 5 V. The Enhanced Product Ion (EPI) scans for MS2 selection recorded the top-5 MS2 product ions, at a scan rate of 1,000 Da/sec, mass range of 225 – 1,000 m/z, CE set at 40, with a collisional energy spread (CES) set to 25. MS/MS/MS (MS3) scan rate was set to 1,000 Da/sec, with Q0 trapping, a fixed fill time of 100 ms, and a mass range of 100 to 1000 m/z. The optimized MS2 excitation energy (AF2) are determined within the MS3 Script. All stages of ion selection are performed iteratively with continuous infusion of the peanut sample (20 µL/min).

Identification of secondary product ions (MS3) was difficult for many peptides. In order to have multiple peptide targets for MRM^3, we investigated using non-fragmented MS2 peptide ions as MS3 targets. To do so, the MS2 m/z value was input for both the MS2 and MS3 m/z's.

B. MRM^3 Cycle Time Optimization

During an MRM^3 method, scheduling is not possible within the software. To circumvent this problem, we divided our experiment into eight periods, wherein scans for each respective target occurred within a defined period. Without scheduling, the method has high cycle times, greatly impacting sensitivity.

C. MRM^3 Peptide Target Selection

MS3 fragmentation spectra of tertiary product ions will be recorded and assessed manually. For inclusion as an MRM^3 target, the ions must have good signal intensity, y- or b-ion(s) originating from MS2 fragment ion sequence, and absent from non-specific MS3 fragmentation events (e.g. loss of NH_3 or H_2O ions) (Korte and Brockmeyer, 2016). The LOD was assessed in a qualitative manner by observing lack of transition detection, shifts in retention time, and poor peak shape.

D. MRM^3 Method Evaluation

The selected tertiary product ions were evaluated for uniqueness against our model analytical samples, sera, peanut, and sera-peanut. The LOD for MRM^3 was performed as previously described using two dilution series of (a) peanut and (b) sera-peanut, according to Table 3-2. The chromatographic separation of peptides is unaffected by the implementation of a second stage MS fragmentation, and no changes were made to the chromatographic settings. Peptide scheduling will be added for MRM^3 acquisitions to achieve decreased cycle times. MRM^3 data were analyzed by Analyst software. Skyline software is incompatible for analysis of MRM^3 data files, due to the addition of an ion trap scan for MRM^3 detection.

VII. Evaluation of Subject Serum Samples by MRM and MRM^3

We evaluated subject samples previously determined positive by Ara h 6 specific sandwich ELISA (JanssenDuijghuijsen et al., 2017b) using our developed MRM and MRM^3 methods. We evaluated two subject sera at baseline (0 hour) and their highest measured two selected time points, 60 and 120 minutes. Complete ELISA results and subject characteristics are described in the original publication. A positive control (blank subject sera-peanut 100:1 (w/w)), was prepared. Following each analytical sample, a blank sample consisting of re-suspension solution was injected. We performed this experiment twice, on two separate occasions. The subject serum was collected from healthy individuals who regularly consume peanuts and are not allergic to peanuts (JanssenDuijghuijsen et al., 2017b).

VIII. Optimization of High-Resolution PRM Acquisitions

Following investigations of multiple targeted methods, PRM, MRM, and MRM^3, we re-investigated the use of high-resolution PRM methods. Due to the specificity required for detection of low abundance serum-proteins, we anticipate the need for high-resolution acquisitions. Confirmation using MRM and MRM^3 methods was unconvincing, even though peanut specific transitions were monitored. Likely, the potential number of interfering serum ions factored into this lack of detection. We have approached this section in the same manner as previously described targeted methods. Our aim was to analyze the maximum level of protein for detection of peanut proteins in sera using the previously developed PRM method. We calculated the amount of serum

necessary to achieve detection, based on the calibration curves from initial PRM studies and estimated concentrations of peanut in serum from JanssenDuijghuijsen et al., 2017b.

Serum used in this set of experiments was collected and analyzed using ELISA by JanssenDuijghuijsen et al., 2017b. Blood was collected from patients into serum separator tubes at the define time points following ingestion (30, 60, 120, 240, 360 minutes). Collected blood was allowed to clot (room temperature, 20 minutes, in the dark), followed by centrifugation (2000 x g, 10 minutes, room temperature), ultimately producing serum. Serum was stored at -80°C (JanssenDuijghuijsen et al., 2017b).

A. Evaluation of Large Scale Digestion Methodology

We prepared a larger volume (45 µL) of our model matrix, non-depleted commercial serum (MP Biomedicals) with peanut extract (50:1 (w/w)) for LC-MS/MS analysis. Commercial serum was initially used to evaluate the optimal amount of digested protein to be injected and separated on the liquid chromatography column.

To accommodate the increased amount of digested protein, we modified the previously developed in-solution trypsin digestion procedures from Chapter 2 (page 80). Commercial and subject (baseline) serum (80 µg/µL) were diluted (45 µL) to 10 µg/µL with acetonitrile and water to achieve a final concentration of 5% acetonitrile at the end of digestion. Samples were mixed with ammonium bicarbonate (50 mM) and reduced by adding dithiothreitol (100 mM), followed by heating at 95°C for either, (A) 5 minutes or (B) 20 minutes. Alkylation was performed with iodoacetamide (100 mM) at room temperature in the dark for 20 minutes. Trypsin (Pierce™ Trypsin Protease, MS Grade,

90057) was added at two enzyme:protein ratios, either 1:50 or 1:100 (w/w) enzyme:protein. Trypsin (1 µg/µL in 5 mM acetic acid) was added to reduced and alkylated protein samples, and allowed to digest for 3 hours at 37°C. A second addition of trypsin was added, resulting in a final enzyme:protein ratio of 1:50 or 1:100 (w/w), respectively, and incubated overnight at 30°C. The digestion was stopped by freezing samples. The remaining tryptic digests were de-salted using Strata-X 33 µm polymeric reversed phase 10 mg/1 mL columns (Phenomenex, Aschaffenburg, Germany). Columns were conditioned with 100 % methanol and equilibrated with 5% acetonitrile, 0.1% formic acid. Digested peptides were loaded into equilibrated columns and washed twice with 5% acetonitrile and 0.1% formic acid. Peptides were eluted from columns by 70% methanol, 0.1% formic acid (500 µL). The elution step was repeated once. De-salted peptides were dried under vacuum by centrifugal evaporation Jouan RC-10.10; RCT-90; Winchester, VA, USA. Peptides were re-suspended to 80 µg/µL in 5% acetonitrile, 0.1% formic acid. Modified digestion and reduction procedures were evaluated by SDS-PAGE.

B. SDS-PAGE Evaluation of Digestion Methodology

Samples for SDS-PAGE were taken at various time points, 0, 3, 18, and 24 hours, and after de-salting and re-suspension procedures. SDS-PAGE was conducted under reducing conditions using NuPAGE Bis-Tris Mini Gels 4-12% (1.0 mm, 12 wells) and constant voltage (200V) for 40 minutes in an XCell SureLock Mini Cell Electrophoresis System (Invitrogen Life Technologies). Non-digested serum (e.g. control) was diluted 22-fold prior to SDS-PAGE sample preparation, to a concentration of 3.6 µg/µL.

Aliquots taken during digestion, diluted control serum (10 µL), and re-suspended peptide digests after de-salting (12 µL) were combined with 4 x concentrated Laemmli buffer and 1% β-mercapto-ethanol (BME). Due to peptide re-suspension following de-salting and solvent evaporation, different volumes were required to maintain equal protein loading of each sample. Reduction occurred by heating all samples for 5 minutes at 95°C. Samples (10 µL) were loaded into each gel well. Precision Plus Protein Dual Xtra Standards (Bio-Rad) were used as the molecular weight (MW) standard. Gels were stained, overnight, in Coomassie Brilliant Blue R-250 (Bio-Rad) stain. Gels were then de-stained (Coomassie Brilliant Blue R-250 Destaining solution, Bio-Rad), rehydrated, and imaged.

C. Evaluation of Protein Loading by PRM

The following section describes the optimized in-solution trypsin digestion procedure from our previous evaluations. Commercial serum (80 µg/µL) was combined with raw peanut extract (50:1 (w/w)). The matrix samples, sera-peanut (45 µL), were diluted to 10 µg/µL with acetonitrile:water achieving a final concentration of 5% acetonitrile at the end of digestion. Samples were digested and de-salted as previously described (page 163) using a longer incubation time for reduction (95°C for 20 minutes) and a 1:100 (w/w) enzyme:protein ratio. A third aliquot of trypsin, at equal volume to previous trypsin aliquots, was added following overnight digestion to insure complete protein digestion, resulting in a final 1:67 final enzyme:protein ratio. Samples were incubated at 30°C for 3 hours following the third addition of trypsin. Following de-salting, peptides were re-suspended into two separate volumes of 5% acetonitrile, 0.1% formic acid to a final concentration of either (A) 11.5 µg/µL or (B) 52.4 µg/µL, ranging

from 11 – 1100 µg protein, chromatographically separated, and evaluated by PRM. PRM method is described in detail below. Injections were performed in increasing order of protein, and a blank sample (5% acetonitrile, 0.1% formic acid) was injected between analytical samples.

We optimized the amount of protein loading primarily at two protein load levels, first, at 600 µg with Sigma sera and subject serum collected by JanssenDuijghuijsen et al., 2017b. We secondly evaluated a higher protein load up to 3000 µg with baseline subject serum.

Then, we evaluated active subject serum at baseline and selected time points (baseline (0), 30 60, 120, 240, 360 minutes) were evaluated using the same sample preparation parameters, the retention times and loop count were adjusted according to the amount of protein loaded.

At each of these protein levels, we evaluated serum collected at all time points (baseline, 30, 60, 120, 240, and 360 minutes) collected by JanssenDuijghuijsen et al., 2017b. At the higher amounts of protein injected, we monitored column pressures during chromatography and blank injection cleaning methods.

D. LC-MS/MS Analysis of Increased Protein Loading by PRM

Peptide digests were chromatographically separated using an UltiMate 3000RSL® liquid chromatography (UPLC) system (Thermo Scientific™) equipped with a Hypersil Gold C18 1.9 µm, 100 x 2.1 mm, 1.9 µm reversed phase column (Thermo Scientific™) attached with an in-line filter cartridge (1 mm ID, 0.2 µm) (Thermo Scientific™), and a

column oven temperature maintained at 35°C. Mobile phases were as follows, (A) 0.1% (v/v) formic acid in water, (B) 0.1% (v/v) formic acid in acetonitrile, (C) 1:1:1:1 acetonitrile, isopropanol, methanol, water, and (D) was 100% methanol.

The separation column was equilibrated at 7.5% mobile phase B for 7.5 minutes. Peptides were separated using a linear gradient of 7.5 – 25% mobile phase B over 20 minutes at a flow rate of 300 µL/min. Following the gradient elution, the column was washed using the same flow rate for 3 minutes at 98% mobile phase B, followed by 100% mobile phase D for 3 minutes. The separation column was re-equilibrated for 4 minutes.

The total protein load was evaluated using the optimized PRM method as previously described, using the same equipment and instrumentation. However, in these experiments, the javelin and pre-column were replaced with an inline filter cartridge. In summary, peptides were separated using a Hypersil GOLD Vanquish (100 x 2.1 mm, 1.9 µm particle) (Thermo Scientific™) separation column with an inline filter cartridge (1 mm ID, 0.2 µm) (Thermo Scientific™). Peptide retention times were input to the inclusion list, and monitored for stability after each protein level injection. If peptide retention times demonstrated shifts, the retention time windows were adjusted iteratively.

The column pressures were monitored during both chromatographic separation, blank injections, and cleaning methods in all higher protein injections.

E. Data Analysis of Optimized PRM Methodology

Peanut peptide detection was evaluated for intensity and retention time stability. Shifts in retention time due to protein load were recorded and evaluated in Skyline (Maclean et al., 2010).

F. Evaluation of Subject Serum Samples with Higher Protein Loads

We evaluated subject samples previously determined positive by Ara h 6 specific sandwich ELISA (JanssenDuijghuijsen et al., 2017b) using our developed MRM and MRM^3 methods. Subject serum numbers 2 and 7, as identified in the original manuscript, were selected for PRM evaluation, but hereafter, the sera samples are referred to as subject sera 1 and 2, respectively. The evaluated sera included the baseline (0 minutes) and their highest measured absorption time points, 60 and 120 minutes. ELISA results and subject characteristics are described in the original publication (JanssenDuijghuijsen et al., 2017b).

Subject sera (40 µL) were enzymatically digested and de-salted as previously described (page 163). Peptides were re-suspended to 30 µg/µL in 7.5% acetonitrile, 0.1% formic acid in water. Each digest was injected (20 µL) in duplicate and analyzed using the optimized PRM method as previously described (page 162) and analyzed by Skyline using the same analysis parameters (page 154)

IX. Optimization of Chromatography Methodology for Higher Protein Loads

Due to the increased protein requiring chromatographic separation, more robust cleaning procedures were required. Two column wash methods were injected after each analytical sample. The first column wash method consisted of an isocratic flow of 100% mobile phase C (1:1:1:1: methanol:acetonitrile:isopropanol:water) (300 µL/min) for 20 minutes.

The second wash phase utilized the same mobile phases as previously described (page 81). The column was equilibrated (300 µL/min) for 7.5 minutes at 7.5% mobile phase B, followed by a linear gradient from 7.5 to 98% mobile phase B. Mobile phase B was maintained at 98% for 8 minutes. The column was washed (300 µL/min) with 100% mobile phase C for 5 minutes, followed by 100% mobile phase D (300 µl/min) for 5 minutes. The column was re-equilibrated (300 µl/min) at 7.5% mobile phase B for 12 minutes.

X. Multiple Injections

In order to load higher amounts of protein on the liquid chromatography column, we needed to maintain a consistent concentration of re-suspended peptides prior to analysis by LC-MS/MS. To do so, the peptide concentration was maintained at < 20 µg/µL, and, we injected each sample multiple times to achieve the desired amount or protein load. Baseline subject serum was combined with peanut extract (50:1 (w/w)) to monitor protein loading. Other key parameters monitored included chromatographic pressures during samples and blank injections, protein recovery, transition ratios, and peptide retention time stability.

Multiple, repeated injections of a single peptide sample were injected for chromatographic separation. The column flow was maintained at the equilibration settings until all protein was successfully injected onto the column. Once all injections were performed, the chromatographic gradient began. Individual injection methods were written in order to load protein on column before beginning chromatographic elution. The same, previously described mobile phases, were used (page 169). The column loading injection methods included injection parameters and an isocratic solvent flow (7.5% mobile phase B). The final injection included the parameters for chromatographic elution and MS acquisition. Optimized loop count and retention times for increased protein loading were used. After each analytical sample injection (i.e. serum sample), the column was cleaned by injections of blank samples (mobile phase C) using the optimized column cleaning methods described above.

A. Serum Time Course Evaluation using Multiple Injections and Heavy Labeled Peptides

We ultimately evaluated the subject sera (JanssenDuijghuijsen et al., 2017b) samples which we determined were positive. The ten selected peanut peptide targets, with their corresponding heavy peptides, were monitored by PRM LC-MS/MS. AQUA peptide standards (Thermo Scientific™) were synthesized with a heavy isotope labeled $^{13}C(6)^{15}N(4)$ C-terminal arginine or $^{13}C^{15}N(2)$ C-terminal lysine with carbamidomethylation (+57.02 Da) of cysteine residues. Peptides were synthesized to 97% purity, as determined by LC-MS/MS.

Peptide concentration was determined by calculating the heavy to light ratios of peptide peak areas.

B. Subject Serum Injections with Heavy Peptides for Quantification

Subject sera 1 and 2, (JanssenDuijghuijsen et al., 2017b) were analyzed (600 µg) or (2800 µg) using the multiple injection protocol and PRM acquisition with heavy peptides. The loop count was modified for inclusion of heavy peptides. We additionally evaluated multiple injections for recovery of heavy peptides. A mix of heavy labeled peptides were prepared in an equimolar manner. Heavy peptides were injected (8 µL) to achieve 100 fmol (12.5 fmol/µL individual heavy peptide) of each heavy peptide for LC-MS/MS evaluation. The previously described chromatographic wash methods were included after each sample injection (page 169).

C. Data Analysis

All data were analyzed using Skyline v 20 (Maclean et al., 2010) (page 154), as previoulsy described.

RESULTS AND DISCUSSION

I. Discovery Identification of Peanut Peptide Targets

The identification of unique peanut peptides from discovery mass spectra allowed us to generate a list of candidate peptides. This list of candidate peptides was further

refined in subsequent developmental phases. Peptides identified as candidate peptide targets met the criteria described in the methods section (page 150). In total, 52 candidate peptides were identified from discovery data, representing each major allergenic protein family (Figure 3-1). We identified 20 – Ara h 1 peptides, 6 – Ara h 2 peptides, 19 – Ara h 3 peptides, and 7 – Ara h 6 peptides. The variation in number of peptides per protein detected are as anticipated. The prolamin proteins, Ara h 2 and Ara h 6, have much shorter protein sequences than the cupins, Ara h 1 and Ara h 3. As such, we would expect few unique peanut peptides to be detected in the shorter protein sequences, Ara h 2 and Ara h 6.

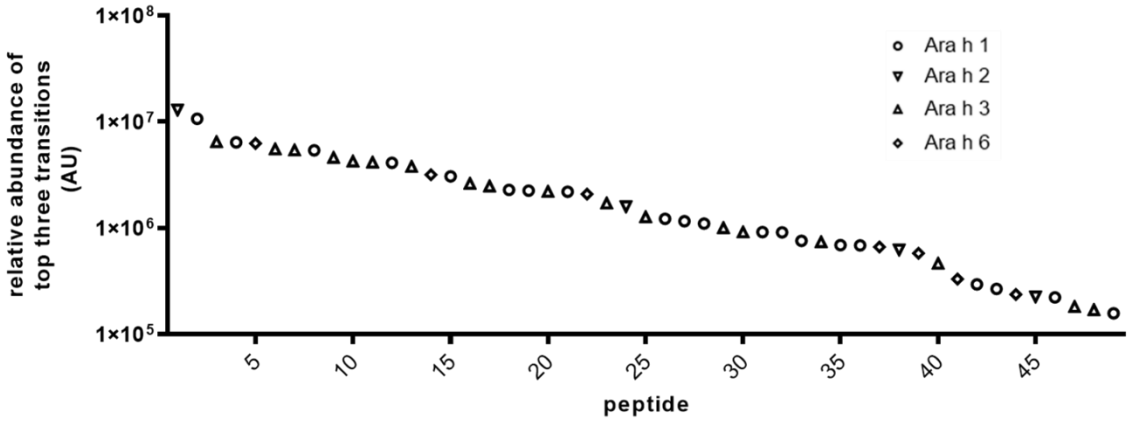


Figure 3-1. The relative abundance of the top peanut peptide targets identified by discovery LC-MS/MS.

By selecting peptide targets using an experimental workflow, we were able to confirm the presence of target peptides in a sample matrix, an advantage over *in silico* peptide selections. Similar workflows using discovery MS for peptide identification for monitoring in targeted acquisitions (e.g. PRM, MRM) has been successfully reported by other investigators (von Bargen et al., 2013; Korte et al., 2016a; Korte et al., 2016b).

Peanuts are typically consumed after roasting, a process known to induce protein aggregation, particularly for Ara h 1 (Koppelman et al., 1999; Schmitt et al., 2010). The heating induced aggregation reduces protein solubility and consequently, overall extractability (Koppelman et al., 1999; Schmitt et al., 2010). As a result of reduced solubility, these aggregated proteins are not efficiently extracted in saline buffer. However, these proteins remain present in the food matrix and are exposed to the immune system after digestion alongside other peanut proteins. To identify all potential unique peanut allergens, all proteins should therefore be included in peptide selections. By using a raw peanut extract, we were able to identify unique Ara h 1 peptides alongside other major allergenic peanut peptides. The incurred matrix (serum-peanut) was prepared at high concentrations of peanut, concentrations greater than the expected *in vivo* levels, to ensure uniqueness of selected peanut peptides in a serum matrix background. It is critical to evaluate each matrix component individually (peanut, serum) as well as in the presence of a matrix background, to eliminate any co-eluting peptide events which may be produced in matrix combination. Secondly, we have chosen to use commercially available human serum as a representative of healthy human adults, which is ultimately our final sample population.

II. PRM

The peptide targets identified in discovery were verified in a targeted PRM method, to confirm their detectability. The same samples were analyzed, and the peptide acceptance criteria were unchanged. After peptide evaluation using PRM acquisition, we reduced our number of targets to 26 peptides. The chromatographic gradient was modified to improve peptide separation and detection. We modified the organic:aqueous solvent gradient by adjusting the total percentage aqueous solvent throughout the course of the run. The time of the gradient was modified, affecting the overall slope. In cases where some peptides were not able to be adequately separated, the peptides with the highest abundance were retained for further evaluations. After chromatographic optimizations, we ultimately selected 13 peanut peptide targets, 4 – Ara h 1, 1 – Ara h 2, 4 – Ara h 3, and 3 – Ara h 6 peptides were selected. The optimal retention times were determined and included during PRM acquisitions for the 12 selected peptides (Figure 3-2).

