STRUCTURAL CHARACTERIZATION OF GROUP 2 MITE ALLERGENS AND AN INVESTIGATION INTO THEIR INTERACTIONS WITH SMALL MOLECULE LIGANDS

By

Christina Marie Linn

A THESIS

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

Biochemistry and Molecular Biology – Master of Science

2024

ABSTRACT

Mites are ubiquitous arthropods inhabiting nearly every type of ecosystem. Domestic mites, which include house dust mites and storage mites, are one of the major sources of indoor aeroallergens worldwide with house dust mites being the leading cause of allergic asthma and rhinitis. It is not the mites themselves that are the cause of allergy, but various proteins found in the mite's fecal pellets, salivary secretions, cuticle sheds, and eggs that are allergy inducing. The International Union of Immunological Societies and the World Health Organization recognize 128 proteins produced by mites as being allergenic. Of these, Groups 1 and 2 are considered major allergens. However, studies suggest that it isn't the protein alone that is responsible for allergy, but the small molecule ligands that bind them likely play a role in the disease process as well. The purpose of this research is to structurally characterize Group 2 allergens from all 8 mite species that are known to produce them using X-ray crystallography methods, and to study the small molecules that bind them. Novel methods of Group 2 allergen expression were used to purify Der f 2, Tyr p 2, and Blo t 2 with Der p 2 being purified using previously established methods. *In silico* studies as well as probe displacement assays were used to study the binding of 9 fatty acids (of varying carbon length and saturation) and cholesterol to the binding pocket of Blo t 2 and Der p 2. The computational studies suggest no binding to cholesterol for either protein. At the same time, these studies indicate that Blo t 2 may bind 7 of the 9 fatty acids tested, whereas Der p 2 may bind 8 of the 9 fatty acids. Initial probe displacement assays indicate binding of Blo t 2 to oleic and pinolenic acid with a possible higher affinity for pinolenic acid.

I would like to dedicate my work to my best friend and husband Craig. You make all of life better!

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my graduate research supervisor Dr. Masksymilian Chruszcz. Your mentorship has provided me with invaluable opportunities for growth, fostering an environment where learning and occasional setbacks were welcomed as essential components of progress. Thank you for your insistence on excellence and holding the bar high in the rigor and integrity of our work. Through your guidance, I have honed my ability to pose insightful questions and approach challenges from diverse perspectives. I also owe a huge thank you to my committee members Dr. Benjamin Orlando and Dr. Cheryl Rockwell. Your thoughtful insights and probing inquiries have significantly enriched my research journey, pushing me to delve deeper into the complexities of my work. To my labmates, especially Kriti Khatri with whom I worked most closely, thank you, thank you, thank you! We have the best lab!! I would also like to thank our collaborators Dr. Alain Jacquet from Chulalongkorn University and Dr. Geoff Miller from the National Institute of Environmental Health and Sciences, as well as the core facilities here at MSU without whom none of this research would be possible.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

HDM	House dust mite
SM	Storage mite
WHO	World Health Organization
IUIS	International Union of Immunological Societies
Der p	*Dermatophagoides pteronyssinus*
Der f	*Dermatophagoides farinae*
LPS	Lipopolysaccharide
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
LB	Lauria broth
TB	Terrific broth
PDI	Protein disulfide isomerase
TEV	Tobacco etch virus
Ni-NTA	Nickel Nitrilotriacetic Acid
IMAC	Immobilized Metal Affinity Chromatography
DSF	Differential scanning fluorimetry
PVDF	Polyvinylidene fluoride
TBST	Tris buffer saline with tween-20
RT	Room temperature
HRP	Horseradish peroxidase
APC	Antigen-presenting cell
IL	Interleukin
MD-2	Myeloid differentiation factor 2
TLR-4	Toll like receptor 4
LPS	Lipopolysaccharide
AIT	Allergy immunotherapy
DAUDA	11-(Dansylamino) undecanoic acid
IgE	Immunoglobulin E

IgG	Immunoglobulin G
POPC	1-palmitoyl-2-oleoylphosphatidylcholine
RMSD	Root mean square deviation

CHAPTER 1

INTRODUCTION

Mites - class Arachnida, subclass acari – are ubiquitous arthropods inhabiting nearly every type of ecosystem: from soil and water to the skin, hair, and fur of animals (including humans), to the dust and stored food in human dwellings (Bensoussan et al., 2016; Coddington & Colwell, 2001). Roughly 55,000 species have been identified (Krantz, 1978) (there are likely thousands more unidentified) but at their tiny size of 0.5-2.0 mm in length they go virtually unnoticed by humans and other animals (Azad, 1986) There are some however that cause major problems for humans around the globe.

Domestic mites, which include storage mites (SM) and house dust mites (HDM), are of the order Sarcoptiformes and found in several families (Figure 1). They are one of the major sources of indoor aeroallergens worldwide (Jeffrey D. Miller, 2019) and the leading cause of allergic rhinitis and asthma affecting an estimated 10-20% of the world's population. This is not only cause for great concern, but the treatment of allergic symptoms puts a tremendous economic burden on healthcare systems (Sanchez-Borges et al., 2018; Yuriev et al., 2023). Both types of mites require warm, humid environments for maturation and reproduction however, populations can survive months of arid conditions by forming desiccation-resistant quiescent protonymphs. When conditions become more favorable these quiescent protonymphs become active, mature into adults, and reproduce. Thus, many individuals with dust mite allergies find their symptoms to be worse in the summer than in the winter (Arlian et al., 2001). It is important to note that mite eggs are capable of surviving temperatures from < -60°F to over 100°F which is well outside the range of human dwellings (J. D. Miller, 2019). It is not the mites themselves that produce allergic response, but various enzymes, proteases, and other proteins found in mite fecal pellets, salivary secretions, cuticle sheds, and eggs that are allergy inducing (Bessot & Pauli, 2011; Khatri et al., 2023).