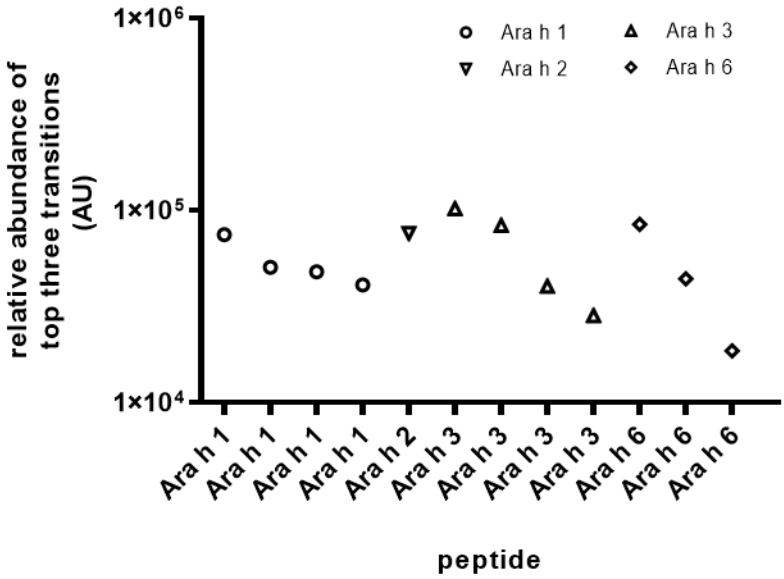


Figure 3-2. Relative abundance of the selected twelve peanut peptides by PRM detection at 0.80 ng peanut protein in a sera-peanut model matrix.

A dilution series of peanut peptides in sera background (sera-peanut) were evaluated by the optimized PRM method and 12 selected peptides (Table 3-3). The best peptide for each protein (Ara h 1, Ara h 2, Ara h 3, Ara h 6) are described in Figure 3-3. Each individual dilution curves is located in Supplementary Figures 3-1 to 3-9. Each sample contained the same amount of serum protein and decreasing amounts of peanut protein (Table 3-1). The dilution series reported detection of all 12 peptides, with a minimum of three transitions per peptide. We were able to detect 0.80 ng peanut protein using the PRM method in a serum matrix (790 ng serum protein) (Figure 3-2). In the dilution series of peanut, lower limits of detection were observed in the peanut only matrix, than when in a non-depleted sera matrix.

Table 3-4 lists the individual limit of detection for each peptide. Protein composition values were determined by Koppelman et al., 2016, and were used here to convert total peanut protein to specific peanut protein (e.g. Ara h 1). Based on the published data, we estimated the total protein composition at 18% Ara h 1, 6% Ara h 2, 70% Ara h 3, and 6% Ara h 6 (Koppelman et al., 2016).

In a comparison of the peanut and sera-peanut matrices, the limit of detection was generally lower in the sera-peanut matrix for most peptides. One Ara h 1 peptide had a lower detection limit in peanut alone. Only this one peptide had a lower detection limit in peanut alone.

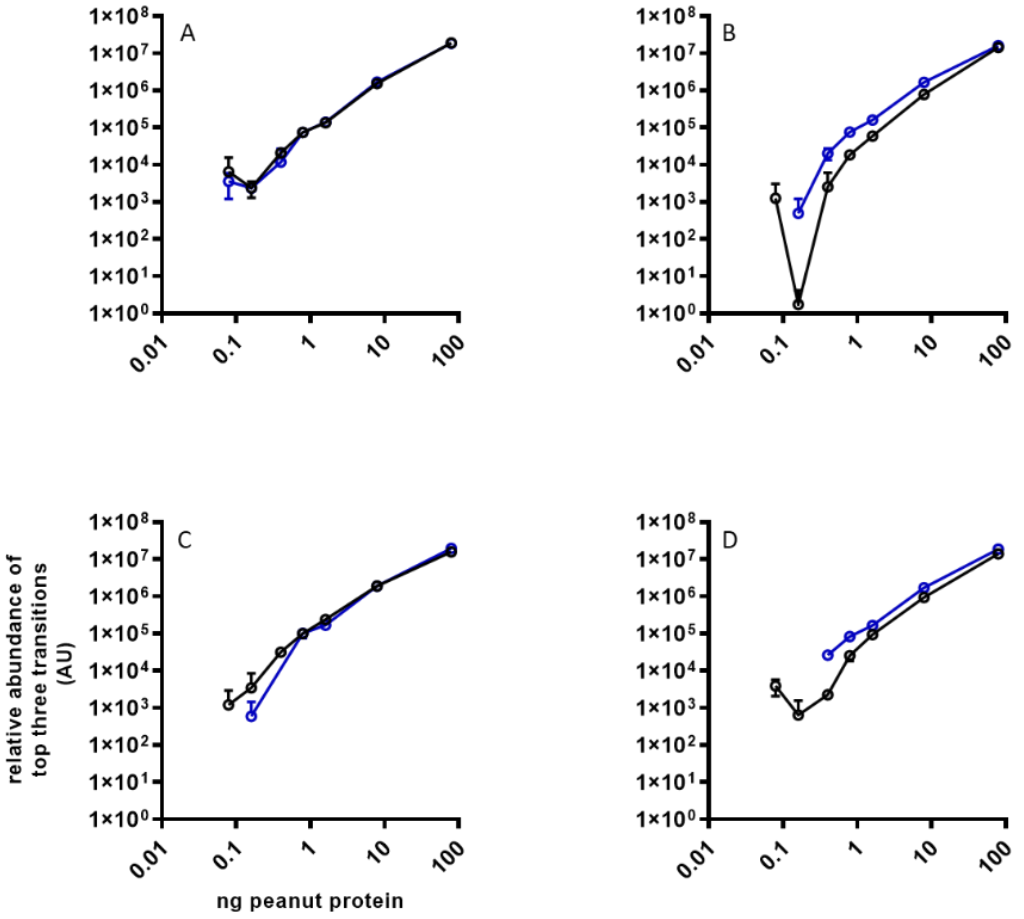


Figure 3-3. Relative abundance of the top three summed transitions for the best performing peptide(s) of (A) Ara h 1, (B) Ara h 2, (C) Ara h 3, (D) Ara h 6 peptides, as detected by PRM, at decreasing quantities of peanut protein in a constant amount of serum. Blue lines represent sera-peanut, where sera remains unchanged. Black lines represent peanut only, in the absence of a sera background.

Protein	Peptide	Peptide Reference Name	m/z	RT
Ara h 1	GTGNLELVAVR	GTGN	564.8222++	23.3
	NNPFYFPSR	NNPF	571.2749++	23.5
	DLAFPGSGEQVEK	DLAF	688.8383++	20.5
	EGALMLPHFNSK	EGAL	448.5641+++	15.5
Ara h 2	NLPQQCGLR	NLPQ	543.2797++	17.2
Ara h 3	TANDLNLLILR	TAND	628.3721++	26.0
	SPDIYNPQAGSLK	SPDI	695.3541++	20.8
	QIVQNLR	QIVQ	435.7614++	17.6
	AHVQVVDSNGNR	AHVQ	432.5532+++	13.9
	WLGLSAEYGNLYR	WLGL	771.3910++	24.9
Ara h 6	CDLDVSGGR	CDLD	489.7191++	16.6
	ELMNLPQQCNFR	ELMN	775.3661++	23.2
	VNLKPCEQHIMQR	VNLK	413.9644++++	18.7

Table 3-3. Peanut peptides evaluated by PRM.

		peanut					sera-peanut				
Peptide	Protein	pg peanut protein	pg Ara h 1	pg Ara h 2	pg Ara h 3	pg Ara h 6	pg peanut protein	pg Ara h 1	pg Ara h 2	pg Ara h 3	pg Ara h 6
DLAFPGSGEQVEK	Ara h 1	400	72				790	142.2			
EGALMLPHFNSK	Ara h 1	790	142.2				790	142.2			
GTGNLELVAVR	Ara h 1	790	142.2				400	72			
NNPFYFPSR	Ara h 1	790	142.2				400	72			
NLPQQCGLR	Ara h 2	1600		96			400		24		
WLGLSAEYGNLYR	Ara h 3	7900			5530		400			280	
TANDLNLLILR	Ara h 3	790			553		400			280	
SPDIYNPQAGSLK	Ara h 3	400			280		400			280	
QIVQNLR	Ara h 3	790			553		400			280	
ELMNLPQQCNFR	Ara h 6	1600				96	790				47.4
CDLDVSGGR	Ara h 6	1600				96	790				47.4
VNLKPCEQHIMQR	Ara h 6	790				47.4	790				47.4

Table 3-4. Detection limit of individual peanut peptides by PRM.

In the sera-peanut matrix, Ara h 3 was consistently detectable at 280 pg Ara h 3 (400 pg peanut protein). In the sera-peanut matrix, Ara h 2 had a lower detection limit of 24 pg Ara h 2 (400 pg peanut protein). Ara h 6 had a lower detection limit in the sera-peanut matrix of 47.4 pg Ara h 6 (790 pg total peanut protein).

The overall limit of detection for all peptides was concluded to be 790 pg peanut protein, which is equivalent to 1.0 ppm peanut protein (4.0 ppm total peanut) for the PRM method.

III. MRM

The MRM settings were optimized for peptide detection and are described in the materials and methods section. The top six transitions reported in the PRM evaluations in both peanut and sera-peanut were evaluated for initial detection by MRM. Transitions were not scheduled during this initial evaluation. We determined the four best transitions for each peptide for the MRM method.

The CE for each peptide, a key fragmentation setting, was first determined by evaluating the relative abundance reported by the selected top four peptide transitions by MRM. This is typically performed using heavy peptides and direct infusion, however due to the uncertainty of our peptide status following gastric digestion, we evaluated the CE using our model samples (peanut, sera-peanut).

Following transition selection and retention time determinations, we evaluated the detection limit for MRM. Figure 3-6 and Table 3-5 demonstrate the peptide LODs for

MRM. Not all peptides included in the PRM method were suitable for MRM analysis, therefore fewer peptides were evaluated by MRM. In our final method, we were able to detect 4 – Ara h 1, 1 – Ara h 2, 4 – Ara h 3, and 1 – Ara h 6 peptides. The peptides QIVQ, VNLK, CDLD were not detectable by the MRM method. The peptides, (QIVQ, CDLD), exhibited poor linearity and relative abundance during PRM method development.

Peptides, and their transitions, included in the MRM LOD evaluations are described in Table 3-6. We monitored a minimum of two transitions per peptide in this evaluation. In general, the detection limits were relatively similar to the PRM detection limits. Similarly, in the sera-peanut matrix, generally equal or lower limits of detection were observed for nearly all peptides. The best performing peptides were Ara h 2 (NLPQ) and Ara h 6 (ELMN) with detection at 15 pg Ara h 2 or 74.4 pg Ara h 6.

One Ara h 3 peptide (WLGL) had a noticeably higher (13100 pg Ara h 3) detection limit in the peanut only matrix for two of its five identified transitions (Table 3-7). Since there are five identified transitions for this peptide, these transitions can be removed from the method. Interestingly, when in the sera-peanut matrix, this peptide (WLGL) exhibits detection at 2600 pg Ara h 3, demonstrating an increased solubility in a serum matrix, observed in both PRM and MRM acquisitions. Since serum is ultimately the anticipated background matrix, we chose to include this peptide for evaluations.

Figure 3-4 describes the relative abundance of serially diluted peanut in the presence and absence of a sera matrix, as detected by MRM. Overall, each transition was detected similarly in both matrices and exhibited linearity over the course of the dilution series.

Protein	pg peanut protein LOD	ppm peanut protein LOD	ppm peanut LOD
Ara h 1	1240	1.53	6.11
Ara h 2	1240	1.53	6.11
Ara h 3	3730	4.58	18.3
Ara h 6	1240	1.53	6.11

Table 3-5. MRM detection limit table determined by the dilution series.

protein	peptide	transition	MS1	MS2
Ara h 2	NLPQ	y7+	543.2797++	858.43
		y6+		381.10
		y7++		429.72
Ara h 3	SPDI	y7+	695.3541++	700.40
		y8+		814.44
		y9+		977.51
Ara h 1	DLAF	y9+	688.8383++	930.45
		b4+		447.22
Ara h 1	GTGN	y6+	564.8222++	686.42
		y3+		345.22
Ara h 6	ELMN	y7+	775.3661++	949.43
Ara h 1	NNPF	y4+	571.2749++	506.27
		y7+		913.46
Ara h 1	EGAL	y7++	448.5641+++	421.73
		y5++		316.66
Ara h 3	WLGL	y5+	771.3910++	622.33
		b3+		357.19
		y6+		785.39
		y8+		985.47
		y7+		914.44
Ara h 3	TAND	y3+	628.3721++	401.29
		y6+		741.50
		y7+		854.58

Table 3-6. MRM peptide transitions, parent and fragment ion m/z.

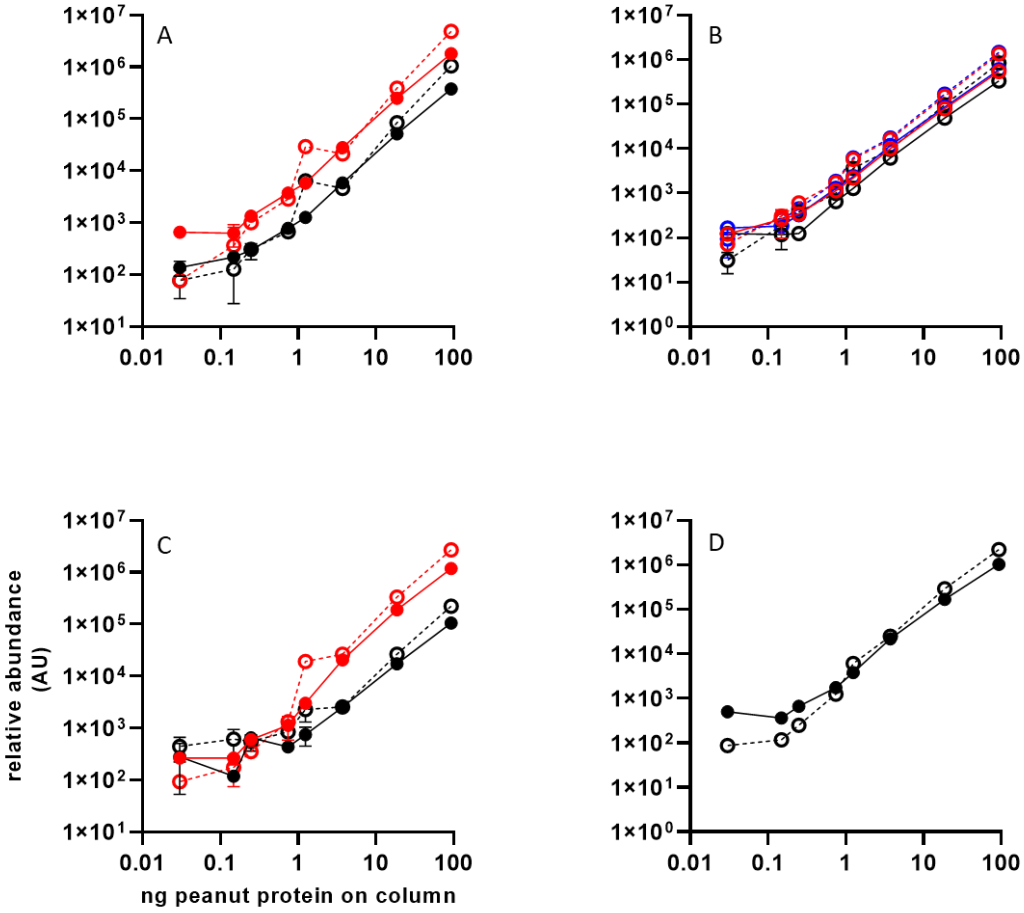


Figure 3-4. Relative abundance of (A) Ara h 1, (B) Ara h 2, (C) Ara h 3, (D) Ara h 6 peptides, as detected by MRM, at decreasing quantities of peanut protein in a constant amount of serum. Each colored line represents individual transitions for a given peanut protein. Solid lines represent sera-peanut matrix, dashed lines represent peanut matrix, in the absence of sera background.

			peanut					sera-peanut				
Peptide	Transition	Protein	pg peanut protein	pg Ara h 1	pg Ara h 2	pg Ara h 3	pg Ara h 6	pg peanut protein	pg Ara h 1	pg Ara h 2	pg Ara h 3	pg Ara h 6
DLAFPGSGEQVEK	b4+	Ara h 1	1240	149				750	90			
	y9+		1240	149				750	90			
EGALMLPHFNSK	y5++	Ara h 1	1240	149				1240	149			
	y7++		1240	149				1240	149			
GTGNLELVAVR	y3+	Ara h 1	750	90				750	90			
	y6+		750	90				750	90			
NNPFYFPSR	y4+	Ara h 1	750	90				750	90			
	y7+		750	90				750	90			
NLPQQCGLR	y6+	Ara h 2	ND		ND			ND		ND		
	y7+		250		15			250		15		
	y7++		1240		74.4			1240		74.4		
WLGLSAEYGNLYR	b3+	Ara h 3	18700			13100		3730			2600	
	y5+		18700			13100		3730			2600	
	y6+		3730			2600		3730			2600	
	y7+		3730			2600		3730			2600	
	y8+		3730			2600		3730			2600	
TANDLNLLILR	y3+	Ara h 3	750			525		750			525	
	y6+		750			525		750			525	
	y7+		750			525		750			525	
SPDIYNPQAGSLK	y7+	Ara h 3	150			105		150			105	
	y8+		150			105		250			175	
	y9+		150			105		250			175	
ELMNLPQQCNFR	y7+	Ara h 6	1240				74.4	1240				74.4

Table 3-7. Detection limit of individual peanut peptides by MRM.

IV. MRM^3

Many of the MS2 peptide fragments had non-specific mass losses, largely attributed to $–NH_3$ or $–H_2O$ losses. To provide the most sensitive detection method, we enabled Q0 trapping. Interestingly, a peptide that produces a good MS2 fragment is not always able to fragment again, producing secondary fragment ions, MS3 ions. Product ions were manually evaluated against a predicted fragment masses derived from each MS2 peptide. MRM^3 transitions were determined using an automated script available in Analyst. The top five resulting product ions were evaluated for fragmentation patterns. We were able to identify six true MS3 fragment ions from Ara h 1, Ara h 2, and Ara h 3. No MS3 ions were identified from Ara h 6. These ion fragments are considered 'true MS3' targets, as they did not have any non-specific mass losses (e.g. $–NH_3$, $–H_2O$).

Due to the challenge of identifying 'true MS3' fragments, we assessed ion detection without the additional fragmentation step, but rather, used the non-fragmented m/z again for detection by MRM^3 in the linear ion trap of the third quadrupole. Using this approach, we were able to detect all peptides included in the MRM method. With the combination of both non-fragmented peptides and MS3 fragmented ('true') peptides, we were able to generate an MRM^3 method for all major allergenic peanut proteins (Table 3-8). We included 18 non-fragmented peptide transitions in the MRM^3 methodology (Table 3-8). In total, 24 peptide transitions were included in the final MRM^3 methodology (Table 3-8). In order to keep cycle times low, we divided the MRM^3 acquisition method into eight periods. Without dividing the method into eight periods, the cycle time for all the MRM^3 transitions would have been very long and reducing overall ion detection.

	Peptide	Peanut protein	ms1 m/z	ms2 m/z	ms2 ion	ms3 m/z	ms3 ion
MS3 fragment ion transitions	NNPFYFPSR	Ara h 1	571.28++	913.32	y7+	408.12	y6++
	NLPQQCGLR	Ara h 2	543.28++	429.71	y7++	633.20	y5+
	WLGLSAEYGNLYR	Ara h 3	771.39++	985.36	y8+	785.27	y6+
	WLGLSAEYGNLYR	Ara h 3	771.39++	914.44	y7+	785.27	y6+
	SPDIYNPQAGSLK	Ara h 3	695.35++	977.37	y9+	700.29	y7+
	TANDLNLLILR	Ara h 3	628.37++	854.47	y7+	341.13	b3+
Non-fragmented MS2 ions	DLAFPGSGEQVEK	Ara h 1	688.84++	930.37	y9+		
				447.29	b4+		
	NNPFYFPSR	Ara h 1	571.27++	506.21	y4+		
	GTGNLELVAVR	Ara h 1	564.82++	345.22	y3+		
				686.33	y6+		
	EGALMLPHFNSK	Ara h 1	448.56+++	421.13	y7++		
				316.63	y5++		
	NLPQQCGLR	Ara h 2	543.28++	858.34	y7+		
				381.1	y6++		
	CDLEVESGGR	Ara h 2	561.25++	604.17	y6+		
	WLGLSAEYGNLYR	Ara h 3	771.39++	622.22	y5+		
				357.07	b3+		
				785.34	y6+		
	SPDIYNPQAGSLK	Ara h 3	695.35++	700.41	y7+		
				814.41	y8+		
	TANDLNLLILR	Ara h 3	628.37++	401.38	y3+		
				741.43	y6+		
	ELMNLPQQCNFR	Ara h 6	775.37++	949.34	y7+		

Table 3-8. MS3 peanut peptide targets for monitoring.

We evaluated the dilution series of the MRM^3 method using the same dilutions as described in the MRM method. As depicted in Figure 3-5, peptides were more readily detected in the sera-peanut matrix, than in peanut alone. These results are consistent with previous dilution series' results (PRM, MRM). The dilution series was performed using all 24 identified targets.

For certain peptides, there was a distinct difference in the reported abundances among transitions. The limit of detection results were highly variable using the MRM^3 method, with some peptides performing much better than others. Figure 3-5 demonstrates detection of serially diluted peanut in the presence and absence of serum. The prolamin peptides, Ara h 2 (NLPQ) and Ara h 6 (ELMN) exhibited low levels of detection at 1.8 – 74 pg Ara h 2 and 45 pg Ara h 6 (Figure 3-5 B and D, respectively). The best performing Ara h 3 peptide (SPDI) ranged from 21 – 525 pg Ara h 3 (Figure 3-5 C). We observed a lack of linearity in some MRM^3 analyzed peanut peptides, indicating these peptides may not be suitable as MRM^3 targets. The Ara h 3 peptide, SPDI, demonstrates high linearity and suitably sensitive detection. Contrastingly, a different Ara h 3 peptide, TAND exhibited poor linearity, and was less sensitive than SPDI (Figure 3-5 C).

Determination of true MS3 fragments was challenging due to the lack of tertiary product ions. Some MS2 peptide targets fragmented more readily than other peptides. As such, peptides which make excellent MS2 targets, are not always suitable MS3 targets.

In summary, we observed no consistent LOD measurements between peptides of the same protein family. However, the Ara h 3 peptide, SPDI, exhibited sufficient sensitivity and linearity with its true MS3 peptide fragment ions. We were able to detect 30 pg peanut protein by this peptide, equivalent to 0.04 ppm peanut protein (0.15 ppm peanut). It is critical to note this observed sensitivity was in a prepared model matrix. If this peptide exists in serum following gastric digestion, it would be a suitable target for low detection monitoring. Furthermore, peanut peptides which have previous *in vivo*

peanut protein detections have been accomplished by ELISA's targeting the prolamins, Ara h 2 and Ara h 6, which are known to remain partially intact following digestion.

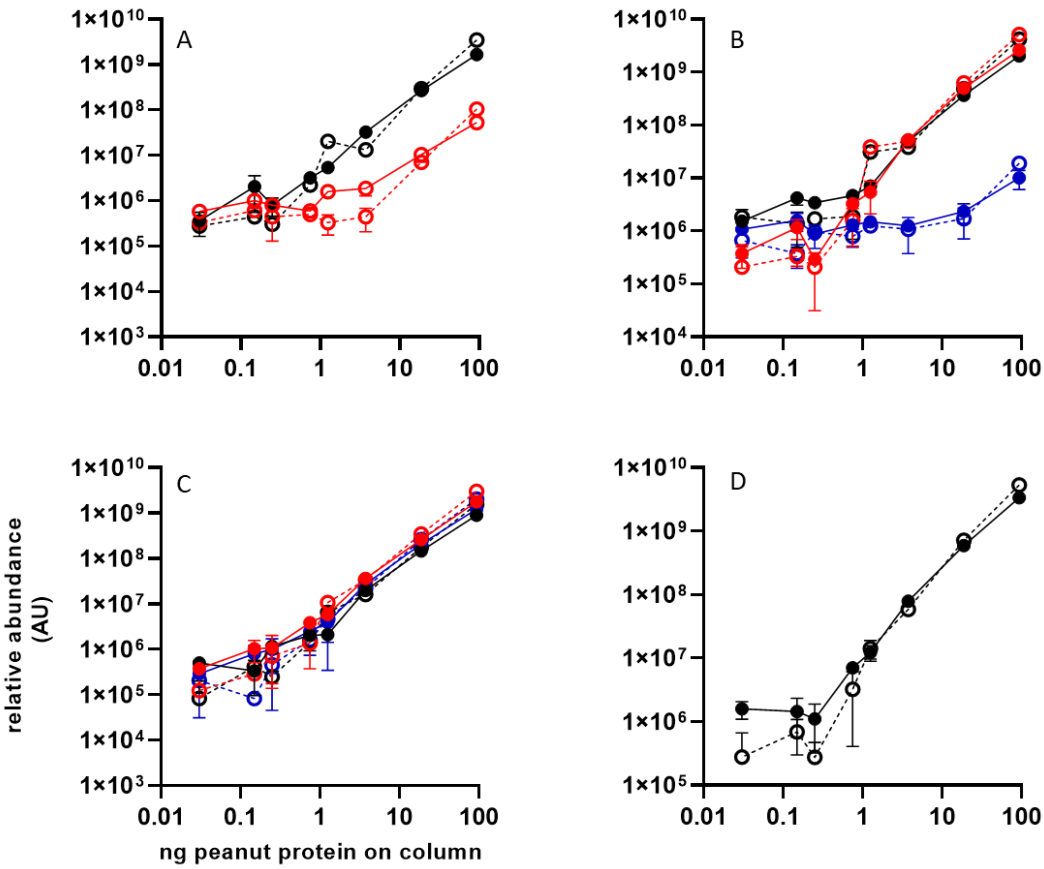


Figure 3-5. Relative abundance of (A) Ara h 1, (B) Ara h 2, (C) Ara h 3, (D) Ara h 6 peptides, as detected by MRM^3, at decreasing quantities of peanut protein in a constant amount of serum. Each colored line represents individual transitions for a given peanut protein. Solid lines represent sera-peanut matrix, dashed lines represent peanut matrix, in the absence of sera background.

To assess any potential differences between MRM^3 peptide transitions, we evaluated the 'true MS3' fragments only using our optimized methodology (Figure 3-6). The results do not show any improvements in overall detection compared to the MRM method. The Ara h 2 peptide, NLPQ, and the Ara h 3 peptide, SPDI, both demonstrate good linearity and similar detection in both peanut and sera-peanut matrices. In comparison to the acquisition method which included both 'true MS3' fragment ions and the non-fragmented MS2 ions, this analysis using only the 'true MS3' ions showed no improvements in peptide detection.

However, the limit of detection does not show any improvements, therefore using the 'true' secondary peptide fragments may not be a suitable method for low detection of peanut proteins in serum.

Overall, for the MRM^3 acquisitions, detection of non-fragmented MS2 transitions was more successful than detection of secondary fragment ions (MS3) in the model sera-peanut matrices.

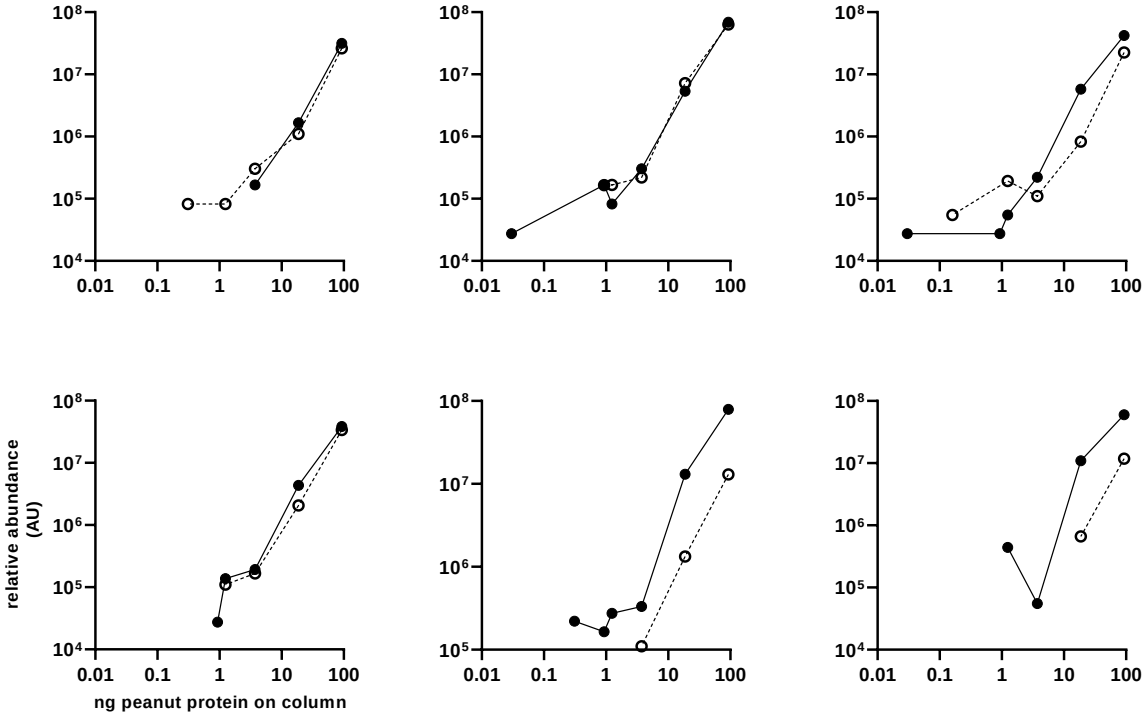


Figure 3-6. Relative abundance of the dilution series while monitoring true secondary fragment ions (MS3 ions).

V. Evaluation of Subject Sera Samples by MRM and MRM^3

Subject sera samples (serum 1, serum 2) at baseline (0 minutes) and selected time points (60, 120 minutes) were evaluated using MRM and MRM^3 methods. We evaluated two individual subjects (JanssenDuijghuijsen et al., 2017b), determined previously positive by an Ara h 6 ELISA. In this experiment, we injected 600 µg total protein for evaluation. However, no positive peanut detections were made in subject serum by either MRM or MRM^3.

One key problem with these evaluations was the use of commercially prepared serum. Commercial serum is prepared from a pool of subject serum which has been passed through a 0.1 micron filter and sterilized. Commercially prepared serum is likely slightly less concentrated than a single subject serum, and as a result, the subject serum samples exhibit a lower signal-to-noise ratio due to poor peak isolation.

Another key issue impacting these experimental results was the individual dilution of serum and raw peanut extract prior to enzymatic digestion for LC-MS/MS analysis. By diluting the analytical matrix components, this may ultimately be less representative of serum following the consumption of peanut. Although we observed LODs for the MRM and some MRM^3 peptides to previously published ELISA results (JanssenDuijghuijsen et al., 2017b), we were unable to replicate these results using our current methodology.

VI. Optimization of High-Resolution PRM

In order to load more protein for LC-MS/MS analysis, we optimized the reduction and trypsin (digestive enzyme) concentrations. We increased the reduction incubation time to 20 minutes. We simultaneously evaluated the trypsin concentration at enzyme:substrate ratios of 1:50 and 1:100 (w/w).

In our final method, we were able to detect ten peptides robustly including, 4 – Ara h 1, 1 – Ara h 2, 4 – Ara h 3, 2 – Ara h 6 peptides. The peptides QIVQ, CDLD were poorly detected in the previous PRM study, and not detectable by MRM, therefore we chose to remove them from our methodology.

The experiments described within this section were conducted using a sera-peanut matrix prepared with subject serum as it is the most representative preparation material for evaluating experimental conditions since it has not been filter sterilized. For all sera-peanut model matrices prepared for evaluations in the following sections, we used baseline subject serum.

SDS-PAGE indicated the current digestion protocol is sufficient for digestion of serum proteins (Figure 3-7). Secondly, the SDS-PAGE demonstrates, the modified digestions were also suitable for digestion procedures. Numerous peptides were visualized <10 kDa, and a majority of these peptides gradually decreased in intensity after 24 hours of digestion. To conserve trypsin, we opted to use the 1:100 w/w ratio, since no differences in overall digestion patterns were observed by SDS-PAGE.

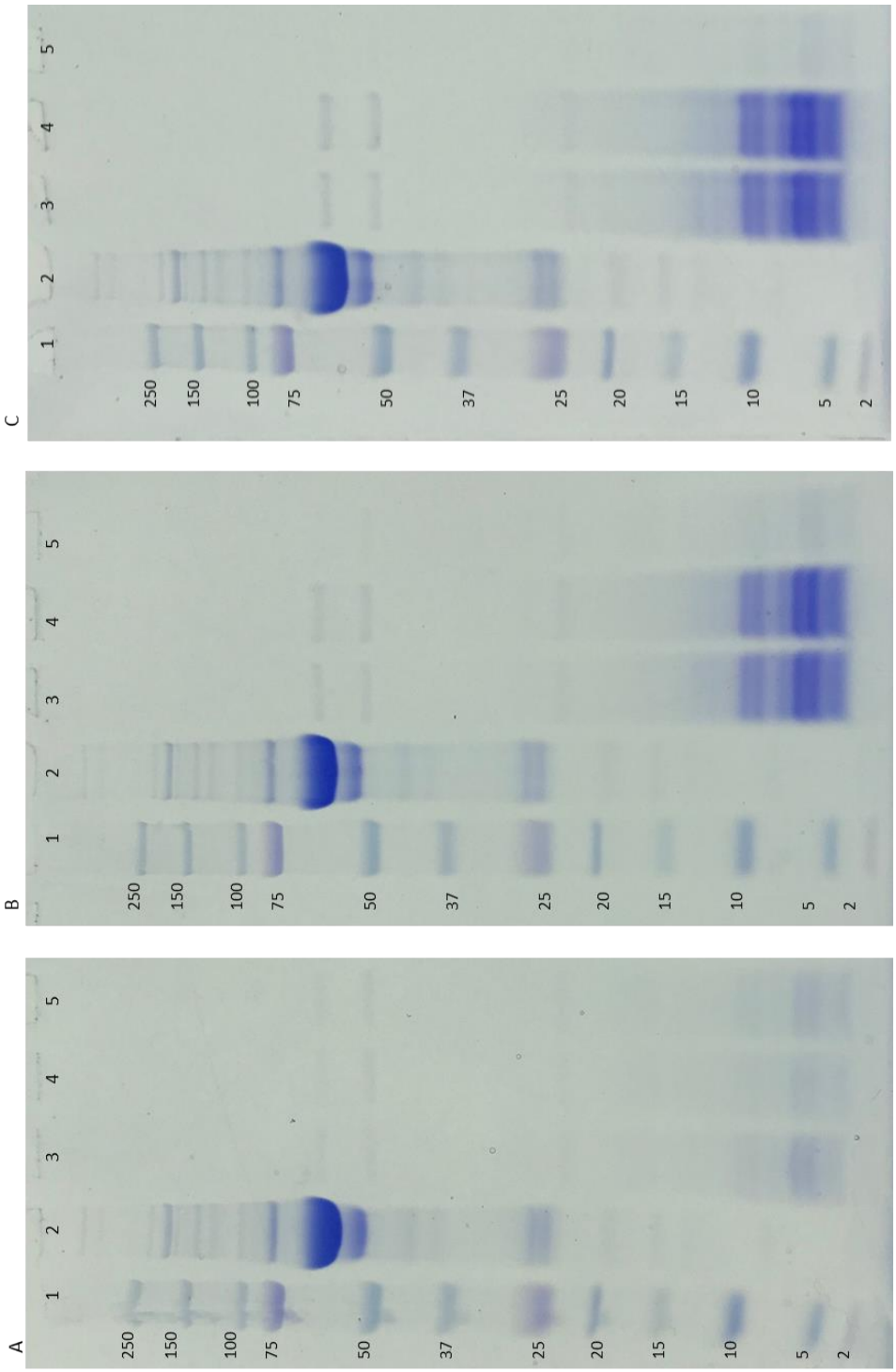


Figure 3-7. SDS-PAGE evaluation of modified protein reduction and trypsin digestion procedures. Each panel represents different experiment conditions (A) 1:50 (w/w) enzyme:substrate ratio with 5 minute reduction at 95°C; (B) 1:50 (w/w) enzyme:substrate ratio with 20 minutes reduction at 95°C; (C) 1:100 (w/w) enzyme:substrate ratio with 20 minutes reduction at 95°C. Lanes 3, 4, and 5 represent digestion aliquots taken at 3, 8, and 24 hours.

VII. Evaluation of Protein Loading for analysis by PRM

In this series of experiments, we reduced the number of PRM targets to the 10 best performing peptides, which are readily detectable and able to be separated chromatographically into distinct peaks (Table 3-9).

Protein	Peptide	m/z
Ara h 1	DLAFPGSGEQVEK	688.8383++
Ara h 1	GTGNLELVAVR	564.8222++
Ara h 1	NNPFYFPSR	571.2749++
Ara h 1	EGALMLPHFNSK	448.5641+++
Ara h 2	NLPQQCGLR	543.2737++
Ara h 3	SPDIYNPQAGSLK	695.3541++
Ara h 3	WLGLSAEYGNLYR	771.3910++
Ara h 3	TANDLNLLILR	628.3721++
Ara h 6	VNLKPCEQHIMQR	413.9644++++
Ara h 6	ELMNLPQQCNFR	775.3661++

Table 3-9. PRM peptide targets monitored with increased protein loading.

In our optimized PRM methodology and acquisition settings, we changed the chromatographic column to a larger column (100 x 2.1 mm, 1.9 um) to facilitate increased protein loading.

We first evaluated a protein load of 600 µg for detection of peanut peptides in serum. In our preliminary studies, we determined 600 µg of protein was suitable for loading due to stable peptide retention times and peak area recovery. Beyond a 600 µg protein load, we observed greater shifts in retention times. Figure 3-8 demonstrates peanut peak area of the top 3 abundant peptide transitions per each increasing amount of protein injected.

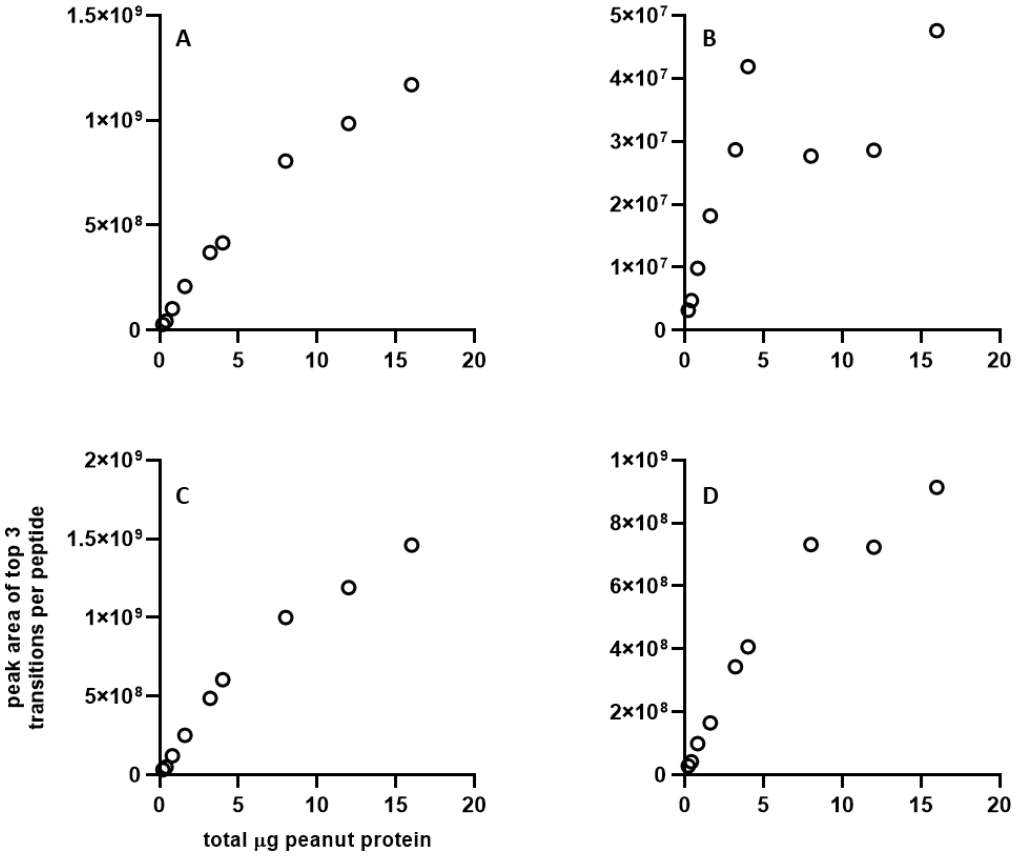


Figure 3-8. Peak area of top 3 transitions per peptide for each amount of protein evaluated; each major allergenic peanut proteins are represented. (A) Ara h 1, DLAF; (B) Ara h 2, NLPQ; (C) Ara h 3 SPDI; (D) Ara h 6, ELMN.

VIII. Evaluation of Subject Samples with an Increased Protein Load

Two subject sera (serum 1, serum 2), at baseline, 60, and 120 minutes, were evaluated at a 600 µg load for PRM analysis. The experiment was performed twice, on two separate occasions. A positive control, (blank subject sera-peanut 100:1 (w/w)), was additionally evaluated.

We observed detection of peanut proteins in subject 1 at 60 minutes post-consumption, in both individual analytical injections by PRM with a 600 µg total protein load. We detected four peanut peptides (SPDI, VNLK, WLGL, TAND), with four detectable transitions for SPDI, VNLK, WLGL and four detectable transitions for TAND (Figure 3-9; Supplementary Table 3-1). The peptide elution times and fragment ion tolerance values were within the acceptance criteria.

We also observed detection of peanut proteins in subject 2 at 120 minutes in both individual analytical injections. The same peptides (SPDI, VNLK, WLGL, TAND) detected in subject serum 1 (60 minutes) were also detected in subject serum 2 (120 minutes). For all detected peptides, five transitions were observed, except for VNLK, which had four detectable transitions. The peptide elution times and fragment ion tolerance values were within the acceptance criteria.