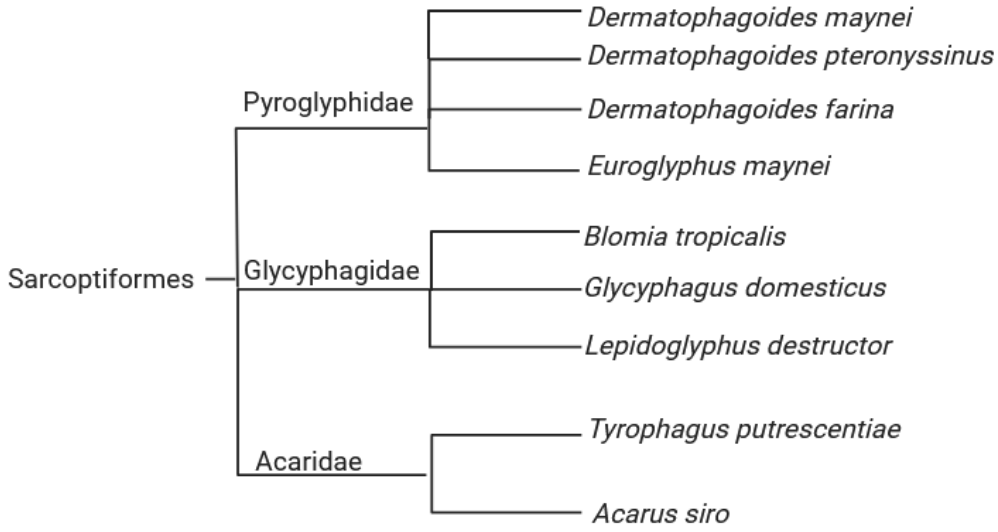


Figure 1. A diagram showing the relationship between house dust mites and storage mites. (Created in BioRender.com)

The immune system's aberrant reaction to mite proteins is classified as an IgE-mediated response or Type I hypersensitivity (Calderon et al., 2015; Larche et al., 2006). Type 1 hypersensitivity occurs in 2 stages, a sensitization phase and an effect phase as depicted in Figure 2 (Abbas et al., 2023). During the sensitization phase antigen-presenting cells (APCs) bind (usually) a protein, also called an antigen, and present it to naïve $CD4^+$ T-cells. These T- cells then differentiate into Th2 cells and begin secreting interleukin (IL)-4 and IL-13. IL-4 plays a role in Th2 cell proliferation as well as IgE synthesis and maintenance. IL-13 is thought to play a role mainly in the effector phase as an inducer of the phenotypic manifestations of allergic disease (Gour & Wills-Karp, 2015). When Th2 cells interact with B cells the B cells class switch and begin producing allergen-specific immunoglobulin E (IgE) (Stavnezer & Schrader, 2014; Vitte et al., 2022). These IgEs then bind to high-affinity receptors (FcεRI) on granulocytes sensitizing them to a specific antigen. The effect phase occurs upon subsequent exposures to an allergen source. In this phase antigens will bind two or more IgEs (this is called cross-linking) that are already bound to a granulocyte causing

the granulocyte to degranulate and release histamine, proteoglycans, and various proteases (Vitte et al., 2022). The release of these pro-inflammatory mediators manifests as an increase in mucous secretions, coughing, rhinitis, atopic dermatitis, abdominal cramping, and potentially anaphylaxis (Abbas et al., 2023; Stavnezer & Schrader, 2014; Vitte et al., 2022).

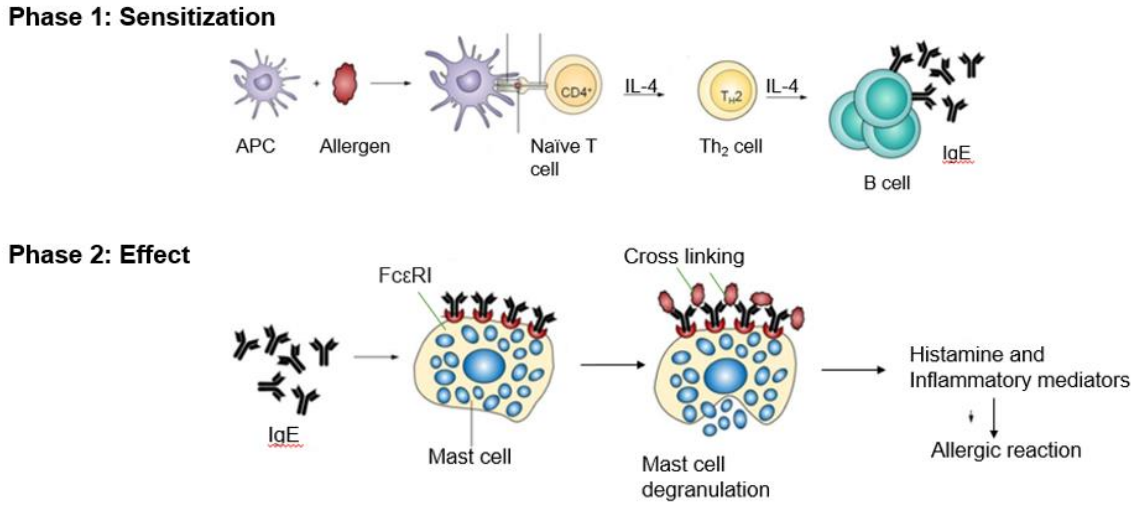


Figure 2. Pictogram showing the 2 stages of Type 1 Hypersensitivity. Phase 1 shows the initial sensitization and phase 2 shows what happens upon subsequent exposures to an antigen. [17-19,56]

Currently, the World Health Organization (WHO) and the International Union of Immunological Societies (IUIS) recognize 128 proteins produced by mites as allergenic, 71 of which belong to *Dermatophagoides pteronyssinus* (Der p) and *Dermatophagoides farinae* (Der f) (allergen.org). Of these 71, allergens belonging to Groups 1 and 2 are recognized as being major allergens (Thomas, 2015). Group 1 allergens are cysteine proteases that can cut transmembrane molecules, destroying the epithelial barrier in the bronchi and bronchioles (Khatri et al., 2023). Group 2 allergens, which are the focus of this research, are structurally similar to myeloid differentiation factor 2 (MD-2) and bind lipopolysaccharide (LPS) in a similar manner allowing them to trigger Toll-like receptor 4 (TLR4). The activation of TLR4 initiates a cascade of protein-protein interactions leading

to the differentiation of naïve T-cells into Th2 cells and a Type 1 hypersensitivity response (H. Xu et al., 2019)

Allergy symptoms are often treated with pharmacological drugs, but these provide only temporary, short-term relief. With HDM allergies being a year-round problem for allergy sufferers, a long-term solution is desirable (Durham & Shamji, 2023) This may be possible with allergy immunotherapy (AIT). AIT has been used for 100 years as a means of desensitizing the immune system and inducing a more tolerant response to allergens. This is often accomplished by administering a modified version of the allergen either by injection or sublingually (Durham & Shamji, 2023; Kim, 2023). The modified allergen is frequently a recombinant protein with mutations in the B-cell epitopes which can induce allergen-specific IgG production through class switching, but not IgE production (Kim, 2023). Allergen-specific IgGs inhibit IgE-allergen interactions which impedes cross-linking and subsequent activation and degranulation of granulocytes (Strobl et al., 2023). Thereby allowing for a more tolerant immune response to the allergen.