The peptide transition ratios are illustrated in Figure 3-9. The same peptide transitions were detected in each individual subject serum at two time points. We detected between 5-8 points across the peaks. Although, it is preferred to have approximately 10-15 points across the peak for quantification, this was convincing evidence of peanut peptide detection in serum. Since we did not evaluate these samples with a peptide standard, we are unable to accurately quantify the amount of peanut

protein detected. Further, we were able to detect multiple peptides from different proteins, strengthening the evidence that these proteins remain detectable *in vivo.*

One additional key factor which must be taken into consideration is the serum was collected from healthy subjects. The rate of uptake of allergenic food proteins is suggested to differ between sensitized and non-sensitized individuals (Reitsma et al., 2014; Samadi et al., 2018).

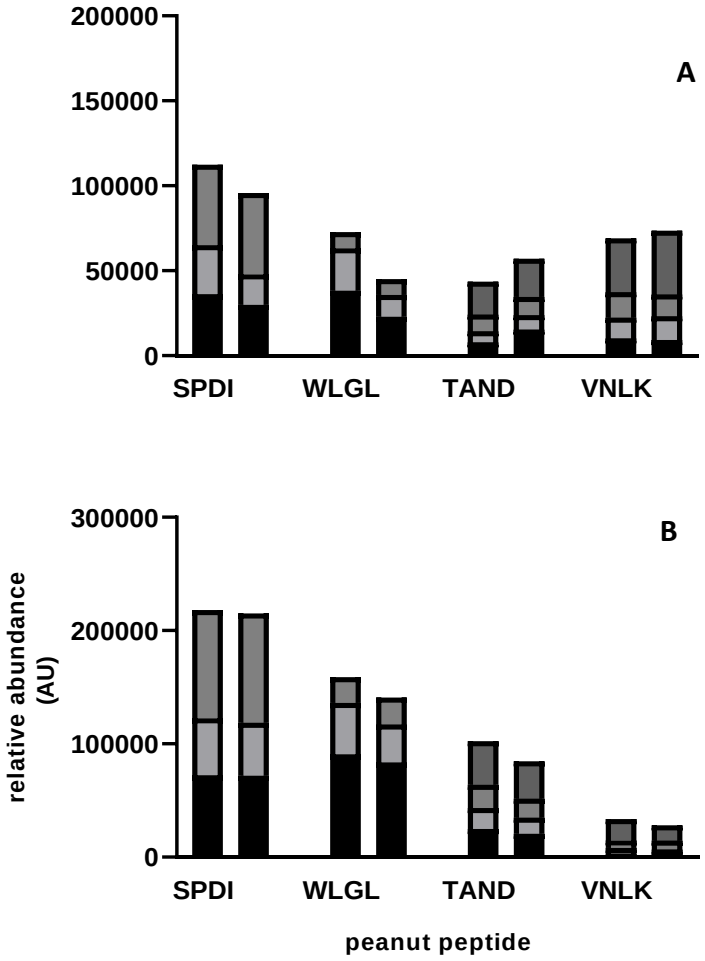


Figure 3-9. Peanut peptide transition patterns detected in human serum collected after peanut consumption; (A) Subject sera 1 collected 60 minutes after consumption; (B) Subject sera 2 collected 120 minutes after consumption.

The QExactive+ offers high mass accuracy detection of fragment ions, which is beneficial in detection of low abundant peptides (Domon and Aebersold, 2010; Gallien et al., 2013).

We then evaluated the entire time course (0, 30, 60, 120, 240, and 360 minutes) of the positive subject serum 1 and 2, to determine the overall absorption. Evaluation of individual absorption curves indicated no positive detection of peanut peptides. This is unexpected since we previously reported positive detection for these same subject sera using the same experimental conditions and PRM acquisition methodology. Upon further evaluation of the collected data, we observed gradual increases in pressure over the course of sequence injections. The increased pressure suggests a slight column blockage due to an accumulation of protein on the column that are not removed by the current cleaning protocols following chromatographic separation, and prior to injection of the subsequent sample. Consequently, we developed improved cleaning methods of the separation column to facilitate higher protein loading.

Upon re-evaluation, we again observed no positive detection of peanut in the same set of subject sera. However, the pressure of the column was largely stabilized and no cumulative increases in pressure were observed (data not shown).

IX. Multiple Injections

Following the evaluation of an entire serum time course and improvements to cleaning procedures, we wanted to again evaluate increased loading since detection in the serum time course samples was unsuccessful. The event of duplicate detection in

separate subject samples with a robust transition pattern was convincing evidence that peanut peptides were detectable in serum. However, detection was not robust or reproducible, which led us to the assumption that by loading more protein, while maintaining a clean column after each injection, we would be able to robustly detect peanut peptides.

The concentration of all serum samples was maintained the same (<20 µg/µL). In preliminary studies, at peptide concentrations >20 µg/µL we observed precipitation in re-suspended sample digests, after being frozen and thawed. One freeze-thaw cycle resulted in opaque particles, which were extremely challenging to re-solubilize prior to injection. There was a significant reduction in detection of peanut peptides prepared at concentrations >20 µg/µL. For all sample digests $\leq$ 20 µg/µL, the detectable area for peanut peptides were comparable. Given our experimental limitations, we developed an alternative injection method, wherein the same sample, from the same vial, was injected repeatedly onto the separation column. Once the desired amount of protein was injected onto the column, gradient elution began. The theoretical binding capacity of the separation column was calculated by determining the volume of cylinder ($V=\pi \bullet r^2 \bullet h$), resulting in cm3 unit or mL. Knowing that the separation column has an 11% carbon load, we assumed the percentage carbon load is equivalent to peptide binding capacity. We then multiplied the carbon load (11%) by the total calculated column volume, resulting in a total binding capacity of 38.09 mg protein. We reduced the calculated value by a factor of 10 as a margin of safety, resulting in a total estimated binding capacity of 3.809 mg protein.

In our experiments we tested up to 3 mg of protein injected (sera-peanut), using our multiple injection protocol. Using the cleaning protocols described previously, we observed no drastic shifts in pressure over the method run and no subtle increases in pressure throughout the sequence run.

X. Multiple Injections with Isotopically Heavy Labeled Peptides and Subject Serum Samples

Isotopically heavy labeled peptides were added to subject serum samples prior to PRM evaluations. We injected 2800 µg (2.8 mg) of protein for evaluation.

We were able to observe detection of heavy peptides in all samples, which eluted at the predicted retention times of the light peptides. We observed no positive detection of peanut peptides in the increased loaded samples, even with stable pressures within the expected ranges and peptide retention times, observed.

Due to the lack of detection at the increased load of protein on column, we suspect high loading is not beneficial to achieving peanut protein detection. It is more likely, peanut peptides are detectable between 600 – 2800 µg protein, with the inclusion of optimized cleaning methods and heavy peptides.

One key difference between the two protein loading amounts, 600 and 2800 µg, were the number of injections. Detection observed at 600 µg protein load was performed in one injection, whereas in the higher protein load (2800 µg), multiple injections were performed. Although according to our preliminary investigations, we determined the multiple injections were sufficient and did not show significant decreases in peak area.

CONCLUSIONS

Throughout Chapter 2 and Chapter 3, we evaluated two sera collections. In sera collected by Baumert et al., 2009, the proteins were encapsulated, and digestion assumingly began once the capsules reached the stomach. The serum samples collected in this manner, are useful in a research and proof of concept studies, but, are less representative of our final sample matrix, serum collected following consumption.

The second set of subject serum evaluated was collected from individuals who consumed roasted peanuts (100 g light roasted peanuts within 10 minutes). In these serum samples, the peanut material was exposed to the entire gastrointestinal tract and collected after consumption, making these sera samples more representative of the ultimate anticipated matrix.

Regardless of the form or the administration method, we were unable to robustly detect peanut proteins in serum. We were able to detect four peanut peptides, from multiple peanut proteins (Ara h 1, Ara h 3, and Ara h 6) in two subjects, at two serum time points collection using a targeted method at a significantly larger protein injection than typical proteomic methods. These subject sera also reported positive detection for Ara h 6 by JanssenDuijghuijsen et al., 2017b. We were unable to reproduce these results by injecting more protein for analysis. This is suggestive there is a median range where these proteins are detectable, likely between 600 – 2000 µg protein. Utilization of thorough and validated cleaning procedures following each sample injection, will likely improve robust detection

The utilization of targeted methods was highly advantageous in comparison to depletion methods and discovery LC-MS/MS. We evaluated three different targeted methods, however, we were only able to achieve detection once we increased our analytical protein injections to 600 µg. PRM methods with their high-resolution are the ideal instrument platform to use for *in vivo* detection of peanut proteins. PRM methods allow for comparison of peptide transition ratios, which is an additional confirmatory step for detection at extremely low peptide concentrations. By increasing our injections 2 and 3-fold by the use of multiple injections we were still unable to detect peanut peptides, suggesting a direct increase in protein load is not equivalent to recording an identical increase in peptide detection. The additional use of heavy peptides during these multiple injection experiments also demonstrated we were able to detect the heavy peptides at consistent ratios, indicating the multiple injections were adequate and did not suffer from significant signal losses. It is probable that the S/N ratio was acceptable for detection of peptides at 600 µg, however, this must be further evaluated in order to develop a routine analytical method.

REFERENCES

Addona, T.A., Abbatiello, S.E., Schilling, B., Skates, S.J., Mani, D.R., Bunk, D.M., Spiegelman, C.H., Zimmerman, L.J., Ham, A.J.L., Keshishian, H., et al. (2009). Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat. Biotechnol. *27*, 633–641.

Adkins, J.N., Varnum, S.M., Auberry, K.J., Moore, R.J., Angell, N.H., Smith, R.D., Springer, D.L., and Pounds, J.G. (2002). Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. Mol. Cell. Proteomics *1*, 947–955.

Anderson, N., and Anderson, N. (2002). The Human Plasma Proteome History, Character, and Diagnostic Prospects. Mol. Cell. Proteomics *1*, 845–867.

Von Bargen, C., Rg Dojahn, J., Waidelich, D., Humpf, H.-U., and Brockmeyer, J. (2013). New Sensitive High-Performance Liquid Chromatography−Tandem Mass Spectrometry Method for the Detection of Horse and Pork in Halal Beef. J. Agric. Food Chem. *61*, 11986–11994.

Baumert, J.L., Bush, R.K., Levy, M.B., Koppelman, S.J., Nordlee, J.A., Hefle, S.L., and Taylor, S.L. (2009). Distribution of Intact Peanut Protein and Digestion-Resistant Ara h 2 Peptide in Human Serum and Saliva. J. Allergy Clin. Immunol. *123*, S268–S268.

Berin, M.C., Kiliaan, A.J., Yang, P.C., Groot, J.A., Kitamura, Y., and Perdue, M.H. (1998). The influence of mast cells on pathways of transepithelial antigen transport in rat intestine. J. IImunology *161*, 2561–2566.

Chehade, M., and Mayer, L. (2005). Oral tolerance and its relation to food hypersensitivities. J. Allergy Clin. Immunol. *115*, 3–12.

Cheriton, D.R., Tran, N.H., Qiao, R., Xin, L., Chen, X., Liu, C., Zhang, X., Shan, B., Ghodsi, A., and Li, M. (2019). Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. Nat. Methods *16*, 63.

Clarke, M., Kilburn, S., Hourihane, J.O., Dean, K., Warner, J., and Dean, T. (1998). Serological characteristics of peanut allergy. Clin. Exp. Allergy *28*, 1251–1257.

Domon, B., and Aebersold, R. (2010). Options and considerations when selecting a quantitative proteomics strategy. Nat. Biotechnol. *28*, 710–721.

Duncan, M.W., Yergey, A.L., and Patterson, S.D. (2009). Quantifying proteins by mass spectrometry: The selectivity of SRM is only part of the problem. Proteomics *9*, 1124–1127.

Fortin, T., Salvador, A., Charrier, J.P., Lenz, C., Bettsworth, F., Lacoux, X., Choquet-Kastylevsky, G., and Lemoine, J. (2009). Multiple Reaction Monitoring Cubed for Protein Quantification at the Low Nanogram/Milliliter Level in Nondepleted Human Serum. Anal. Chem. *81*, 9343–9352.

Gallien, S., Duriez, E., Demeure, K., and Domon, B. (2013). Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. J. Proteomics *81*, 148–158.

Hands, C.M., Sayers, R.L., Nitride, C., Gethings, L.A., and Mills, E.N.C. (2020). A multiple reaction monitoring method for determining peanut (Arachis hypogea) allergens in serum using quadrupole and time-of-flight mass spectrometry. Anal. Bioanal. Chem.

Hourihane, J.O., Allen, K.J., Shreffler, W.G., Dunngalvin, G., Nordlee, J.A., Zurzolo, G.A., Dunngalvin, A., Gurrin, L.C., Baumert, J.L., and Taylor, S.L. (2017). Peanut Allergen Threshold Study (PATS): Novel single-dose oral food challenge study to validate eliciting doses in children with peanut allergy. J. Allergy Clin. Immunol. *139*, 1583–1590.

JanssenDuijghuijsen, L., van Norren, K., Grefte, S., Koppelman, S., Lenaerts, K., Keijer, J., Witkamp, R., and Wichers, H. (2017a). Endurance Exercise Increases Intestinal Uptake of the Peanut Allergen Ara h 6 after Peanut Consumption in Humans. Nutrients *9*, 84.

JanssenDuijghuijsen, L.M., Wichers, H.J., van Norren, K., Keijer, J., Baumert, J.L., de Jong, G.A., Witkamp, R.F., and Koppelman, S.J. (2017b). Detection of peanut allergen in human blood after consumption of peanuts is skewed by endogenous immunoglobulins. J. Immunol. Methods *440*, 52–57.

Koppelman, S., Bruijnzeel-Koomen, C., Hessing, M., and de Jongh, H. (1999). Heat-induced conformational changes of Ara h 1, a major peanut allergen, do not affect its allergenic properties. J. Biol. Chem. *274*, 4770–4777.

Koppelman, S.J., Wensing, M., Ertmann, M., Knulst, A.C., and Knol, E.F. (2004). Relevance of Ara h 1, Ara h 2 and Ara h 3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil-histamine release and intracutaneous testing: Ara h 2 is the most important peanut allergen. Clin. Exp. Allergy *34*, 583–590.

Koppelman, S.J., Jayasena, S., Luykx, D., Schepens, E., Apostolovic, D., de Jong, G.A.H., Isleib, T.G., Nordlee, J., Baumert, J., Taylor, S.L., et al. (2016). Allergenicity attributes of different peanut market types. Food Chem. Toxicol. *91*, 82–90.

Korte, R., and Brockmeyer, J. (2016). MRM3-based LC-MS multi-method for the detection and quantification of nut allergens. Anal. Bioanal. Chem. *408*, 7845–7855.

Korte, R., Lepski, S., and Brockmeyer, J. (2016a). Comprehensive peptide marker identification for the detection of multiple nut allergens using a non-targeted LC-HRMS multi-method. Anal. Bioanal. Chem. 3059–3069.

Korte, R., Monneuse, J.-M., Gemrot, E., Metton, I., Humpf, H.-U., and Brockmeyer, J. (2016b). New High-Performance Liquid Chromatography Coupled Mass Spectrometry Method for the Detection of Lobster and Shrimp Allergens in Food Samples via Multiple Reaction Monitoring and Multiple Reaction Monitoring Cubed. J. Agric. Food Chem. *64*, 6219–6227.

Kumar Dey, K., Wang, H., Niu, M., Bai, B., Wang, X., Li, Y., Cho, J.-H., Tan, H., Mishra, A., High, A.A., et al. (2019). Deep undepleted human serum proteome profiling toward biomarker discovery for Alzheimer's disease. Clin Proteom *16*, 16.

Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring

for quantitative proteomics: a tutorial. Mol. Syst. Biol. *4*.

Maclean, B., Tomazela, D.M., Shulman, N., Chambers, M., Finney, G.L., Frewen, B., Kern, R., Tabb, D.L., Liebler, D.C., and Maccoss, M.J. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinforma. Appl. NOTE *26*, 966–968.

Mose, A.P., Mortz, E., Skov, P.S., Mortz, C.G., Eller, E., Sprogøe, U., Barington, T., and Bindslev-Jensen, C. (2019). The quest for ingested peanut protein in human serum. Allergy all.14109.

Nwaru, B.I., Hickstein, L., Panesar, S.S., Muraro, A., Werfel, T., Cardona, V., Dubois, A.E.J., Halken, S., Hoffmann-Sommergruber, K., Poulsen, L.K., et al. (2014). The epidemiology of food allergy in Europe: A systematic review and meta-analysis. Allergy Eur. J. Allergy Clin. Immunol. *69*, 62–75.

Percy, A.J., Chambers, A.G., Yang, J., and Borchers, C.H. (2013). Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma. Proteomics *13*, 2202–2215.

Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., and Coon, J.J. (2012). Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics.

Picotti, P., and Aebersold, R. (2012). Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. Nat. Methods *9*, 555–566.

Rauniyar, N. (2015). Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. Int. J. Mol. Sci. *16*, 28566–28581.

Reitsma, M., Westerhout, J., Wichers, H.J., Wortelboer, H.M., and Verhoeckx, K.C.M. (2014). Protein transport across the small intestine in food allergy. Mol. Nutr. Food Res *58*, 194–205.

Ronsein, G.E., Pamir, N., von Haller, P.D., Kim, D.S., Oda, M.N., Jarvik, G.P., Vaisar, T., and Heinecke, J.W. (2015). Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. J. Proteomics *113*, 388–399.

Samadi, N., Klems, M., and Untersmayr, E. (2018). The role of gastrointestinal permeability in food allergy. Ann. Allergy, Asthma Immunol.

Schmitt, D.A., Nesbit, J.B., Hurlburt, B.K., Cheng, H., and Maleki, S.J. (2010). Processing can alter the properties of peanut extract preparations. J. Agric. Food Chem. *58*, 1138–1143.

Schocker, F., Baumert, J., Kull, S., Petersen, A., Becker, W.-M., and Jappe, U. (2016). Prospective investigation on the transfer of Ara h 2, the most potent peanut allergen, in human breast milk. Pediatr. Allergy Immunol. *27*, 348–355.

Sherman, J., McKay, M.J., Ashman, K., and Molloy, M.P. (2009). How specific is my

SRM?: The issue of precursor and product ion redundancy. Proteomics *9*, 1120–1123.

Sicherer, S., Muñoz-Furlong, A., Godbold, J., and Sampson, H. (2010). US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immunol. *125*, 1322–1326.

Stone, K.D., Prussin, C., and Metcalfe, D.D. (2010). IgE, mast cells, basophils, and eosinophils. J. Allergy Clin. Immunol. *125*.

Taylor, S.L., and Baumert, J.L. (2012). Food Allergies and Intolerances. In Present Knowledge in Nutrition, (Wiley-Blackwell), pp. 1222–1235.

Venter, C., Hasan Arshad, S., Grundy, J., Pereira, B., Bernie Clayton, C., Voigt, K., Higgins, B., and Dean, T. (2010). Time trends in the prevalence of peanut allergy: three cohorts of children from the same geographical location in the UK. Allergy *65*, 103–108.

CHAPTER 4: INTERMOLECULAR ASSOCIATION OF PEANUT ALLERGENS AND THE IMPACT OF THERMAL PROCESSING

ABSTRACT

The expression patterns and sequences of peanut proteins are well studied, however, the intermolecular arrangements of seed storage proteins in plants are largely unstudied. These protein arrangements may have implications to food allergy sensitization, elicitation, and protein behavior during gastrointestinal digestion.

To study peanut protein intermolecular arrangements, we used size-exclusion chromatography (SEC) with offline mass spectrometry (MS) to detect and quantify peptides. Raw and roasted peanuts (*Arachis hypogaea*) were extracted in conditions designed to maintain native (TBS) or denatured protein conformations (GuHCl). Native MW of proteins were determined using Gaussian modelling allowing for the analysis of isoforms and complexes. Bottom-up analysis of SEC fractions, and Gaussian modelling for native MW determination, proved an effective offline separation technique. Most proteins elute as expected, however we noted several exceptions. Ara h 3 exists as an oligomer, but at an equilibrium between trimers or hexamers. Some of the prolamins (Ara h 2 and 6) eluted at higher MWs than expected, alongside the cupins (Ara h 1 and 3). We additionally characterized post-translationally processed Ara h 1 into two, distinct and unassociated forms

The analytical workflow presented, and automated assignment of MW using Gaussian fitting, allows for the study of protein conformation in complex protein mixtures.

INTRODUCTION

The sequences and expression patterns of individual allergenic peanut proteins have been well-studied (Burks et al., 1998; Chruszcz et al., 2011; Koppelman et al., 2005, 2016). However, the intermolecular associations between peanut seed storage proteins, which are deposited in seed storage vacuoles, a high-density and low water environment, are not well-understood (Müntz, 1998). These molecular arrangements of peanut allergens may have implications for food allergen sensitization and behavior during gastric digestion. Furthermore, the impact of food processing on these molecular arrangements is unknown.

Peanut allergens have been extensively studied, due to their severe potency and high prevalence rates of sensitization (van Boxtel et al., 2008; Koppelman et al., 2004; Gupta et al., 2019; Maleki et al., 2000a). The major peanut allergens (Ara h 1, 2, 3, and 6) are seed storage proteins, which provide a reservoir of amino acids for use during plant germination and growth (Shewry et al., 1995). During seed storage protein translation the proteins are ultimately deposited in the storage vacuoles where they form protein bodies, however, the arrangements of seed storage proteins within these protein bodies is unknown (Shewry et al., 1995). Therefore, the aim of our study was to assess arrangements of these peanut seed storage proteins.

Ara h 1 and 3 are globulin proteins belonging to the cupin superfamily (Burks et al., 1995b; Kleber-Janke et al., 1999; Rabjohn et al., 1999;). Ara h 1 is classified as a 7S vicilin with a monomeric molecular weight of 63.5 kDa that associates into a 180 kDa trimer (Maleki et al., 2000b; van Boxtel et al., 2006). The 7S vicilin (Ara h 1) is

translated as a pre-pro-protein, which undergoes two post-translational proteolytic cleavage events to produce mature Ara h 1. In this process, the signal peptide (~25 amino acids) is cleaved first, followed by a secondary cleavage of the subsequent 53 – 59 amino acids of the N-terminal sequence (de Jong et al., 1998; Wichers et al., 2004). As a result, mature Ara h 1 exists as a truncated sequence compared to the full length sequence (Burks et al., 1995b). Ara h 3, the 11S legumin, is a 60 kDa monomer that associates as a hexamer (360 kDa) composed of two homo-trimers (Koppelman et al., 2003; Jin et al., 2009). Recent elucidation of the peanut genome has allowed characterization of the many Ara h 3 isoforms in peanuts (Bertioli et al., 2019). Ara h 3 contains acidic and basic subunits which remain associated until they undergo gastrointestinal digestion (Marsh et al., 2008). It appears the acidic subunit is more allergenic than the basic subunit (Marsh et al., 2008).

The 2S albumins, Ara h 2 and 6, belong to the prolamin superfamily and are highly homologous proteins sharing 59% sequence identity and a common conserved disulfide stabilized alpha helical protein core structure (Koppelman et al., 2005; Moreno and Clemente, 2008). These structural characteristics confer thermal stability and resistance towards gastric digestion, contributing to the severe allergenic potency of Ara h 2 and 6 (Blanc et al., 2009; Flinterman et al., 2007; Klemans et al., 2013).

Raw peanuts typically undergo thermal processing (e.g. roasting, boiling, frying, blanching) prior to consumption, likely influencing structural arrangement(s) and conformation (Beyer et al., 2001; Guillon et al., 2016; Zhang et al., 2016). Roasting, in particular, induces protein denaturation and aggregation causing reduced protein solubility in aqueous buffer systems (e.g. TBS, PBS etc.) (Kopper et al., 2005; Schmitt et

al., 2010). Reports of increased allergenicity have been attributed to thermal processing due to increased epitope and IgE binding sites (Beyer et al., 2001; Mondoulet et al., 2005). Furthermore, the formation of oligomeric structures due to thermal processing, may also contribute to reported increases in allergenicity (Maleki et al., 2000a).

Peanut allergy is one of the most prevalent food allergies, with current estimates at 2% of children and 0.6% of adults diagnosed with peanut allergy, and, nearly all individuals are sensitized to multiple peanut proteins (Ara h 1, 2, 3, or 6) (Gupta et al., 2011; Sicherer et al., 2010). The cumulative dose of allergenic protein dictates individual sensitization and overall reaction severity (Ballmer-Weber et al., 2015; Turner et al., 2016). Individuals who are only sensitized to Ara h 2 and 6, reported more severe reactions than those sensitized to Ara h 1 and 3, and as such, the prolamins (Ara h 2, 6) are predictors of reaction severity (Koppelman et al., 2004; Kukkonen et al., 2015). The molecular arrangement(s) and initial presentation of these allergenic proteins to the immune system are suspected to influence sensitization and reaction severity, therefore understanding the molecular arrangements of these peanut proteins is important.

Our aim was to study the intermolecular associations between peanut proteins using size-exclusion chromatography (SEC) coupled with offline discovery liquid chromatography tandem mass spectrometry (LC-MS/MS) to detect and quantify peptides in individual size fractions. SEC is an established technique used for separation of proteins and protein complexes based on their molecular radii in which large proteins elute rapidly and smaller proteins elute more slowly (Fekete et al., 2014). We analyzed raw and roasted peanuts in the presence or absence of chaotropic reagents to solubilize heavily aggregated material, which is absent from most studies. The mass spectrometry

(MS) data were analyzed against our peanut database derived from the peanut genome (Bertioli et al., 2019), allowing us to distinguish between peanut protein isoforms. The combination of SEC with offline LC-MS/MS offers a data-rich method allowing us to describe detailed intermolecular associations of seed storage proteins.

MATERIALS AND METHODS

I. Reagents

All reagents for protein preparation (defatting, extraction) and size exclusion chromatography were of analytical reagent grade. All reagents for MS sample preparation and analyses were of MS grade.

The following reagents were used during gel filtration protein preparations: hexane, deionized distilled water, tris-buffered saline (Tris-HCl, NaCl), GuHCl (GuHCl, Tris-HCl, EDTA), acetone.

The following reagents were used during LC-MS/MS sample preparation and data acquisition: water, acetonitrile (ACN), ammonium bicarbonate (ABC), dithiothreitol (DTT), iodoacetamide (IAA), trypsin, formic acid (FA), methanol (MeOH), and acetic acid.

II. Sample Preparation

Raw peanuts (var. Runner) were obtained from the Golden Peanut Company (Alpharetta, GA). Peanuts (25 g) were roasted in a conventional oven (Groen Combination oven, Model No. CC10-E) at 160°C for 13.5 minutes. All peanuts (raw,

roasted) were ground into a powder under liquid nitrogen in a freezer mill (Spex 6850 CentriPrep Freezer/Mill (Metuchen, NJ) followed by manual grinding with a mortar and pestle until a fine powder was achieved. Ground peanut powder was defatted twice in hexane (1:20 w/v) and allowed to dry overnight at room temperature.

III. Protein Extraction

Peanut proteins were extracted (1:25 w/v) by mixing 0.5 g defatted peanut powder with 12.5 mL of extraction buffer either (a) tris buffered saline (TBS, 50 mM Tris-HCl, 150 mM NaCl, pH 7.4) or (b) guanidine hydrochloride (GuHCl, 5 M GuHCl, 50 mM Tris, 5 mM EDTA) in a sonicating water bath for 20 minutes at room temperature (20°C). Extracts were clarified by centrifugation (3500 x g, 10°C, 10 minutes). The supernatant(s) were removed and dialyzed overnight using 3.5 kDa molecular weight cut off (MWCO) dialysis tubing (Snakeskin dialysis tubing, 3.5K MWCO, 16 mm ID; Prod. No. 88424, ThermoScientific™, Rockford, IL, USA) against the appropriate extraction buffer at room temperature. The dialysates were filter sterilized using 0.2 µm syringe filters (Fisherbrand, Cat. No. 09-719C). Protein concentrations were determined by 2D Quant protein assay (GE Healthcare, 80648356). The protein fractions extracted in TBS represent soluble proteins, whereas proteins extracted in GuHCl, a chaotropic buffer, represent the total peanut protein composition.

IV. Size Exclusion Chromatography (SEC)

Peanut protein solutions (1.5 mg/mL TBS extracts, 5 mg/mL GuHCl extracts) were applied to a Superdex 200 column (16 x 600 mm) (HiLoad 16/600 Superdex 200 prep grade) attached to the AKTA Avant 25 chromatography system (GE Healthcare Life Sciences NJ, USA) and monitored at 214 and 280 nm. For all extracts, the column was equilibrated and eluted with TBS at a 1 mL/min flow rate. All SEC extracts were analyzed in duplicate. Fractions were collected by volume (4 mL). The SEC column was calibrated by using the molecular weight markers: cytochrome c (12.4 kDa), carbonic anhydrase (29 kDa), albumin (66 kDa), alcohol dehydrogenase (150 kDa), beta-amylase (200 kDa), and blue dextran (2000 kDa) (Gel Filtration Molecular Weight Markers Kit for Molecular Weights 12,000–200,000 Da; MWGF200, Sigma-Aldrich, St. Louis, MO, USA). Blue dextran was used to determine column void volume.

V. Sample Preparation for LC-MS/MS

Following gel filtration separation, and protein quantification (by 2D quant assay (GE Healthcare, ID)), aliquots (200 µL – determined to be less than 30 µg), the fractions obtained from TBS extracts (raw and roasted peanut) were precipitated with acetone. Aliquots of fractions obtained from GuHCl extracts were also precipitated with acetone, but the aliquot volume was limited by the maximum fraction concentration found, such that the maximum total protein precipitated was 30 µg. As such, 45 µL – raw peanut and 30 µL – roasted peanut were used. Precipitation was conducted by adding four volumes of -20°C acetone to each aliquoted fraction(s) and incubated for 2 hours at -80°C. After incubation, samples were centrifuged (16,000 x g, 10 minutes, 5°C) and the supernatant

was removed. The precipitated pellet was washed twice, with 7.5 volumes of chilled (-20°C) 4:1 acetone:water, followed by centrifugation at (16,000 x g, 10 minutes, 5°C) and dried.

Acetone precipitated protein fractions were prepared for enzymatic digestion using a modified in-solution trypsin digestion protocol (In-Solution Tryptic Digestion and Guanidination Kit, ThermoScientific, Rockford, IL, USA). Protein pellets were re-solubilized in 31.5 µL of 15% (v/v) acetonitrile followed by an addition of 45 µL of 50 mM ammonium bicarbonate. Samples were then reduced by the addition of 4.5 µL of 100 mM dithiothreitol and heated at 95°C for 5 minutes. Following reduction, 9 µL of 100 mM iodoacetamide was added to and incubated in the dark at room temperature for 20 minutes. Proteins were digested with 300 ng trypsin (100 ng/µL in 5 mM acetic acid) at 37°C for 3 hours followed by a second addition of 300 ng trypsin with overnight incubation at 30°C. Peptide digests were desalted using C-18 spin columns (Pierce C-18 spin columns, ThermoScientific, Rockford, IL, USA) according to manufacturers' instructions, eluted in 50% (v/v) acetonitrile, and dried under a vacuum by centrifugal evaporation (Jouan RC-10.10; RCT-90; Winchester, VA, USA). Peptides were re-solubilized in 50 µL of 5% (v/v) acetonitrile, 0.1% (v/v) formic acid prior to injection on LC-MS/MS.

VI. LC-MS/MS Separation and Acquisition

Tryptic digests (2 µL injection) were chromatographically separated using an UltiMate 3000RSL® liquid chromatography (UPLC) system (Thermo Scientific™) equipped with a Hypersil Gold C18 1.9 µm, 100 x 1 mm reversed phase column (Thermo

Scientific™) with a pre-column (20 x 2.1 mm reversed phase, 1.9 µm, Thermo Scientific™) set at 35°C. Mobile phase A contained 0.1% (v/v) formic acid in water and mobile phase B contained 0.1% (v/v) formic acid in acetonitrile. Peptides were separated using a linear gradient of 2 – 40% mobile phase B over 70 minutes at a flow rate of 60 µL/min. Data dependent acquisitions (DDA) were performed on a Thermo Q Exactive Plus™ Hybrid Quadrupole-Orbitrap™ mass spectrometer (Thermo Scientific™) operating in positive ion mode using a top-20 method.

Survey scan mass spectra (400 – 1400 m/z) were acquired at a nominal resolution of 70,000 FWHM (200 m/z) and an AGC target of 3e6. Fragmentation spectra were acquired at a nominal resolution of 70,000 FWHM with normalized collisional energy (NCE) set at 27 and dynamic exclusion of 20 seconds, AGC target of 1e5, and a maximum inject time of 240 ms. The electrospray ionization settings were as follows: sheath gas 15 AU, auxiliary gas flow 4.2 AU, spray voltage 3500 kV, capillary temperature 320°C, S-lens RF level 60.

VII. Data Analysis

Acquired mass spectral data were analyzed using PEAKS version 8.5 (Bioinformatics Solutions Inc., Waterloo, Ontario, Canada). Identification and label-free quantification of proteins were performed using a custom database derived from the peanut genome (Bertioli et al., 2019; Marsh et al., 2020). Peptides were selected to represent allergen protein families as described in Marsh et al., (2020). Data were analyzed using the following criteria: 5 ppm parent ion mass tolerance, FDR 1%, parent ion charge states 2, 3, 4+, and normalized to original extract volume. For each allergen

protein family, three selected peanut peptides were chosen due to their presence in multiple isoforms. Peptides unique to allergen isoforms were also chosen for individual isoform quantification. Data were normalized to original SEC sample volume. Data presented in this manuscript were analyzed using allergen protein families, unless otherwise noted. In order to determine the observed peptide MW, we applied Gaussian curve fitting (nonlinear regression) to determine the mean elution time. We then calculated the MW based upon the determined elution time using the standard curve generated from the MW markers.

RESULTS

I. Determination of Mean Protein Elution with Gaussian Distribution Fitting

In order to determine the observed (e.g. calculated) protein MW, we applied a Gaussian curve fit to the elution profile of identified peptides corresponding to specific peanut proteins using PRISM. Generally, chromatographic peaks exhibit symmetrical Gaussian shapes (Le Vent 1995; Uversky 1993). This distribution is a theoretical curve fit to the SEC chromatographic data, which determines the mean elution time, standard deviation, and amplitude for individual peptides along the Gaussian fitted curve. We were further able to determine the molecular weight of the identified peptides with excellent precision between observed MWs for each protein. However, since we do not know the true value of our protein MWs, we are unable to determine the accuracy of our observed MWs. Using the mean protein elution time, we are able to determine the observed (e.g. calculated) protein MWs by inference from our standard curve.

II. Peanut Protein Elution

By pairing offline SEC with LC-MS/MS, we identified a majority of the peanut peptides present within our genomically derived peanut database. The observed elution time(s) for these peanut specific peptides (Table 4-1), determined by a Gaussian fitted curve, were used to calculate the protein MWs and hereafter referred to as observed MW. The results presented describe the elution profile of raw peanuts extracted in TBS (saline extractions), unless otherwise noted. In the GuHCl extractions, nearly all the protein eluted in the aggregate fraction(s) with a large MW unit.

The major allergenic peanut proteins, Ara h 1, 2, 3, and 6, eluted predominantly as expected and were present as either monomers or large oligomers (Figure 4-1). The cupins, Ara h 1 and 3, were predominantly present as oligomers whereas Ara h 2 and 6, the majorly allergenic prolamins, were present as monomers. The minor peanut allergens, Ara h 7 and 8 eluted as monomers whereas as Ara h 9, 10, and 11 eluted as octamers and most likely associated with the cupins (Figure 4-1). Peanut peptides derived from the same parent protein (e.g. Ara h 2) should elute identically resulting in the same calculated MW. However, in several instances, we observed unexpected elution profiles due to differences in individual protein isoforms (e.g. Ara h 3), post-translational cleavage of proteins (e.g. Ara h 1), and intermolecular associations between proteins, each of which are described in further detail in the following sections.

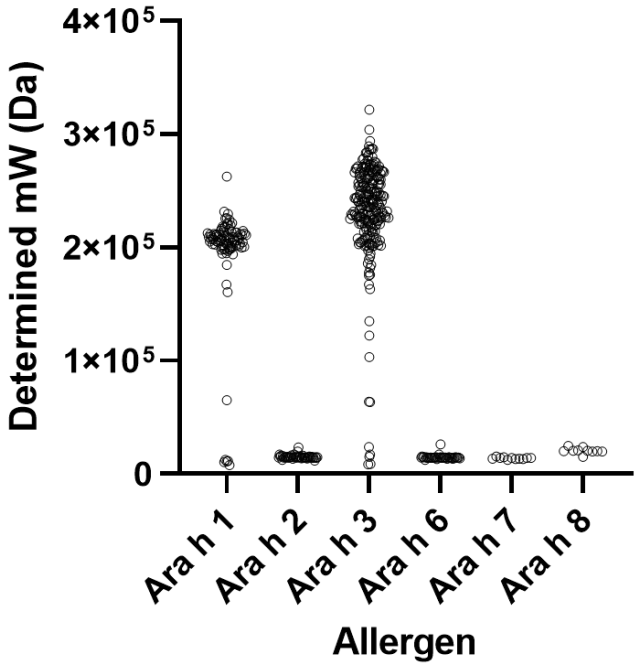


Figure 4-1. Observed MWs of TBS extracted peanut proteins with each point representing a single peptide

Accession	Peptide sequence(s)	m/z	z	Calculated mW (kDa)	Observed mW (kDa)
Ara h 1.01/02	GTGNLELVAVR	564.8	2	62	210
	NTLEAAFNAEFNEIR	869.9	2		
	SFNLDEGHALR	629.8	2		
Ara h 2.01/02	ANLRPC(+57.02)EQHLMQK	542.3	3	17	14
	C(+57.02)C(+57.02)NELNEFENNQR	576.2	3		
	NLPQQC(+57.02)GLR	543.3	2		
Ara h 6.01/02	C(+57.02)DLDVSGGR	489.7	2	15	14
	ELMNLPQQC(+57.02)NFR	775.4	2		
	VNLKPC(+57.02)EQHIMQR	826.9	2		
Ara h 3.01/11	EGQILLVPQNFAVGK	807.0	2	71	250
	FYLAGNPEEEHPETQQQQPQTR	876.4	3		
	IESQGGITETWNSNHPELR	1084.5	2		
Ara h 3.04/05/10/13/17/20	FNLAGNHEQEFLR	787.9	2	58	240
	RPFYSNAPQEIFIQQGR	684.4	3		
	SPDIYNPQAGSLK	695.4	2		
Ara h 3.07/16	AQSENYEYLAFK	731.9	2	57	270
	SSNPDIYNPQAGSLR	809.9	2		
	VYDEELQEGHVLVVPQNFAVAAK	1278.2	2		
Ara h 3.02/12	AGSDAFDWVAIK[1]	640.3	2	49	230
	GVMEIVVTGC(+57.02)R	610.8	2		
	TSDNPIINTLAGELSLVR[1]	957.0	2		
Ara h 3.03/06/19	LGLSAEYGSIHR	434.9	3	49	180
	SQSEHFLYVAFK	728.4	2		
	TVNELDLPILNR	698.9	2		
Ara h 3.08/15	LPILADLQLSAER	480.3	3	48	230
	NIVMVEGGLDVVRPEPGSR	506.8	4		
	FYIAGNTEDEHGEGGR	876.4	2		
Ara h 3.09/14	EGQILIVPQQFVVAK	835.0	2	44	240
	GLLLPHYINAPR	455.3	3		
	NDQFQC(+57.02)VGVSALR	747.4	2		
Ara h 3.18/-	EIVQNLR	436.3	2	50	220
	IDSEGGFIETWNPK	796.9	2		
	QEQEFLQYQHQHGGPR	496.2	4		

Table 4-1. Elution of raw peanut proteins extracted in TBS detected by shared peptides. Calculated MWs were determined by Expasy. Observed MWs determined by Gaussian modeling.

III. Differences in Elution of Ara h 3 Isoforms

We identified 21 individual Ara h 3 isoforms in our genomically derived peanut protein database, and grouped isoforms into families based on shared identical peptide sequences (Bertioli et al., 2019; Marsh et al., publication in progress). A table of shared peptides among Ara h 3 isoforms are listed in Table 4-1. As shown in Figure 4-2, we demonstrate Ara h 3 exists largely as an oligomer (178 – 280 kDa) in aqueously extracted raw peanuts (Figure 4-3 and Table 4-2). We observed Ara h 3 exists as a mixture of trimer and hexamer. Additionally, Ara h 3 isoforms appeared to distribute between trimer and hexamer differently, as evidenced by calculated MWs that could not be explained by primary sequence alone (Figure 4-3).

The expected monomeric MW of each Ara h 3 isoform was determined using the known isoform sequence, removing known signal peptides if present, and molecular weight determination using Expasy, compute pI/MW tool which calculated the MW based on entered peptide sequence(s) (https://web.expasy.org/compute_pi/) (ExPASy: SIB Bioinformatics Resource Portal). The number of subunits was calculated using the ratio of observed MW versus expected MW determined by SEC elution and LC-MS/MS. As an example, the expected monomeric MW of Ara h 3.01/11 is 71 kDa, and was observed to aggregate into a trimer (246 kDa) (Figure 4-2). Contrastingly, Ara h 3.09/14, with an expected monomeric MW of 45 kDa, was determined to be existing as a hexamer (280 kDa). The determined number of subunits was between 3 and 6 subunits with many of the Ara h 3 isoforms calculated to exist between four or five subunits, assuming the aggregates are composed of identical monomers. We suspect these aggregates may exist

as a mixture of peanut proteins, such as Ara h 3 associating with Ara h 6. This is suggestive of a mixed ratio of oligomeric states, between the trimeric and hexameric arrangements, and could exist in an equimolar state between the oligomeric states. No differences in MW were determined for isoforms of Ara h 1, 2, and 6.