In order to produce safe and effective allergens for AIT, the allergens themselves must be well characterized (Scheiner, 1993). Allergens have defined 3D structures that determine their molecular surface and antibody binding epitopes which are conformational rather than linear. This means antibodies are interacting with amino acid residues on the allergen that are close in space due to their tertiary structure but not always close sequentially (Pomes et al., 2020). Theoretically, epitopes could be located anywhere on the allergen, but studies show that antibodies are very specific about the epitopes they will recognize and there seem to be areas on allergens that are more preferential for antibody binding (Pomes et al., 2020).

Characterizing allergens and their antibody binding epitopes is not an easy task. Researchers frequently turn to X-ray crystallography and NMR, as high-resolution data provides detailed information about the three-dimensional structure of allergens, epitope regions, and chemical interactions (Pomes et al., 2020). Over 88% of the experimental models deposited in the Protein Data Bank (PDB) were determined using X-ray crystallography (Pomes et al., 2020). One of the challenges with employing these methods, however, is the requirement for large quantities of pure and homogenous

protein preparations (Pomes et al., 2020). Crude allergen extracts often have a mixture of allergen isoforms that are difficult to separate and/or the allergen of interest is at such a low concentration it is impossible to obtain in sufficient amounts (Tscheppe & Breiteneder, 2017). To get around this difficulty, researchers have turned to recombinant protein technology. Recombinant protein expression and purification methods allow for the production of allergens in both high yield and purity (Tscheppe & Breiteneder, 2017).

Not only is it important to understand the tertiary structure of allergens and their antibody binding epitopes, but also their ability to bind small molecule ligands. About a decade ago it was noted that ligands can influence the allergenic properties of proteins through direct interaction with the immune system or by altering the intrinsic properties of allergens and thereby increasing their ability to interact with the immune system (Chruszcz et al., 2021; Khatri et al., 2023). It has been found that a significant fraction of major allergen families binds lipids and/or lipid derivatives (Khatri et al., 2023). Group 2 allergens are one such family. In particular, the Group 2 allergens from *Dermatophagoides pteronyssinus* (Der p 2) and *Dermatophagoides farinae* (Der f 2) have been shown to bind LPS and cholesterol. As discussed earlier, Der p 2 in complex with LPS triggers the TLR4 signaling pathway. Der p 2 has also been shown to bind 1-palmitoyl-2-oleoylphosphatidylcholine (POPC), a diacylglycerol phospholipid frequently used in studying lipid rafts (Khatri et al., 2023; Wanderlingh et al., 2017). Crystal structures of Der p 2 and Der f 2 show the structural flexibility of this family of proteins making it likely that they can bind a variety of sterols (Khatri et al., 2023). However, it is unknown if the presence of a small molecule ligand in the binding pocket of these proteins influences their interactions with antibodies (Khatri et al., 2023).

The focus of this project has been primarily on Group 2 allergen structural studies via X-ray crystallography. This requires milligrams of high purity protein and many crystallization conditions to get diffraction quality-crystals. Consequently, a significant aspect of the project has been developing methods for the production of Group 2 allergens in *E. coli*. The other part of this project has focused on small molecule ligand binding of Group 2 allergens with a particular focus on free fatty acid binding. Identifying the ligands that bind these proteins will give us insight not only into

the physiological role they play for the mite, but also into their ability to induce allergy in humans.

CHAPTER 2

STRUCTURAL ANALYSIS OF GROUP 2 ALLERGENS FROM MITES

2.1 Group 2 Allergens

Group 2 allergens belong to the NPC2 family of proteins. They are relatively small with an average size of ~15 kDa. They are a single domain protein consisting of six β-strands that form two β-sheets which are stabilized by three disulfide bonds. When folded properly a hydrophobic cavity is formed (Ichikawa et al., 1998). Proteins in this family have high structural homology but not always high sequence similarity as can be seen in Figure 3. However, a sequence alignment of Group 2 allergens shows that the 6 cysteines involved in disulfide bond formation are 100% conserved across all eight mite species for which there are registered Group 2 allergens. (Figure 4.)

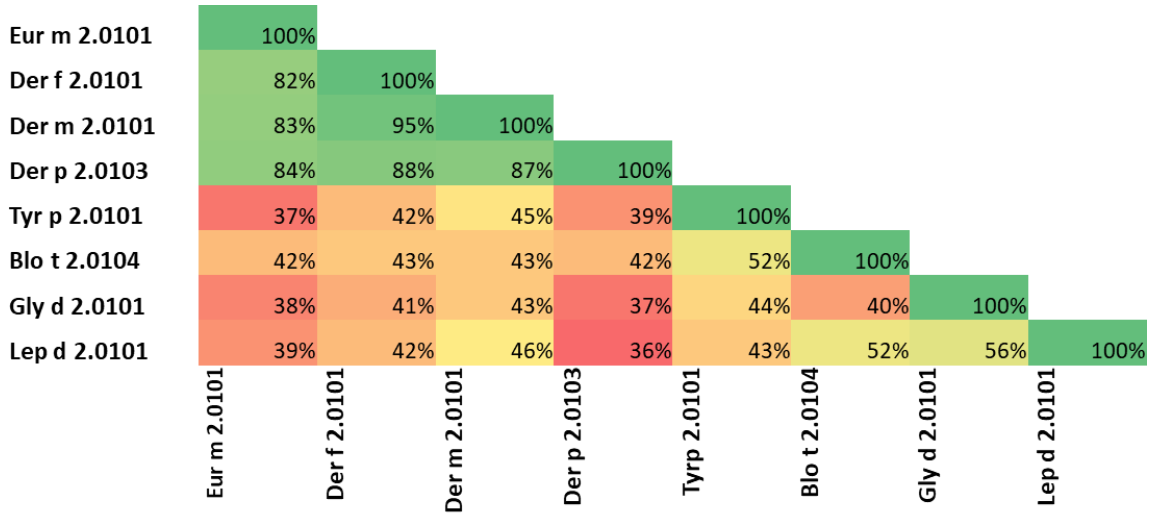


Figure 3. Diagram showing the percent sequence identity for Group 2 allergens across all mite species known to induce allergy. The isoform used for the sequence comparison was 2.0101 for *E. maynei, D. farinae, D. maynei, T. putrescentiae, G. domesticus,* and *L. destructor.* Isoform 2.0103 was used for *D. pteronyssinus,* and 2.0104 was used for *B. tropicalis.* These were the same isoforms used in protein expression and purification experiments.