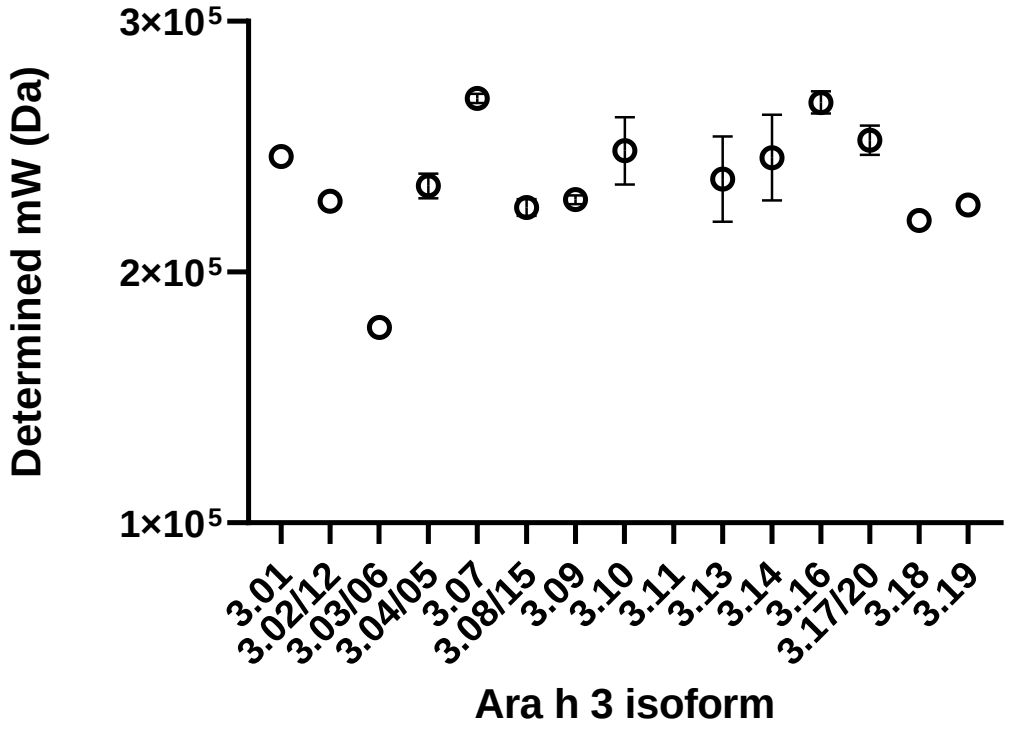


Figure 4-2. Observed MWs of TBS extracted Ara h 3 proteins detected by unique peptides.

Protein	Peptide	m/z	charge	monomeric MW (Expasy)	Gaussian determined MW (kDa)	Calculated number of subunits
Arah3_1_02	R.ISSANSLTFPILR.W	709.90	2	71	250	3.5
Arah3_2_02	K.GVMEIVVTGC(+57.02)R.A	610.81	2	49	230	4.7
Arah3_2_02	K.AGSDAFDWVAIK.T	640.32	2	49	230	4.7
Arah3_2_02	K.TSDNPIINTLAGELSLVR.A	957.02	2	49	230	4.7
Arah3_3_06	R.GLSILVPAER.R	527.82	2	49	180	3.7
Arah3_4_06	K.FFVPPSQQSLR.A	653.35	2	58	240	4.1
Arah3_5_06	K.SQSDNFEYVAFK.T	717.83	2	58	230	4.0
Arah3_5_06	R.GENESEEEGAIVTVK.G	795.88	2	58	240	4.1
Arah3_7_06	R.GEEQENEGNNIFSGFAQEFLQHAFQVDR.E	1080.83	3	57	270	4.7
Arah3_7_06	R.QGGEENEC(+57.02)QFQR.L	741.31	2	57	270	4.7
Arah3_7_06	R.ILNPDEEDESSR.S	702.32	2	57	270	4.7
Arah3_9_06	R.LTAEEAINLK.K	551.31	2	44	230	5.2
Arah3_9_06	K.LVALEPSK.R	428.76	2	44	230	5.2
Arah3_9_06	K.TVAESLGIDMGIAGK.V	731.38	2	44	230	5.2
Arah3_10_06	K.FFVPPSQQSPR.A	645.34	2	58	230	4.0
Arah3_10_06	R.GENESEEEGAIVTVR.G	809.88	2	58	250	4.3
Arah3_10_06	K.TDSRPSIANLAGENSVIDNLPEEVVANSYGLPR.E	1166.59	3	58	270	4.7
Arah3_13_16	K.SPDEEEEYDEDEYAEEER.Q	1131.92	2	58	210	3.6
Arah3_13_16	K.FFVPPFQQSPR.A	675.35	2	58	240	4.1
Arah3_13_16	R.AGQEQENEGGNIFSGFTSEFLAQAFQVDDR.Q	1097.83	3	58	270	4.7
Arah3_14_16	K.LVALEPTK.R	435.77	2	44	230	5.2
Arah3_14_16	R.LTAEEAISLK.K	537.81	2	44	230	5.2
Arah3_14_16	K.TVAESLDIDMGIAGK.V	760.39	2	44	280	6.4
Arah3_15_16	R.FYIAGNTEDEHGEGGR.E	876.59	2	48	220	4.6
Arah3_15_16	K.NIVMVEGGLDVVRPEPGSR.A	56.77	4	48	230	4.8
Arah3_15_16	K.LPILADLQLSAER.G	480.28	3	48	230	4.8
Arah3_16_16	R.FYLAGNQEQEFLR.Y	807.9	2	57	260	4.6
Arah3_16_16	R.FQVGQDDPSQQQQDSHQK.V	700.65	3	57	270	4.7
Arah3_16_16	R.ILSPDEEDESSR.S	688.81	2	57	270	4.7
Arah3_17_16	R.QILQNLR.G	442.77	2	58	240	4.1
Arah3_17_16	R.GENESDEQGAIVTVR.G	802.38	2	58	250	4.3
Arah3_18_16	R.IDSEGGFIETWNPK.S	796.88	2	50	220	4.4
Arah3_19_16	K.GGLSILVPPEWR.Q	662.37	2	49	230	4.7
Arah3_20_16	R.VFDEELQEGHVLVVPQNFAVAGK.S	1263.16	2	58	260	4.5

Table 4-2. Unique Ara h 3 peptides for Gaussian MW determinations

IV. Post Translational Cleavage of Ara h 1 *N*-Terminal Sequence

The Ara h 1 gene product is processed into two forms, the N-terminal region (pre-pro) protein and the mature Ara h 1 product (Hurlburt et al., 2014; Wichers et al., 2004). We observed evidence of post-translational processing of Ara h 1 by identifying N-terminal peptides present as distinct monomeric molecular entities from our raw peanut extracts (Figure 4-3). Peptides (Table 4-3; Figure 4-4) associated with the N-terminal region were identified at 12 kDa, independent from the mature Ara h 1 peptides as a distinct molecule (Figure 4-3). This type of processing has been observed in certain tree nuts and other legumes (Downs et al., 2014; Marcus et al., 1999).

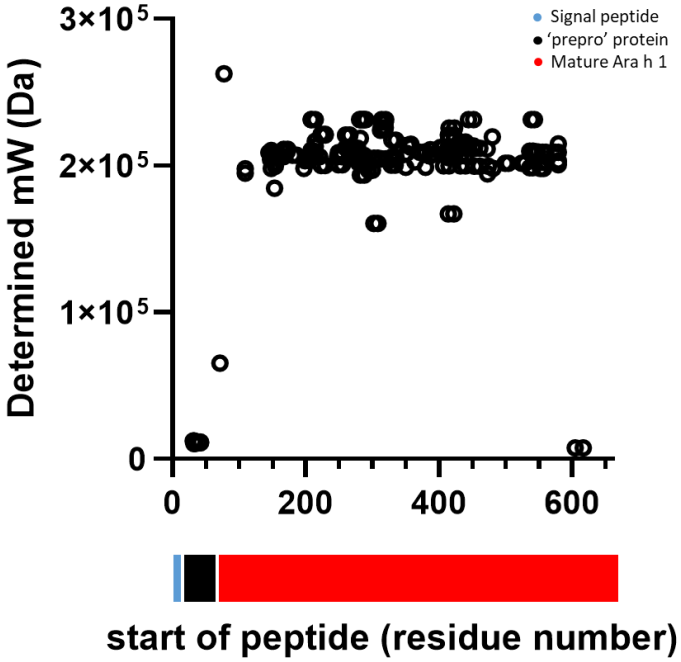


Figure 4-3. Determined MW of Ara h 1 peptides plotted by beginning residue of amino acids.

Accession	Peptide sequence(s)	m/z	z	Calculated MW (kDa)	Observed MW (kDa)
Ara h 1.01/02 (unique peptides)	GTGNLELVAVR	564.82	2	62	210
	NTLEAAFNAEFNEIR	869.92	2		
	SFNLDEGHALR	629.81	2		
Ara h 1.01/02 (N-terminal peptides)	K.TENPC(+57.02)AQR.C	552.27	2	9.2	11
	R.C(+57.02)LQSC(+57.02)QQEPDDLK.Q	812.85	2		
	R.C(+57.02)VYDPR.G	405.18	2		

Table 4-3. Shared and N-terminal peptides used for Ara h 1 protein elution from TBS extractions.

>Arah1_2_19

*MRGRVSPLMLLLGILVLASVSATHA*KSSPYQKKTENPCAQRCLQSCQQEPD
DLKQKACESRCTKLEYDPRCVYDPRGHTGTTNQRSPPGERTRGRQPGDY
DDDRRQPRREEGGRWGPAGPREREREEDWRQPREDWRRPSHQQPRKI
RPEGREGEQEWGTPGSHVREETSRNNPFYFPSRRFSTRYGNQNGRIRVLQ
RFDQRSRQFQNLQNHRIVQIEAKPNTLVLPKHADADNILVIQQGQATVTV
ANGNNRKSFNLDEGHALRIPSGFISYILNRHDNQNLRVAKISMPVNTPGQ
FEDFFPASSRDQSSYLQGFSRNTLEAAFNAEFNEIRRVLLEENAGGEQEER
GQRRWSTRSSENNEGVIVKVSKEHVEELTKHAKSVSKKGSEEEGDITNPIN
LREGEPDLSNNFGKLFEVKPDKKNPQLQDLDMMLTCVEIKEGALMLPHF
NSKAMVIVVVNKGTGNLELVAVRKEQQQRGRREEEEDEDEEEEGSNREV
RRYTARLKEGDVFIMPAAHPVAINASSELHLLGFGINAENNHRIFLAGDKD
NVIDQIEKQAKDLAFPGSGEQVEKLIKNQKESHFVSARPQSQSQSPSSPEK

Figure 4-4. Ara h 1 protein sequence. Italicized letters represent signal peptide; underlined letters represent N-terminal ('prepro') peptide. Calculated MWs were determined by Expasy. Observed MWs determined by Gaussian modeling.

V. Intermolecular Associations between Peanut Proteins

The prolamins, Ara h 2, 6, and 7 exist predominantly as monomers, but we observed the elution of the prolamins in higher MW containing fractions (Figure 4-5). The elution profiles of each individual prolamin peptide demonstrate a small proportion (1.5%) of the prolamin peptides were present in the cupin containing fractions (170 - 320 kDa), suggesting an association between prolamin (Ara h 2, 6, 7) and cupin peanut proteins (Ara h 1, 3). Nearly all of the identified prolamin peptides eluting in higher MW fractions were predominantly cysteine-containing peptides. This is interesting because work with purified proteins has suggested Ara h 2 and 3 are cross-reactive, however in light of our observations, this may not be the case (Bublin et al., 2013; Smit et al., 2015). This elution behavior was also observed in roasted peanut extracts as well; suggesting the co-elution of prolamins with cupins is not solely attributed to thermal processing.

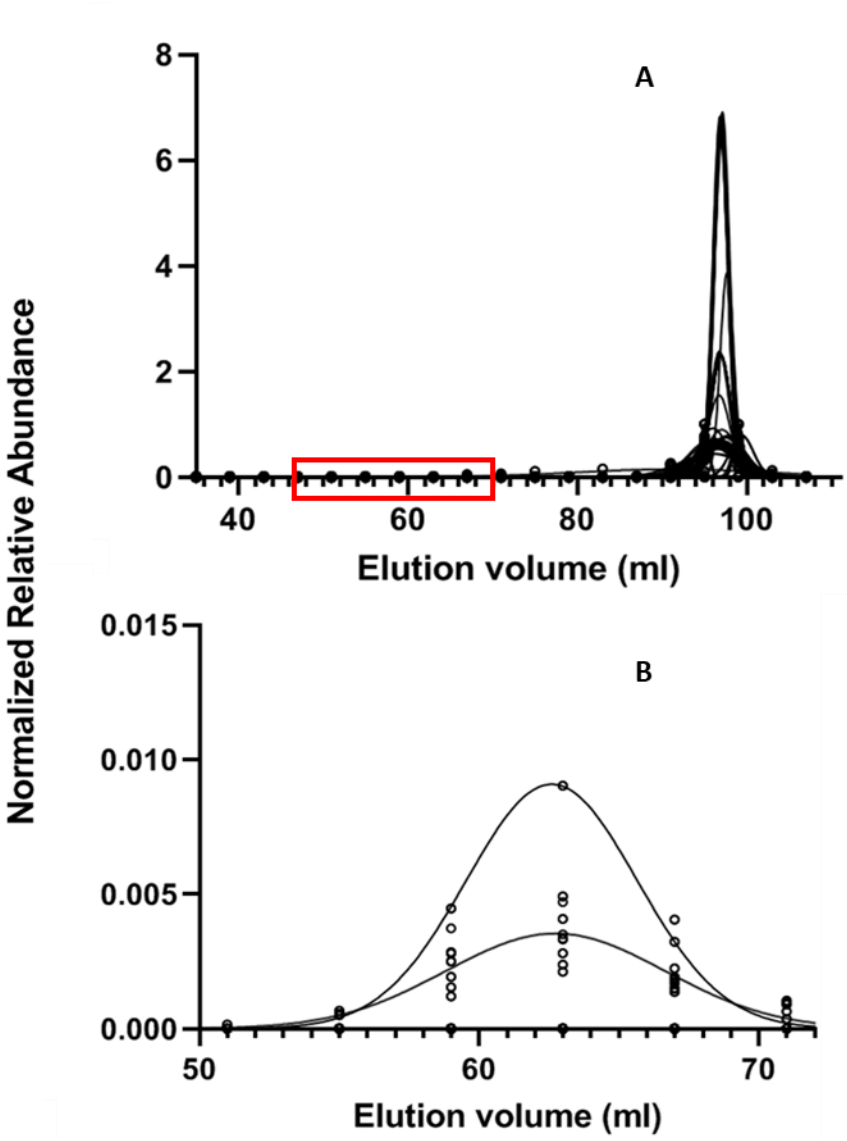


Figure 4-5. Elution profiles of prolamin peptides extracted in TBS (Ara h 2, Ara h 6). Individual points represent a single peptide. Panel (A) shows the entire elution profile. Panel (B) shows the elution profile during a 50-70 minute window, highlighted in Panel (A).

VI. Effect of Thermal Processing

We additionally analyzed a set of roasted peanuts to evaluate any changes in molecular arrangements due to thermal treatment. Roasted peanut were prepared from the same raw peanut starting material and roasted as described in Methods section. We observed a clear increase in calculated MWs for roasted peanuts (Figure 4-6). Thermal processing likely induces aggregation causing changes in the conformation of the peanut proteins. Some proteins, particularly Ara h 1, underwent significant aggregation and were poorly extracted under aqueous conditions. However, when we extracted roasted peanut proteins under chaotropic conditions (e.g. GuHCl), we were able to recover Ara h 1 (data not shown), supporting thermally induced protein aggregation of Ara h 1. As expected, Ara h 2 and 6 were largely present as monomers in roasted peanuts.

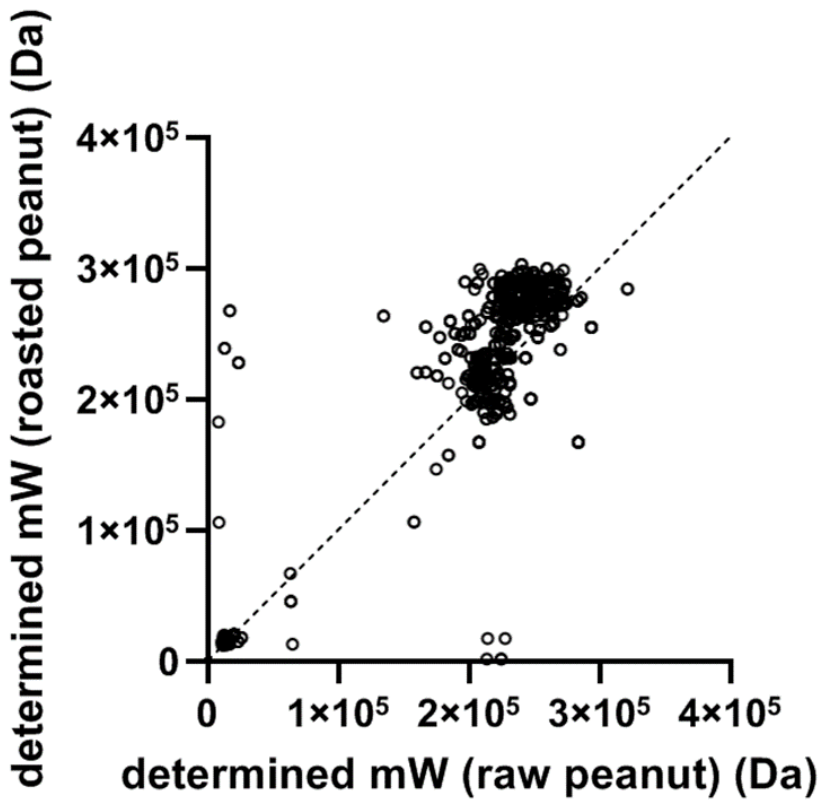


Figure 4-6. Observed peptide MWs in raw and roasted peanut proteins extracted in TBS.

DISCUSSION

Proteins in the foods we eat are often poorly characterized, and the overall protein contents unknown. In many cases, limited sequence information is available for individual foods. Protein conformation in foods is important for understanding overall digestibility and allergy, for some foods, and is nearly entirely unstudied. The combination of SEC with offline, bottom-up proteomics based-LC-MS/MS and semi-automated MW calculation is a powerful and surprisingly precise technique. Our described methodology is able to identify multiple isoforms with exceptional protein coverage. The use of the recently derived genomic peanut database, describing all potential peanut isoforms, further enhanced protein characterization. By extracting raw peanuts in TBS, we were able to assess the arrangements of the soluble major seed storage proteins. The majority of the peanut proteins behave as expected, however we observed several interesting anomalies, providing insight into previously suspected peanut protein behavior(s).

Ara h 3 is known to exist in many isoforms, but not all isoforms have been fully elucidated in publically available databases. Using our methodology, we detected 80% (17/21) individual isoforms with exceptional peptide coverage. Purification of each Ara h 3 isoform would be a massive undertaking, but using our methodology, we were able to demonstrate the presence of multiple, individual Ara h 3 isoforms, and their observed MWs. We determined Ara h 3 to exist as an oligomer, but specifically as a mixture of monomers based on the number of calculated subunits present for each isoform (Table 4-2). Since a majority of the Ara h 3 conformations were composed of four or five subunits, it is likely that Ara h 3 exists in a state of equilibrium between the trimeric and

hexameric arrangements. These mixed aggregates may be composed of one individual Ara h 3 isoform, or more likely, various Ara h 3 isoforms. The individual isoform potency is unknown; however, it is likely all isoforms contribute to an individual's overall reaction severity. Furthermore, the potential mixed aggregates, composed of several Ara h 3 isoforms will additionally influence reaction severity.

Similar to other studies of allergenic protein post-translational processing, the *N*-terminal region of Ara h 1 exists as a distinct protein molecule due to post-translational processing (Aalberse et al., 2019; Downs et al., 2014). This *N*-terminal cleavage pattern has been previously described, but lack of association with the mature protein is novel (Wichers et al., 2004). The cleaved *N*-terminus has distinct physicochemical properties distinct to the mature Ara h 1 protein, and could potentially act as an independent sensitization molecule (Aalberse et al., 2019). Similar to our findings, Aalberse et al., (2019) determined the pro-peptide to be present as an individual molecule. According to the IUIS repository, the N-terminal region, or Ara h 1 propeptide, demonstrates allergenic activity without the mature Ara h 1 present, and is therefore recognized as an allergen (IUIS designation: Ara h 1.0101 (26-84)). (Aalberse et al., 2019; Burks et al., 1995a).

A minor proportion of the prolamins (Ara h 2, Ara h 6), which are small, monomeric proteins, eluted uncharacteristically in cupin containing fractions. Previous studies of purified proteins have suggested Ara h 2 and Ara h 3 are cross-reactive, however in light of our observations this may not be the case (Bublin et al., 2013; Smit et al., 2015). Due to their structural arrangements, individuals may be co-sensitized to both prolamins and cupins simultaneously. In both peanut preparations, raw and roasted, we observed identical elution patterns of the prolamins first, predominantly as monomers and

secondly, as a subset co-eluting as oligomers with the cupins. Prolamin peptides co-eluting in the same fraction as the cupins were predominantly cysteine-containing peptides, which may be indicative of interactions between the proteins. These protein arrangement characteristics between prolamins and cupins may be indicative of sensitization patterns among peanut allergic individuals.