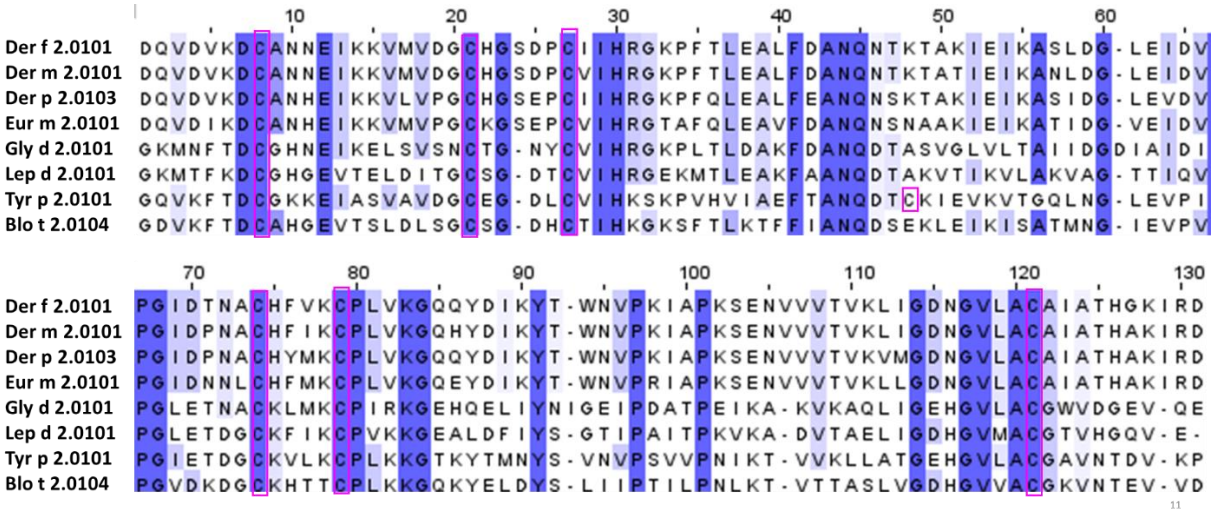


Figure 4. Sequence alignment of Group 2 allergens across all mite species known to induce allergy. The isoforms used for the sequence alignment were 2.0101 for *E. maynei, D. farinae, D. maynei, T. putrescentiae, G. domesticus,* and *L. destructor.* Isoform 2.0103 was used for *D. pteronyssinus,* and 2.0104 was used for *B. tropicalis.* These were the same isoforms used in protein expression and purification experiments. The six conserved cysteines are boxed out in pink. Tyr p 2 has a 7th cysteine which is also boxed in pink.

What is not so well known about these proteins is the physiological role they play for mites. Studies suggested that they bind cholesterol in a manner similar to human NPC2 which is an essential cholesterol transporter (Reginald & Chew, 2019; Y. Xu et al., 2019). Other studies show Der p 2 and Der f 2 bind lipopolysaccharide (LPS), an endotoxin and major component of the outer membrane of Gram-negative bacteria (Bertani & Ruiz, 2018; Bessot & Pauli, 2011). Therefore, identification of small molecule ligands of Group 2 allergens will not only allow for elucidation of physiological role of these proteins in mites but will also improve our understanding of allergies.

2.2 Protein Expression and Purification Methods

2.2.1 Expression and Purification

Constructs encoding the mature protein forms of Eur m 2, Tyr p 2, and Der f 2 were ordered from Synbio Technologies. All constructs were inserted into pET-15b(+) vector with C-terminal 6x-Histidine tag added for purification purposes. Blo t 2 construct was in pET-15b(+) vector with an N-terminal 6x-Histidine tag and was gifted by Dr. Alain Jacquet's from Chulalongkorn University in Thailand. All constructs were transformed by heat shock protocol into *E. coli* T7 SHuffle Express cells (New England Bio Labs

(NEB)). Once transformed the cells were incubated overnight in 5 mL of Luria broth (LB), 300 µL of Recovery Media (SIGMA Aldrich), and 100 µg/mL of ampicillin at 37°C and shaken at 200 rpm. The next day the overnight cultures were transferred to 1 L of LB with an additional 100 µg/mL of ampicillin and 1 mL of Recovery Media and grown at 35°C and shaken 200 rpm until an OD of 0.6-0.8 was reached. The cultures were then cooled to 16°C and induced with 400 mM of isopropyl ß-D-1-thiogalactopyranoside (IPTG). The induced cultures were grown overnight at 16°C and shaken at 150 rpm. The next day the cultures were collected by centrifugation in a Sorvall RC6 Plus Superspeed Centrifuge at 16,000 g for 8 min. Collected pellets were stored at -80°C until purification.

Der p 2 was expressed in terrific broth (TB) using the same protocol as described for other Group 2 proteins but without the recovery media and in co-expression with protein disulfide isomerase (PDI). PDI was in a vector with kanamycin resistance so along with 100 mg/L of ampicillin, kanamycin was added at 50 mg/L. The vector used to express Der p 2 was pET-28b(+) with an N-terminal 6x – Histidine tag followed by a tobacco etch virus (TEV) cleavage site for purification purposes.