Prior to consumption, peanuts generally undergo some type of thermal processing (e.g. boiling, roasting, add one more), inducing protein aggregation by disrupting intramolecular forces causing structural and conformational shifts (Schmitt et al., 2010; Koppelman et al., 1999). These structural shifts lead to aggregation by exposing the hydrophobic groups to the solvent phase where they interact with other matrix components or proteins (Hebling et al., 2012). Other chemical interactions occur between peanut proteins and/or peanut proteins and a food matrix. We observed a 20-fold decrease in Ara h 1 protein recovery after roasting, which is unsurprising as Ara h 1 is highly susceptible to heating (> 80°C) readily forming higher molecular arrangements (MW 500 - 600) (van Boxtel et al., 2006; Chruszcz et al., 2011; Koppelman et al., 1999). Several studies have suggested roasting causes protein aggregation and co-elution of peanut proteins (e.g. prolamins and cupins) however, based on our results, it is apparent these associations exist as a product of plant protein synthesis.

Most often, peanuts are consumed after roasting and possess a different intermolecular protein arrangement than native raw peanuts. Regardless of processing steps, an individual's immune system is exposed to all peanut proteins present in the seed. This has been suspected to be a cause of poly-sensitization as most individuals are sensitized to multiple peanut proteins. However, diagnostic materials may not be

representative of the entire peanut seed. In preparation of diagnostic materials, such as skin prick tests (SPT), it is important to consider protein arrangements, type (raw, roasted), and extraction conditions. Generally, individuals who undergo SPT's are only exposed to the soluble proteins present in the prepared test extract. Understanding these intermolecular arrangements in different peanut preparations, will better inform clinicians and researchers.

Improving our knowledge of the intermolecular associations of complex proteins is important in understanding sensitization patterns and overall protein behavior. Most allergic individuals are sensitized to multiple peanut proteins presenting varying degrees of reaction severity, which is often attributed to roasting (or other thermal treatments).

CONCLUSIONS

To study intermolecular arrangements of peanut proteins, pairing SEC with offline LC-MS/MS, and automated MW calculation generated thousands of identified peptides detected across the SEC column. The combination of these techniques in one comprehensive methodology, allowed for the remarkably precise derivation of native protein MWs and multiple determinations of individual peptides. The use of automated MW determinations by Gaussian modeling further enhances the study of protein conformations, isoforms, and complexes. This methodology can be readily applied to basic protein research for studying interactions, and would be particularly helpful in characterizing allergenic foods as well as biopharmaceuticals.

REFERENCES

Aalberse, R.C., Mueller, G.A., Derksen, N.I.L., Aalberse, J.A., Edwards, L.L., Pomés, A., Lidholm, J., Rispens, T., and Briza, P. (2019). Identification of the amino-terminal fragment of Ara h 1 as a major target of the IgE-binding activity in the basic peanut protein fraction. Clin. Exp. Allergy cea.13554.

Ballmer-Weber, B.K., Lidholm, J., Fernandez-Rivas, M., Seneviratne, S., Hanschmann, K.M., Vogel, L., Bures, P., Fritsche, P., Summers, C., Knulst, A.C., et al. (2015). IgE recognition patterns in peanut allergy are age dependent: perspectives of the EuroPrevall study. Allergy *70*, 391–407.

Bertioli, D.J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., Leal-Bertioli, S.C.M., Ren, L., Farmer, A.D., Pandey, M.K., et al. (2019). The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat. Genet. *51*, 877–884.

Beyer, K., Morrow, E., Li, X.M., Bardina, L., Bannon, G.A., Burks, A.W., and Sampson, H.A. (2001). Effects of cooking methods on peanut allergenicity. J. Allergy Clin. Immunol. *107*, 1077–1081.

Blanc, F., Adel-Patient, K., Drumare, M.-F., Paty, E., Wal, J.-M., Bernard, H., and Bernard, H. (2009). Capacity of purified peanut allergens to induce degranulation in a functional in vitro assay: Ara h 2 and Ara h 6 are the most efficient elicitors. Clin. Exp. Allergy *39*, 1277–1285.

van Boxtel, E.L., Van Beers, M.M.C., Koppelman, S.J., van den Broek, L.A.M., and Gruppen, H. (2006). Allergen Ara h 1 occurs in peanuts as a large oligomer rather than as a trimer. J. Agric. Food Chem. *54*, 7180–7186.

van Boxtel, E.L., van Den Broek, L.A.M., Koppelman, S.J., and Gruppen, H. (2008). Legumin allergens from peanuts and soybeans: Effects of denaturation and aggregation on allergenicity. Mol. Nutr. Food Res. *52*, 674–682.

Bublin, M., Kostadinova, M., Radauer, C., Hafner, C., Szépfalusi, Z., Varga, E.M., Maleki, S.J., Hoffmann-Sommergruber, K., and Breiteneder, H. (2013). IgE cross-reactivity between the major peanut allergen Ara h 2 and the nonhomologous allergens Ara h 1 and Ara h 3. J. Allergy Clin. Immunol. *132*, 118-124.e12.

Burks, A., Cockrell, G., Stanley, J., Helm, R., and Bannon, G. (1995a). Isolation, identification, and characterization of clones encoding antigens responsible for peanut hypersensitivity. Int. Arch. Allergy Immunol. *107*, 248–250.

Burks, A.W., Cockrell, G., Stanley, J.S., Helm, R.M., and Bannon, G.A. (1995b). Recombinant peanut allergen Ara h I expression and IgE binding in patients with peanut hypersensitivity. J. Clin. Invest. *96*, 1715–1721.

Burks, W., Sampson, H.A., and Bannon, G.A. (1998). Peanut allergens. Allergy *53*, 725–730.

Chruszcz, M., Maleki, S.J., Majorek, K.A., Demas, M., Bublin, M., Solberg, R., Hurlburt, B.K., Ruan, S., Mattisohn, C.P., Breiteneder, H., et al. (2011). Structural and

immunologic characterization of Ara h 1, a major peanut allergen. J. Biol. Chem. *286*, 39318–39327.

Downs, M.L., Semic-Jusufagic, A., Simpson, A., Bartra, J., Fernandez-Rivas, M., Rigby, N.M., Taylor, S.L., Baumert, J.L., and Mills, E.N.C. (2014). Characterization of Low Molecular Weight Allergens from English Walnut (Juglans regia). J. Agric. Food Chem. *62*, 11767–11775.

Fekete, S., Beck, A., Veuthey, J.L., and Guillarme, D. (2014). Theory and practice of size exclusion chromatography for the analysis of protein aggregates. J. Pharm. Biomed. Anal. *101*, 161–173.

Flinterman, A.E., Van Hoffen, E., Den Hartog Jager, C.F., Koppelman, S., Pasmans, S.G., Hoekstra, M.O., Bruijnzeel-Koomen, C.A., Knulst, A.C., and Knol, E.F. (2007). Children with peanut allergy recognize predominantly Ara h 2 and Ara h 6, which remains stable over time. Clin. Exp. Allergy *37*, 1221–1228.

Guillon, B., Bernard, H., Drumare, M.-F., Hazebrouck, S., and Adel-Patient, K. (2016). Heat processing of peanut seed enhances the sensitization potential of the major peanut allergen Ara h 6. Mol. Nutr. Food Res. *60*, 2722–2735.

Gupta, R.S., Springston, E.E., Warrier, M.R., Smith, B., Kumar, R., Pongracic, J., and Holl, J.L. (2011). The prevalence, severity, and distribution of childhood food allergy in the United States. Pediatrics *128*, e9–e16.

Gupta, R.S., Warren, C.M., Smith, B.M., Jiang, J., Blumenstock, J.A., Davis, M.M., Schleimer, R.P., and Nadeau, K.C. (2019). Prevalence and Severity of Food Allergies Among US Adults. JAMA Netw. Open *2*, e185630.

Hebling, C.M., Mcfarland, M.A., Callahan, J.H., and Ross, M.M. (2012). Global Proteomic Screening of Protein Allergens and Advanced Glycation Endproducts in Thermally Processed Peanuts. J. Agric. Food Chem. *61*, 5638–5648.

Hurlburt, B., McBride, J., Nesbit, J., Ruan, S., and Maleki, S. (2014). Purification of Recombinant Peanut Allergen Ara h 1 and Comparison of IgE Binding to the Natural Protein. Foods *3*, 642–657.

Jin, T., Guo, F., Chen, Y. wei, Howard, A., and Zhang, Y.Z. (2009). Crystal structure of Ara h 3, a major allergen in peanut. Mol. Immunol. *46*, 1796–1804.

de Jong, E.C., van Zijverden, M., Spanhaak, S., Koppelman, S.J., Pellegrom, H., and Penninks, A.H. (1998). Identification and partial characterization of multiple major allergens in peanut proteins. Clin. Exp. Allergy *28*, 743–751.

Kleber-Janke, T., Crameri, R., Appenzeller, U., Schlaak, M., and Becker, W.M. (1999). Selective cloning of peanut allergens, including profilin and 2S albumins, by phage display technology. Int. Arch. Allergy Immunol. *119*, 265–274.

Klemans, R.J.B., Otte, D., Knol, M., Knol, E.F., Meijer, Y., Gmelig-Meyling, F.H.J., Bruijnzeel-Koomen, C.A.F.M., Knulst, A.C., and Pasmans, S.G.M.A. (2013). The diagnostic value of specific IgE to Ara h 2 to predict peanut allergy in children is comparable to a validated and updated diagnostic prediction model. J. Allergy Clin.

Immunol. *131*, 157–163.

Koppelman, S., Bruijnzeel-Koomen, C., Hessing, M., and de Jongh, H. (1999). Heat-induced conformational changes of Ara h 1, a major peanut allergen, do not affect its allergenic properties. J. Biol. Chem. *274*, 4770–4777.

Koppelman, S., Knol, E., Vlooswijk, R., Wensing, M., Knulst, A., Hefle, S., Gruppen, H., and Piersma, S. (2003). Peanut allergen Ara h 3: isolation from peanuts and biochemical characterization. Allergy *58*, 1144–1151.

Koppelman, S.J., Wensing, M., Ertmann, M., Knulst, A.C., and Knol, E.F. (2004). Relevance of Ara h 1, Ara h 2 and Ara h 3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil-histamine release and intracutaneous testing: Ara h 2 is the most important peanut allergen. Clin. Exp. Allergy *34*, 583–590.

Koppelman, S.J., de Jong, G.A.H., Laaper-Ertmann, M., Peeters, K.A.B.M., Knulst, A.C., Hefle, S.L., and Knol, E.F. (2005). Purification and immunoglobulin E-binding properties of peanut allergen Ara h 6: evidence for cross-reactivity with Ara h 2. Clin. Exp. Allergy *35*, 490–497.

Koppelman, S.J., Jayasena, S., Luykx, D., Schepens, E., Apostolovic, D., de Jong, G.A.H., Isleib, T.G., Nordlee, J., Baumert, J., Taylor, S.L., et al. (2016). Allergenicity attributes of different peanut market types. Food Chem. Toxicol. *91*, 82–90.

Kopper, R.A., Joey, O.N., Sen, M., Helm, R.M., Stanley, J.S., and Burks, A.W. (2005). Peanut protein allergens: the effect of roasting on solubility and allergenicity. Int. Arch. Allergy Immunol. *136*, 16–22.

Kukkonen, A.K., Pelkonen, A.S., Makinen-Kiljunen, S., Voutilainen, H., and Makela, M.J. (2015). Ara h 2 and Ara 6 are the best predictors of severe peanut allergy: a double-blind placebo-controlled study. Allergy *70*, 1239–1245.

Maleki, S.J., Chung, S.-Y., Champagne T., E., and Raufman, J.-P. (2000a). The effects of roasting on the allergenic properties of peanut proteins. J. Allergy Clin. Immunol. *106*, 763–768.

Maleki, S.J., Kopper, R.A., Shin, D.S., Park, C.-W., Compadre, C.M., Sampson, H., Burks, A.W., and Bannon, G.A. (2000b). Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. J. Immunol. *164*, 5844–5849.

Marcus, J.P., Green, J.L., Goulter, K.C., and Manners, J.M. (1999). A family of antimicrobial peptides is produced by processing of a 7S globulin protein in Macadamia integrifolia kernels. Plant J. *19*, 699–710.

Marsh, J., Rigby, N., Wellner, K., Reese, G., Knulst, A., Akkerdaas, J., Van Ree, R., Radauer, C., Lovegrove, A., Sancho, A., et al. (2008). Purification and characterisation of a panel of peanut allergens suitable for use in allergy diagnosis. Mol. Nutr. Food Res. *52*, S272–S285.

Marsh, J.T., Jayasena, S., Gaskin, F., Baumert, J.L., and Johnson, P. (2020). Thermal processing of peanut impacts detection by current analytical techniques. Food Chem. *313*, 126019.

Mondoulet, L., Paty, E., Drumare, M.F., Ah-Leung, S., Scheinmann, P., Willemot, R.M., Wal, J.M., and Bernard, H. (2005). Influence of thermal processing on the allergenicity of peanut proteins. J. Agric. Food Chem. *53*, 4547–4553.

Moreno, F.J., and Clemente, A. (2008). 2S Albumin Storage Proteins: What Makes them Food Allergens? Open Biochem. J. *2*, 16–28.

Müntz, K. (1998). Deposition of storage proteins.

Rabjohn, P., Burks, A.W., Sampson, H.A., and Bannon, G.A. (1999). Mutational analysis of the IgE-binding epitopes of the peanut allergen, Ara h 3: a member of the glycinin family of seed-storage proteins. J. Allergy Clin. Immunol. *103*, S101.

Schmitt, D.A., Nesbit, J.B., Hurlburt, B.K., Cheng, H., and Maleki, S.J. (2010). Processing can alter the properties of peanut extract preparations. J. Agric. Food Chem. *58*, 1138–1143.

Shewry, P.R., Napier, J.A., and Tatham, A.S. (1995). Seed storage proteins: structures and biosynthesis. Plant Cell *7*, 945–956.

Sicherer, S., Muñoz-Furlong, A., Godbold, J., and Sampson, H. (2010). US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. J. Allergy Clin. Immunol. *125*, 1322–1326.

Smit, J.J., Pennings, M.T., Willemsen, K., Van Roest, M., Van Hoffen, E., and Pieters, R.H. (2015). Heterogeneous responses and cross reactivity between the major peanut allergens ara h 1, 2,3 and 6 in a mouse model for peanut allergy. Clin. Transl. Allergy *5*, 13.

Turner, P.J., Baumert, J.L., Beyer, K., Boyle, R.J., Chan, C.H., Clark, A.T., Crevel, R.W.R., DunnGalvin, A., Fernández-Rivas, M., Gowland, M.H., et al. (2016). Can we identify patients at risk of life-threatening allergic reactions to food? Allergy Eur. J. Allergy Clin. Immunol. *71*, 1241–1255.

Uversky, V.N. (1993). Use of Fast Protein Size-Exclusion Liquid Chromatography To Study the Unfolding of Proteins Which Denature through the Molten Globule. Biochemistry *32*, 13288–13298.

Le Vent, S. (1995). Simulation of chromatographic peaks by simple functions. Anal. Chim. Acta *312*, 263–270.

Wichers, H.J., de Beijer, T., Savelkoul, H.F.J., and van Amerongen, A. (2004). The Major Peanut Allergen Ara h 1 and Its Cleaved-off N-Terminal Peptide; Possible Implications for Peanut Allergen Detection. J. Agric. Food Chem. *52*, 4903–4907.

Zhang, W., Zhu, Q., Zhang, T., Cai, Q., and Chen, Q. (2016). Thermal processing effects on peanut allergen Ara h 2 allergenicity in mice and its antigenic epitope structure. Food Chem. *212*, 657–662.

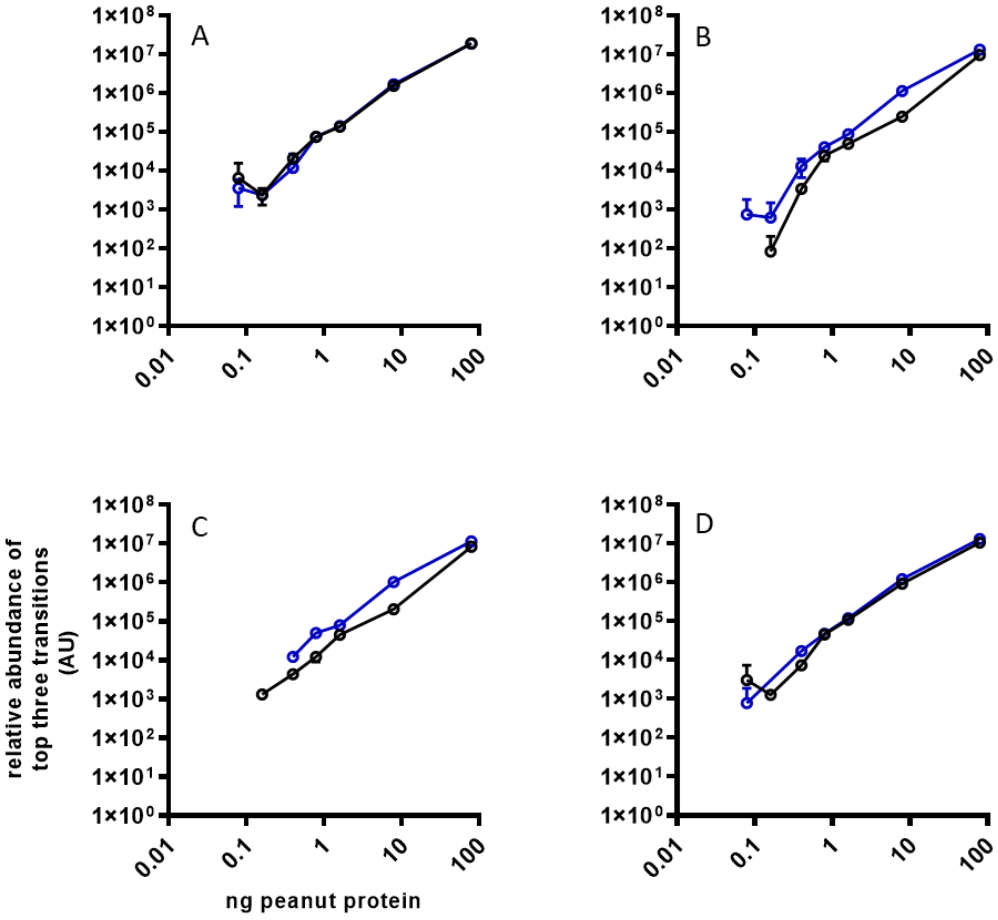
SUPPLEMENTARY

CHAPTER 2

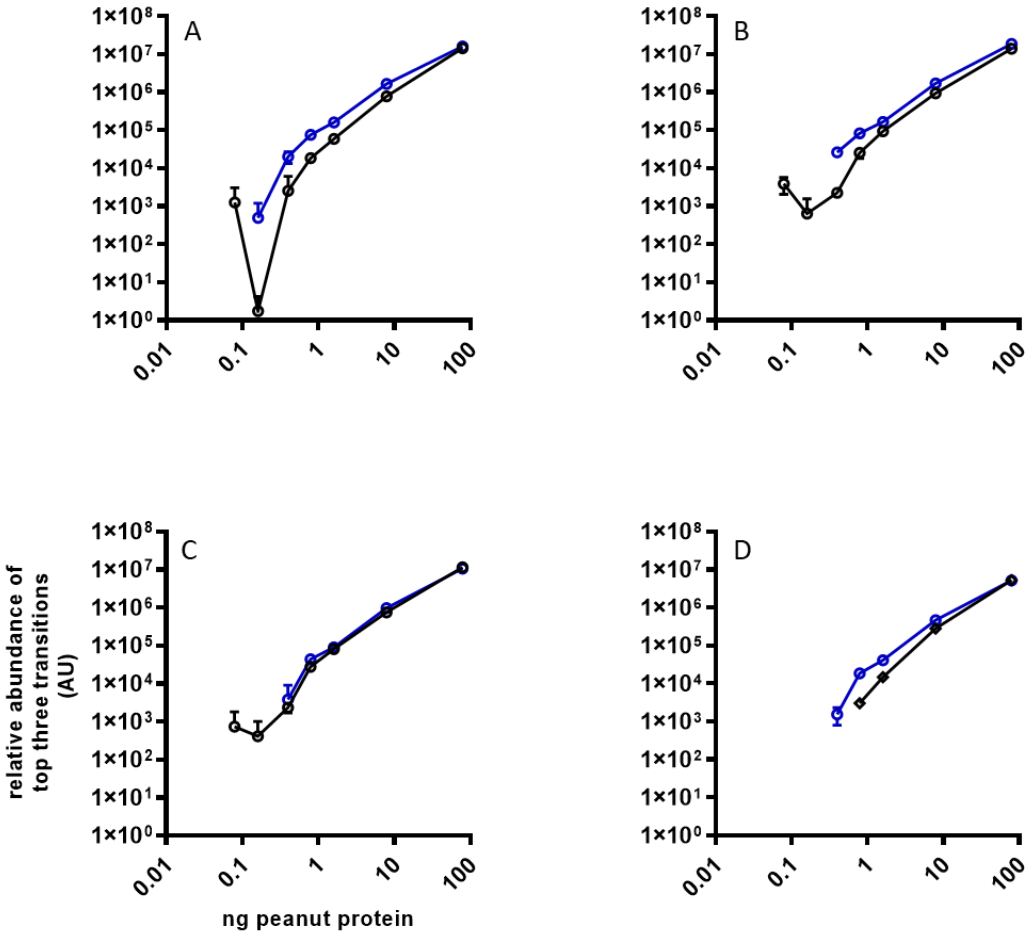
			SPT (AU)	IgE (ISU)					ICAP IgE (kUa/L)
Patient ID	Dilution	Other atopy	Peanut	Ara h 1	Ara h 3	Ara h 2	Ara h 6	Peanut	Peanut
PN2060	1:20	yes	9.0	51.7	27.1	64.5	76.0	224	53.0
PN2073	1:50	yes	35	60.4	15.1	31.3	64.5	141	575
PN2075	1:20	yes	20	50.2	14.5	24.0	50.2	378	72.5
DD	1:20	no	15	57.7	5.71	18.4	21.3	n.d.	85.5
PN2079	1:50	no	12	96.7	58.2	98.5	124.3	173	787
Pool	1:50	n.a.	n.a.	53.4	15.2	39.3	76.5	251	n.d.

Supplementary Table 2-1. Peanut allergic subject characteristics and serology. Dilution value represents serum dilution prior to serologic measurements. SPT indicates the value for skin prick test, IgE (ISU) represents, ICAP represents the Immunocap measured responses.

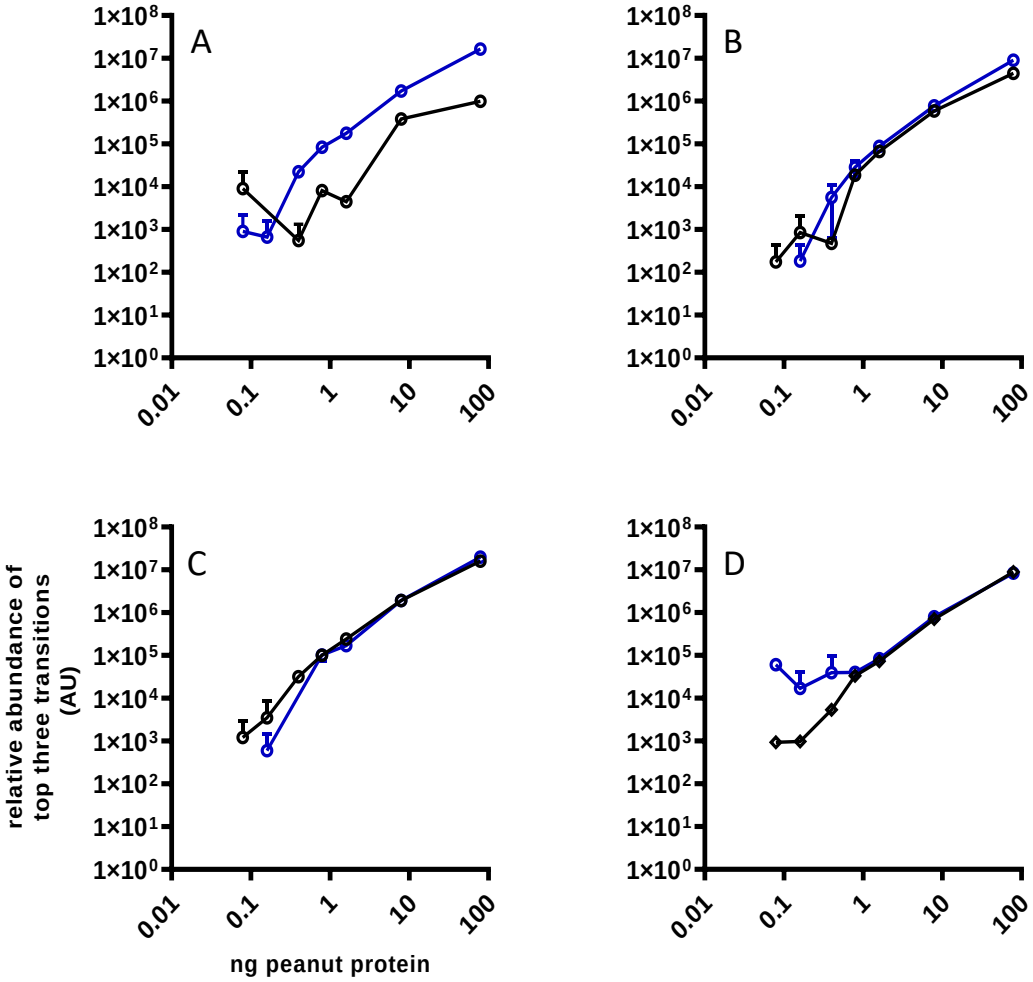
CHAPTER 3



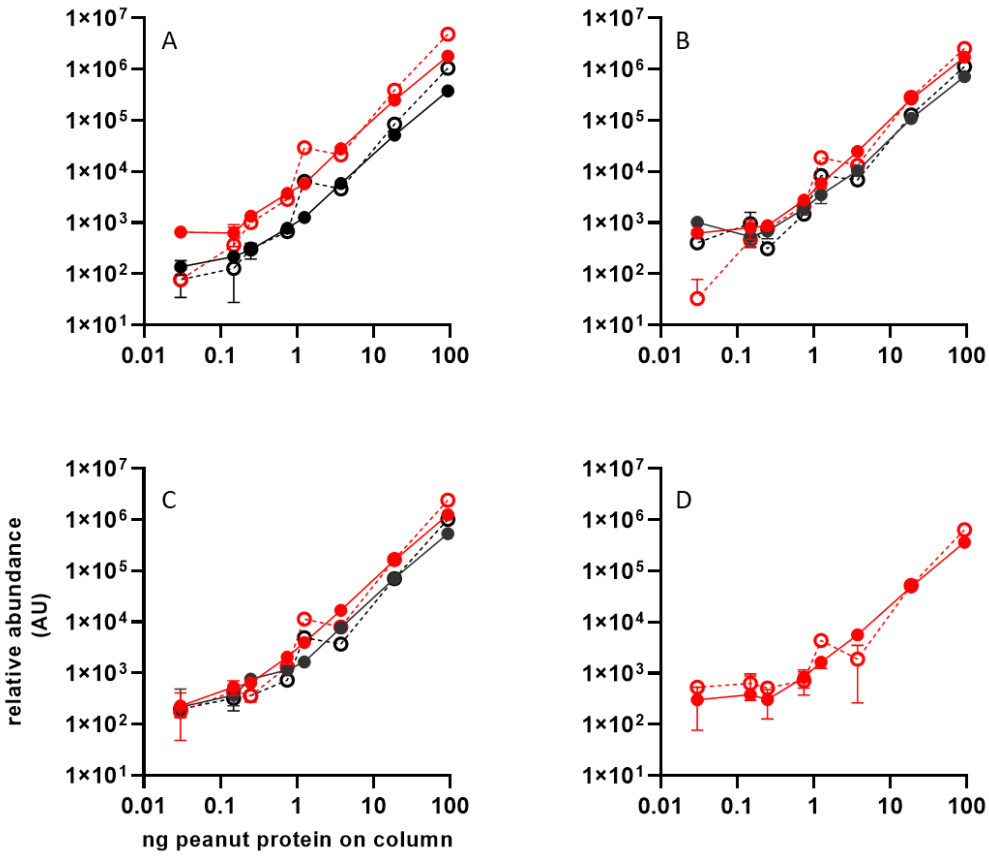
Supplementary Figure 3-1. Relative abundance of individual Ara h 1 peptides (A-D) at decreasing quantities of peanut protein in a constant amount of serum in PRM. Blue lines represent sera-peanut matrices; black lines represent peanut only in the absence of a sera background.



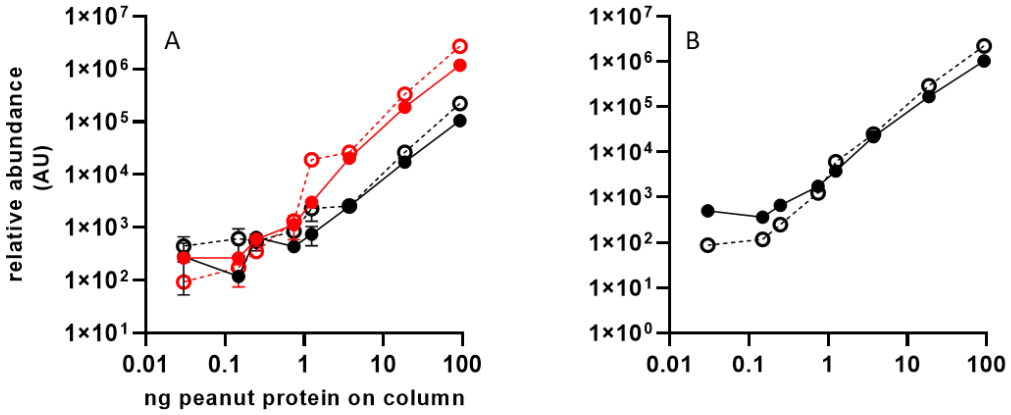
Supplementary Figure 3-2. Relative abundance of individual (A) Ara h 2 and (B-D) Ara h 6 peptides at decreasing quantities of peanut protein in a constant amount of serum in PRM. Blue lines represent sera-peanut matrices; black lines represent peanut only in the absence of a sera background.



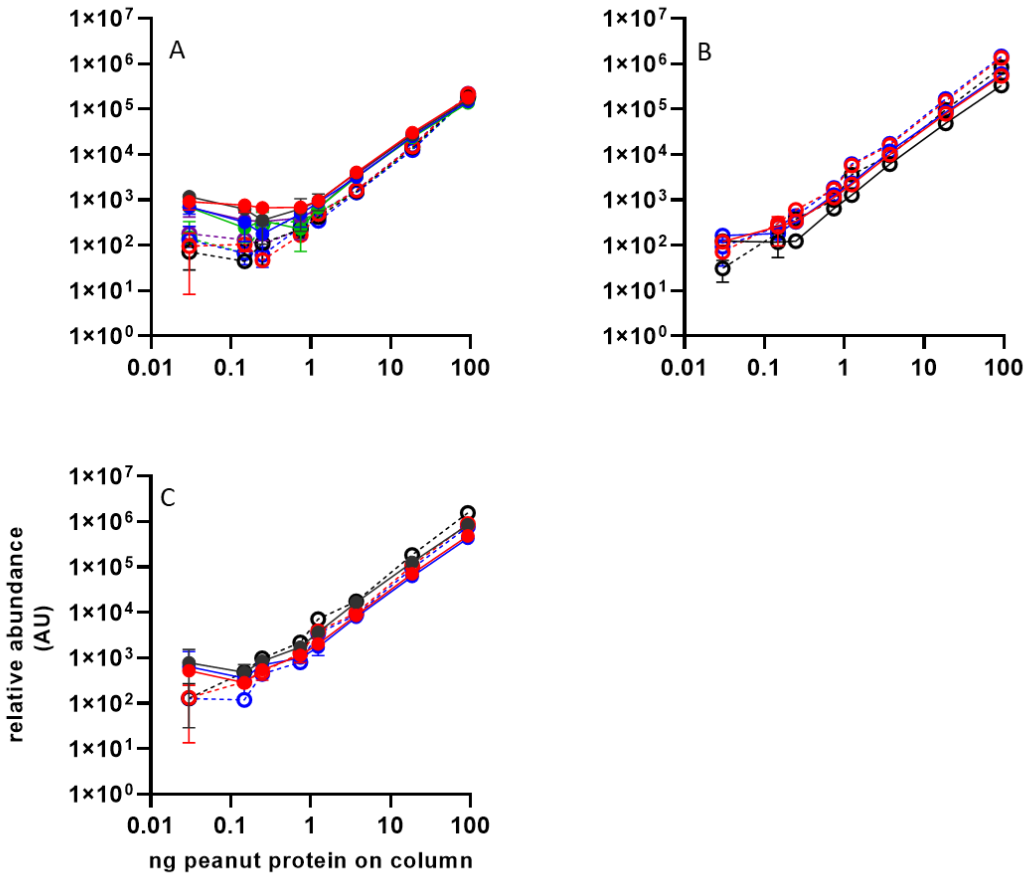
Supplementary Figure 3-3. Relative abundance of individual Ara h 3 peptides (A-D) at decreasing quantities of peanut protein in a constant amount of serum in PRM. Blue lines represent sera-peanut matrices; black lines represent peanut only in the absence of a sera background.



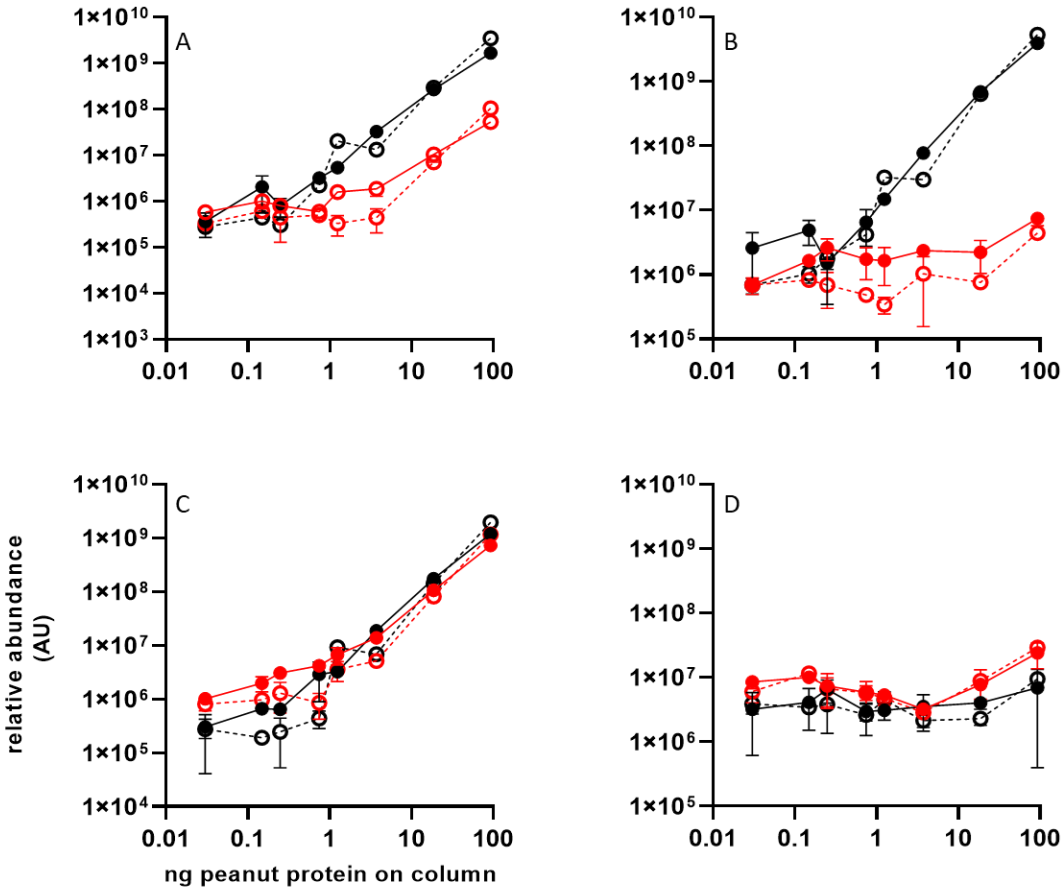
Supplementary Figure 3-4. Relative abundance of individual Ara h 1 peptides (A-D) at decreasing quantities of peanut protein in a constant amount of serum in MRM. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.



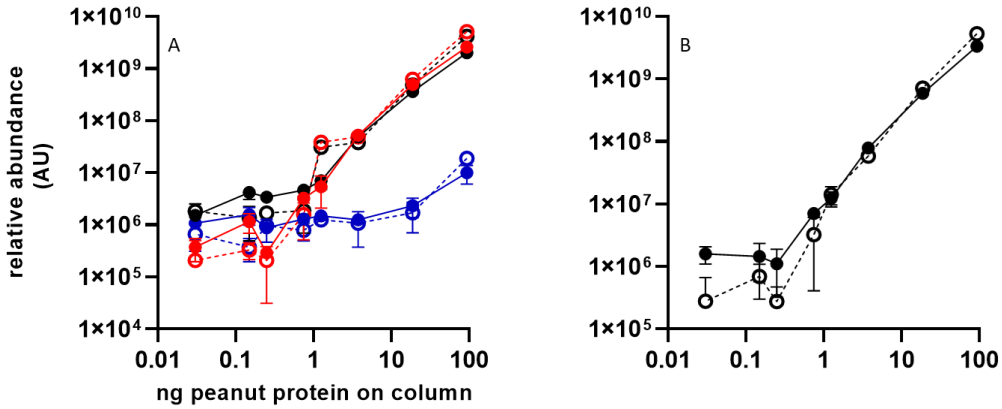
Supplementary Figure 3-5. Relative abundance of individual (A) Ara h 2 (B) Ara h 6 peptides at decreasing quantities of peanut protein in a constant amount of serum in MRM. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.



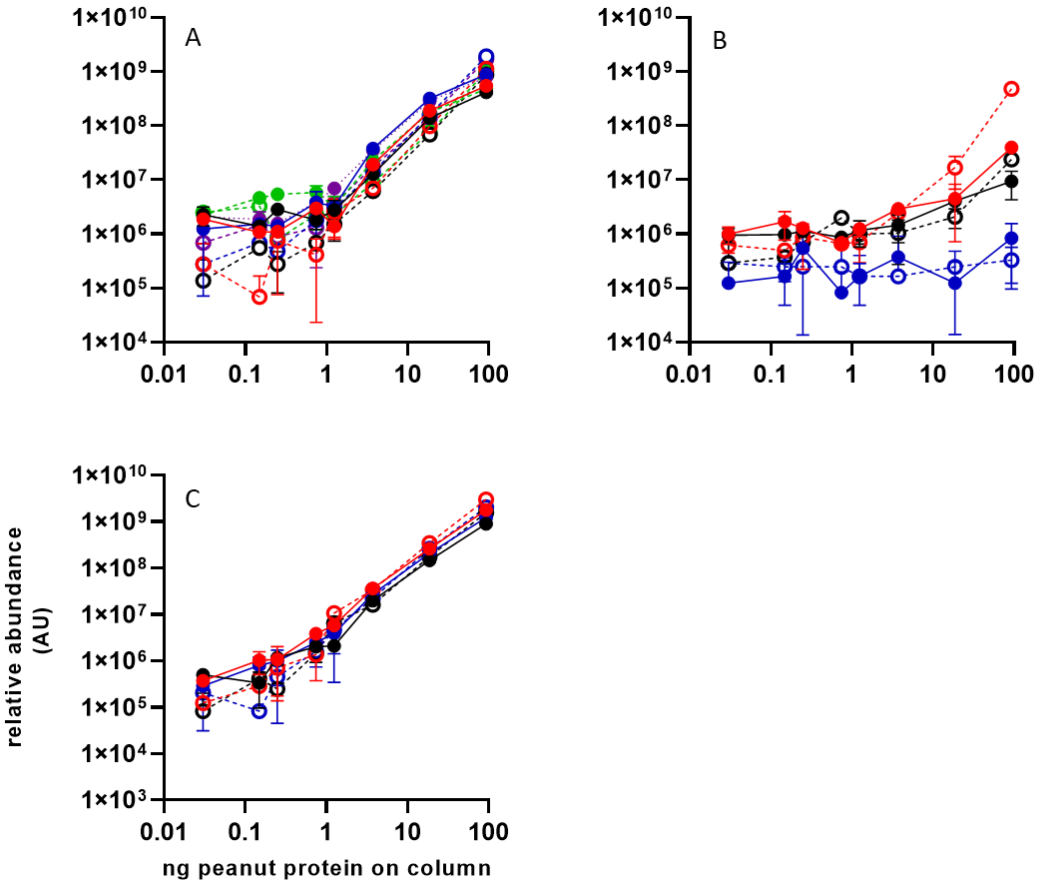
Supplementary Figure 3-6. Relative abundance of individual Ara h 3 peptides (A-C) at decreasing quantities of peanut protein in a constant amount of serum in MRM. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.



Supplementary Figure 3-7. Relative abundance of individual Ara h 1 peptides (A-D) at decreasing quantities of peanut protein in a constant amount of serum in MRM^3. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.



Supplementary Figure 3-8. Relative abundance of individual (A) Ara h 2 (B) Ara h 6 peptides at decreasing quantities of peanut protein in a constant amount of serum in MRM^3. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.



Supplementary Figure 3-9. Relative abundance of individual Ara h 3 peptides (A-C) at decreasing quantities of peanut protein in a constant amount of serum in MRM^3. Transitions are represented by different colored lines. Solid lines represent sera-peanut matrices; dashed lines represent peanut only in the absence of a sera background.

Protein	Peptide	transitions			
Ara h 3	SPDIYNPQAGSLK	y9 - 977.5051+	y8 - 814.4417+	y7 - 700.3988+	
Ara h 3	WLGLSAEYGNLYR	y9 - 1072.5058+	y8 - 985.4738+	y5 - 622.3307+	
Ara h 3	TANDLNLLILR	y9 - 1083.6521+	y8 - 969.6091+	y7 - 854.5822+	y6 - 741.4981+
Ara h 6	VNLKPCEQHIMQR	y5 - 684.3610+	y4 - 547.3021+	y3 - 434.2180+	y9 - 599.7764++

Supplementary Table 3-1. List of peptide transitions detection in positive subject sera samples (subject 1, 2).