Purification was done using Immobilized Metal Affinity Chromatography (IMAC) and Nickel Nitrilotriacetic Acid (Ni-NTA). Cell pellets were lysed using binding buffer, 20 mM Na_2HPO_4, 500 mM NaCl, 5 mM imidazole, pH 8.0 with a QSonica sonicator at 50% amplitude for 3 minutes with a pulse rate of 10 sec on and 40 sec off. After sonication the lysate was centrifuged in J-26S XPI Beckmann Coulter Ultracentrifuge at 8,817 g for 45 minutes. The supernatant was poured onto a 5 mL Ni-NTA gravity column equilibrated with binding buffer. After all the flow through had been collected, the column was washed with 3-4 column volumes of 20 mM Na_2HPO_4, 500 mM NaCl, 20 mM imidazole, pH 8.0. Lastly, the protein was eluted in multiple fractions from the column using 20 mM Na_2HPO_4, 500 mM NaCl, 300 mM imidazole, pH 8.0. All fractions containing protein, as confirmed by Bradford Protein Assay Reagent (Thermo Fisher), were pooled and ran on a HiPrep DEAE sepharose FF 16/10 (Cytiva) column attached to a GE Healthcare ÄKTA-Pure FPLC. Buffer A was 20 mM Na_2HPO_4, pH 8.0, and buffer B (for elution) was 20 mM Na_2HPO_4, 2 M NaCl, pH 8.0. The protocol called for a 50% increase of buffer B for every column volume of buffer flowed through the column.

Fractions containing the protein of interest were collected, pooled, and run on a HiLoad 16/600 Superdex 200 pg gel filtration column using 20 mM Na_2HPO_4 and 150 mM NaCl as the mobile phase. The fractions containing the protein were collected and the presence and purity of the protein of interest checked by SDS-PAGE and western blot.

Proper folding of the protein was checked by 1D-NMR and thermal stability was checked by differential scanning fluorimetry (DSF). Fractions containing well folded protein were concentrated and used for crystallographic and ligand binding studies.

2.2.2 Western Blot

Proteins contained within an SDS-PAGE gel were electro transferred to Amersham™ Hybond™ 0.2 µM polyvinylidene fluoride (PVDF) blotting membrane (Cytiva) via a Mini Trans-Blot® cell (Bio-Rad) using a PowerPac Universal (Bio-Rad) as the power supply. The Mini Trans-Blot® cell containing the membrane and the gel was placed in ice to maintain a temperature of 4°C. Transfer buffer, 25 mM Tris-HCl, 192 mM glycine, 20% (v/v) methanol at pH 8.3, was added to the cell and the voltage set to 100V for 30-45 minutes. Once the transfer was complete the membrane was blocked overnight with 5% (w/v) instant nonfat dry milk in tris buffer saline with Tween-20 (TBST) (50 mM Tris-HCl, 150 mM NaCl, 0.1% (v/v) Tween-20, pH 7.5). Blocking occurred at room temperature (RT) on a digital rocker (Thermo Fisher) with a tilt of 10° and a speed of 30 rpm.

The next day the membrane was washed 3 x 5 minutes with TBST at RT, then incubated with mouse anti-histidine antibody conjugated with horseradish peroxidase (HRP) at a concentration of 1:3,000 for 1 hour at RT on the digital rocker with a 10° tilt and a speed of 30 rpm. After incubation with the antibody, the membrane was again washed 3 x 5 minutes in TBST at RT. Lastly the membrane was incubated with Clarity™ Western ECL Substrate (Bio-Rad) for 5 min at RT then visualized on a Sapphire biomolecular imager (azure biosystems) using Sapphire capture software.

2.3 Confirmation of Tertiary Structure of Proteins

2.3.1 Differential Scanning Fluorimetry

Conditions for all DSF experiments were: ~72 µM of protein, 2 µM BODIPY™ FL L-cystine (Thermo Fisher), a final DMSO concentration of 0.02%, and a final reaction volume of 20 µL in a 96-well PCR plate (Bio-Rad Hard-Shell PCR-96-thin well, white

shell/clear well plate) (Hofmann et al., 2016). The buffer was 50 mM Tris-HCl with an increasing concentration of NaCl from 0.25-0.75 mM and an increasing pH from 6.5-9.0. DSF experiments were run in triplicate and monitored in a CFX Opus96 Real-time PCR system. For all experiments the starting temperature was 4°C and held for 2 min. The melt curve was 20-100°C increasing in increments of 2°C and held at each temperature for 5 seconds.

2.3.2 NMR

All NMR experiments were 1D ^{1}H-NMR and conducted using a Bruker 500 or a Bruker 600 both with a CryoProbe Prodigy located at MSU's NMR core facility. A minimum of 250 µM of protein was prepared in a solution of 90% buffer (20 mM NaH_2PO_4, 150 mM NaCl, pH 8.0) and 10% D_2O. For the Bruker 600 64 scans were run with a 2 second delay between scans and water suppression at 298K. For the Bruker 500 128 scans were run with a 2 second delay between scans and water suppression at 298 K. All data was analyzed using MestReNova software.

2.4 Crystallization

Purified Blo t 2 was used for crystallization experiments employing the sitting drop vapor diffusion method. Protein was plated at a concentration of 6.1 mg/mL in 96-well, 2 drop plates (Intelliplate). The commercial screens used for plating were JCSG+, Wizard I and II, Ammonium Sulfate, Index, Sodium Malonate, MPD, and PEG 6000 (Hampton Research). Crystals were also found in the microcentrifuge tube that Blo t 2 was stored in while in Arginine and phosphate buffer.

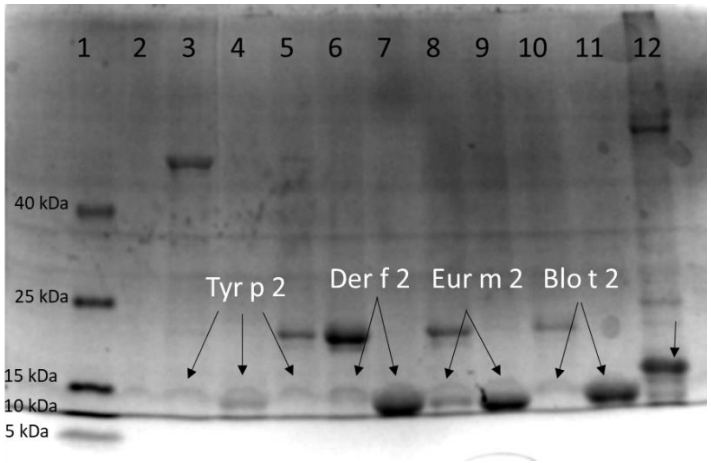
2.5 Results of Group 2 Allergens Expression

2.5.1 Protein Expression

Group 2 proteins from the mite species *D. farinae* (Der f 2)*, E. maynei* (Eur m 2)*, B. tropicalis* (Blo t 2)*,* and *T. putrescentiae* (Tyr p 2) were successfully expressed and were in the soluble fraction of the cell lysate from *E. coli* T7 SHuffle Express cells. In the case of *D. pteronyssinus* (Der p 2) protein, *E. coli* BL21(DE3) cells were used, and protein was found in the soluble fraction as well. The average quantity of protein eluted from the nickel column per liter of cell culture was as follows: Tyr p 2 was ~16 mg/L, Blo t 2 was ~15 mg/L, Eur m 2 was ~32 mg/L, and Der f 2 was ~5 mg/L. The amount of protein eluted from the nickel column per liter for Der p 2 has not yet been measured.

Expression of all proteins was confirmed through SDS-PAGE and Western Blot as seen in Figure 5.

A



B

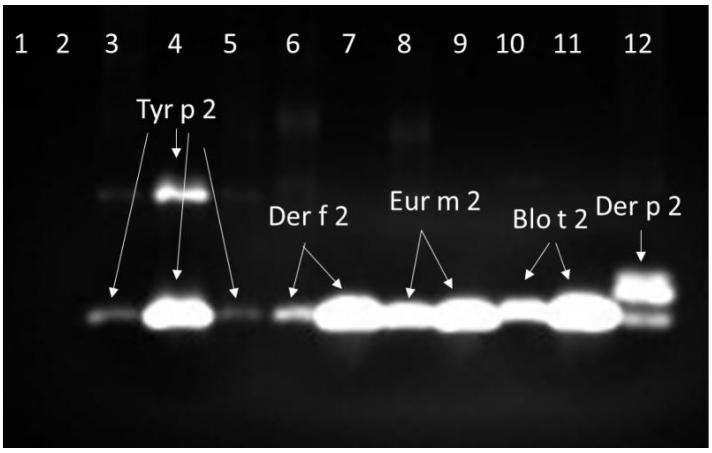


Figure 5. SDS-PAGE gel under non-reducing conditions and Western blot of all Group 2 allergens expressed and purified. (A) SDS-PAGE gel of Group 2 protein fractions collected after gel filtration. Lane 1 is a low range protein ladder; lane 2 is empty; lane 3 is a possible oligomer of Tyr p 2; lanes 5, 6, 8 and 10 are fractions collected that correspond to the size of the protein in dimeric form; lanes 4, 7, 9, and 11 are fractions collected that correspond to the size of the protein in monomeric form. (B) Western blot of all expressed and purified Group 2 proteins. The Western blot indicates that each protein was eluted from the gel filtration column in both the fraction that would correspond to the molecular weight of a dimer and in the fraction corresponding to the molecular weight of a monomer. Except for Tyr p 2 none of the proteins formed covalent dimers. Tyr p 2 has a seventh, unpaired cysteine allowing it to form an intermolecular bond between two monomers of itself.

2.5.2 Evaluation of protein Folding

DSF was used to evaluate the thermal stability or average melting temperature of all five proteins, with melting curves shown in Figure 6. Der p 2, Eur m 2, and Der f 2 had an average melting temperature of 68˚C, Tyr p 2 had an average melting temperature of 72°C, and Blo t 2 had an average melting temperature of 64°C. Not all proteins gave off the same level of fluorescence, but the individual data for each protein showed a clear transition phase, allowing for calculation of the melting temperature.

NMR data gave an indication of good folding for Blo t 2 and Der f 2 (Figure 7) with peaks in the amide proton region of 8-11 ppm and in the methyl proton region of -0.5-1.5 ppm. NMR data for Tyr p 2 indicated it was not folded and NMR experiments have not yet been performed for Eur m 2 and Der p 2.

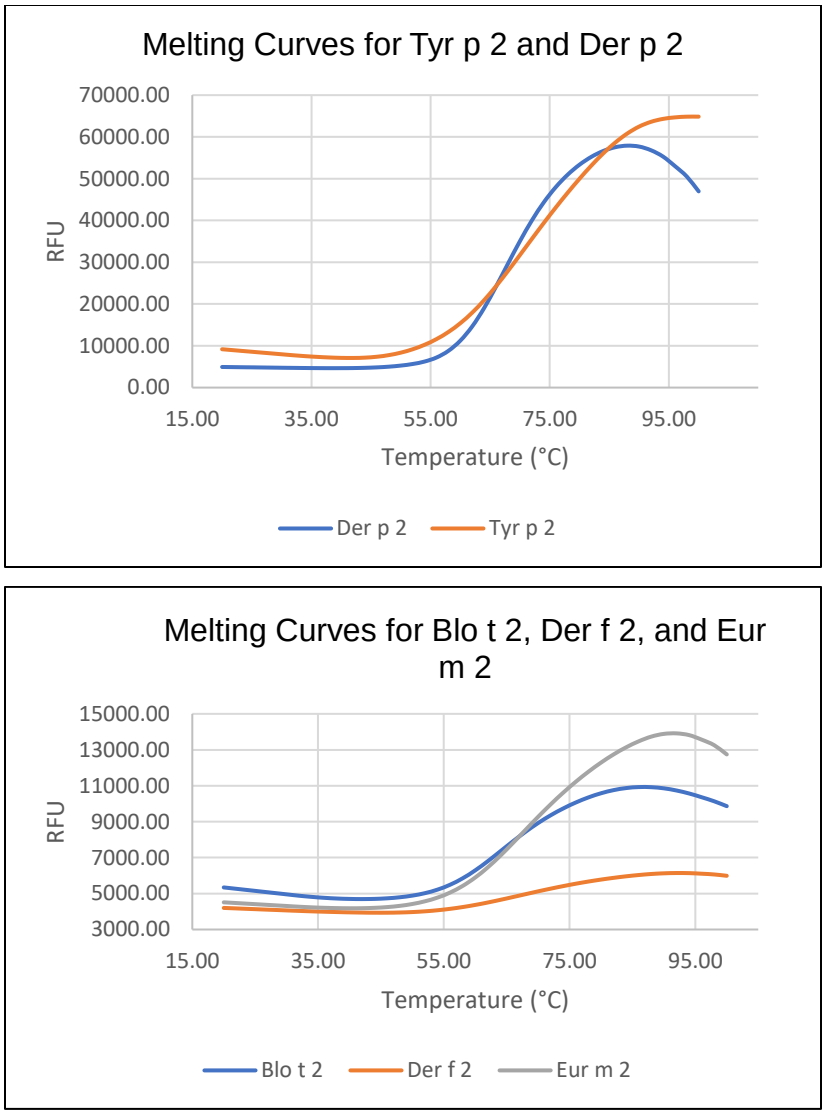
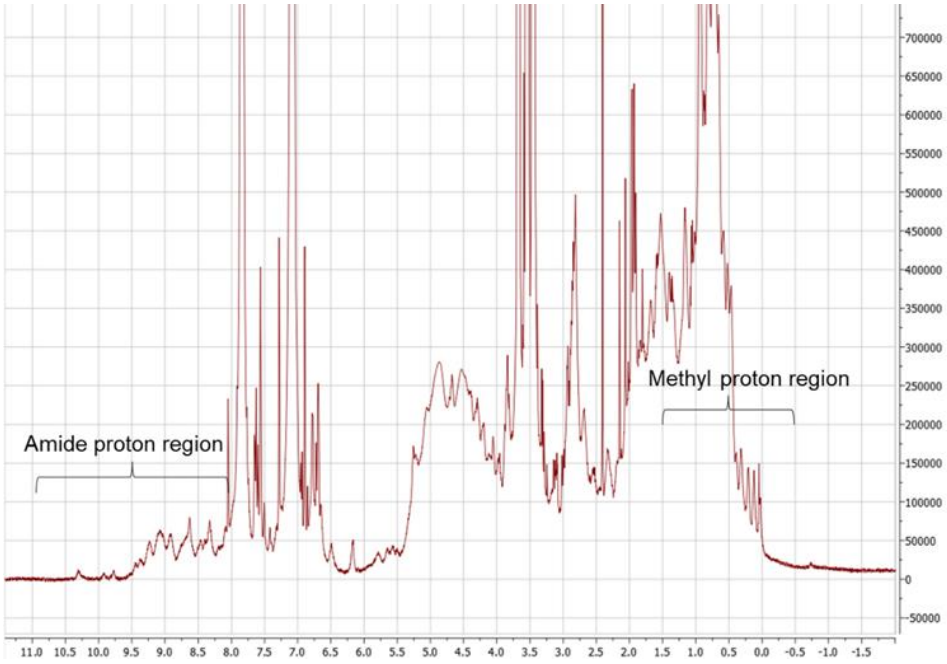


Figure 6. Representative Melting curves for Der p 2, Tyr p 2, Blo t 2, Der f 2, and Eur m 2. All experiments were performed in 50 mM Tris-HCl, 250 mM NaCl, pH 8.0.

A



B

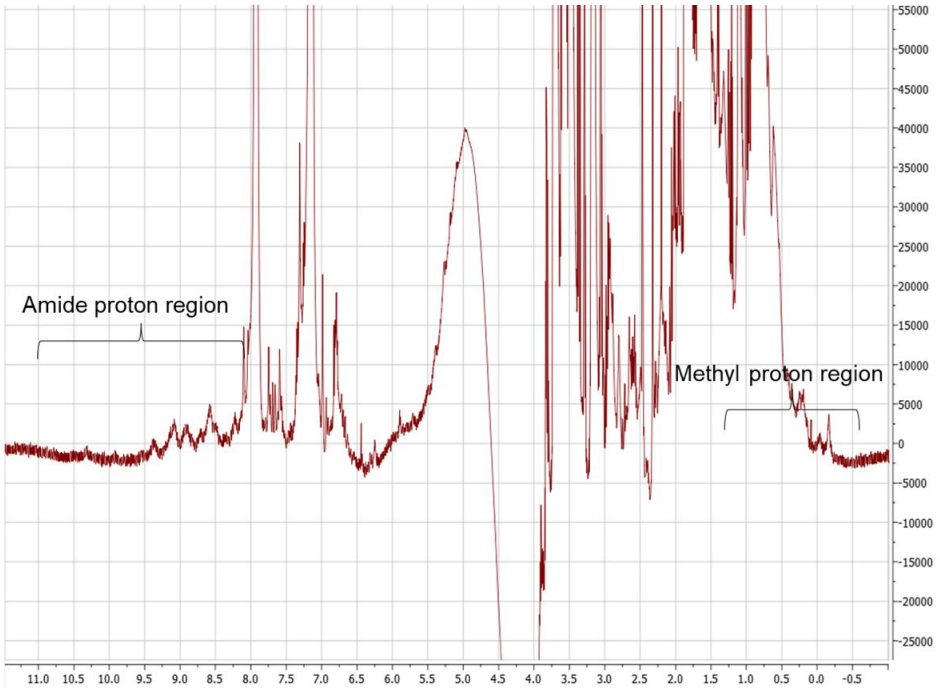


Figure 7. NMR spectra for Blo t 2 and Der f 2. (A) NMR spectrum for Blo t 2 shows peaks in the amide proton region as well as the methyl region, indicating tertiary structure. (B) NMR spectrum for Der f 2 shows peaks in the amide proton region as well as the methyl region, indicating tertiary structure.

2.6 Crystallization Results

Crystals of differing morphologies (Figure 8) were found in multiple wells from plates that had been set with JCSG+ and Wizard I and II. Crystals were collected and sent for diffraction, however most have not diffracted. The two crystals that diffracted were composed of some small molecules or salt crystals. Crystals collected from the microcentrifuge tube that Blo t 2 was stored in were also sent for diffraction, but the data suggested these too were composed of small molecules and not Blo t 2.

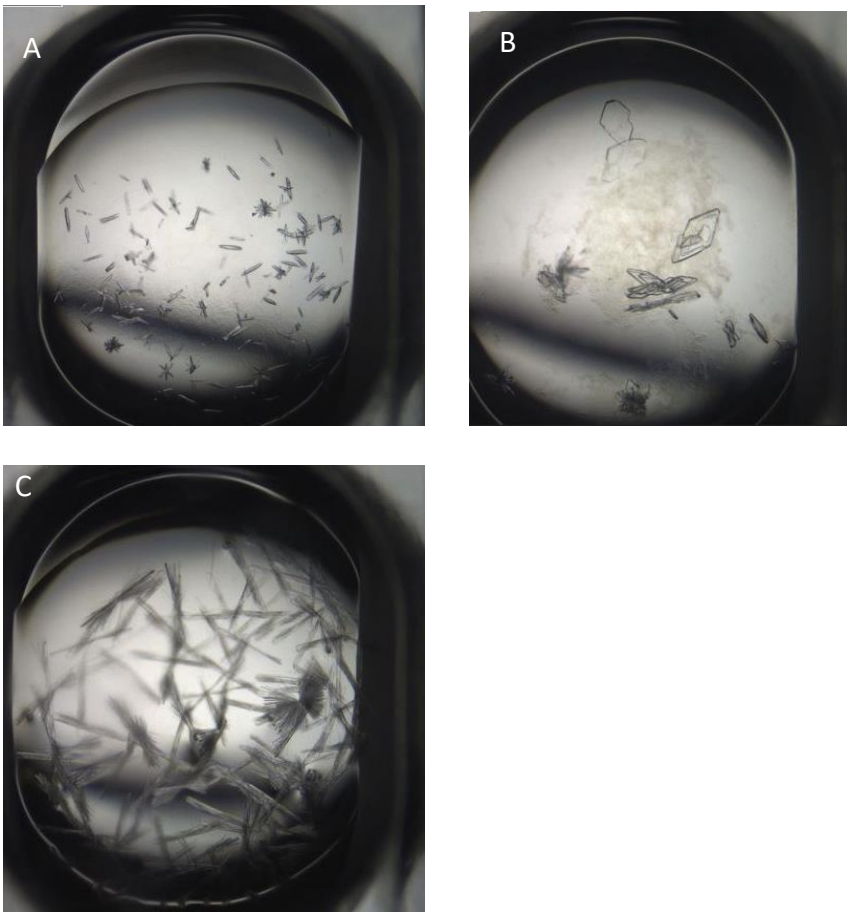


Figure 8. Example morphologies of crystals found in crystallization plates set with Blo t 2. (A) Crystals grown in JCSG+ screen: 0.14M CaCl2, 30% (v/v) glycerol, 14% (v/v) 2-propanol, 0.07M Na acetate pH 4.6. (B) Crystals grown in Wizard 1 screen: 30% (v/v) PEG 400, 0.2M Ca acetate, 0.1M Na acetate pH 4.5. (C) Crystals grown in Wizard II screen: 0.2M $MgCl_2$, 2.5M NaCl, 0.1M Tris pH 7.0.

2.7 Discussion

Escherichia coli was the first expression host of recombinant proteins and is still considered one of the major workhorses in the field. This is due to their rapid doubling time, potential for high yield, and low cost (Brondyk, 2009). However, one of the big challenges of producing recombinant protein in *E. coli* is the formation of inclusion bodies (IB). IBs are aggregates of macromolecules (predominantly protein) that form in the nucleus, cytoplasm, and/or periplasm of the host cell (Bhatwa et al., 2021). This can be particularly problematic for proteins containing cysteine residues involved in disulfide bond formation, such as those observed in Group 2 allergens. They have a high propensity for forming mismatched disulfide bonds as well as intermolecular disulfide bridges which leads to aggregation (Vincentelli et al., 2004). The general protocol for dealing with these aggregates is to treat them first with urea, or another denaturant, and after purification, to implement a refolding protocol. This process often leads to very low yields and refolding protocols are not straightforward and are often protein dependent. This proved to be the case for Group 2 allergens from mites.

When the construct containing Blo t 2 was gifted by Dr. Alain Jacquet (Chulalongkorn University, Thailand), multiple attempts were made to purify Blo t 2 from inclusion bodies by treating them first with urea then refolding by dialyzing against either Tris-HCl or PBS buffers at both pH 7.5 and pH 8.0. This method yielded very low quantities of soluble protein and what was collected showed no indication of proper folding when checked using NMR.

Since refolding attempts were unsuccessful, different *E. coli* strains were explored including C41(DE3) and LOBSTR BL21(DE3) (Andersen et al., 2013). There were also attempts to co-express Blo t 2 with protein disulfide isomerase (PDI). All attempts ended with the protein of interest predominantly in inclusion bodies and very little in the soluble fraction; proper folding was not confirmed for protein in the soluble fraction. T7 SHuffle express cells (New England Biolabs) containing DsbC, a disulfide bond isomerase that promotes the correction of mis-formed disulfide bonds (www.neb.com) were then used. Some of the protein was found in inclusion bodies, but a

significant amount was in the soluble fraction as observed on an SDS-PAGE gel. To ensure it was the protein of interest in the gel, a Western blot was run as described in 2.2, allowing for final confirmation of protein identity. Once the presence of Blo t 2 was confirmed, NMR data as well as DSF data was obtained to assess protein folding (Figures 6 and 7). NMR showed the protein was likely folded with peaks in the amide proton region of 8-11 ppm and in the methyl region of -0.5-1.5 ppm (Page et al., 2005; Poulsen, 2003). DSF was run using BODIPY FL L-Cystine (Thermo Fisher Scientific). This fluorophore was chosen because Group 2 allergens have a hydrophobic cavity making fluorophores that bind hydrophobic residues unsuitable. BODIPY FL L-Cystine is designed to fluoresce when forming mixed disulfides during thiol exchange with biomolecules(Lackman et al., 2007; Yu et al., 2003) making it suitable for Group 2 allergens which have 3 disulfide bonds. DSF gives information about the stability of the protein as well as some indication of folding. The melt curve of a protein which has not formed disulfide bridges will not have a transition phase, whereas the melt curve for a protein that has formed disulfide bridges will show the transition phase as the protein denatures. It is important to note that DSF does not give any indication of which cysteines formed disulfide bridges.

Once there was confirmation that Blo t 2 was well folded, constructs of Eur m 2, Der f 2, and Tyr p 2 were ordered. All recombinant proteins were produced using the same protocols and checks as described in Sections 2.2 and 2.3 with mixed results. The presence of protein in the soluble fraction of cell lysate for all three constructs was confirmed by SDS-PAGE and Western Blot (Figure 5) and DSF data was successfully collected (Figure 6). Obtaining NMR data, however, proved to be more difficult. NMR data was collected for Der f 2 (Figure 7) and there were indications of folding, but Eur m 2 has not yet been purified in sufficient quantity to obtain NMR data, as protein expression is significantly lower than for the other Group 2 proteins.

Tyr p 2 has proven to be the most difficult to express consistently. It has an odd number of cysteines, making it likely for the protein to form dimers by means of an intermolecular disulfide bridge during purification and/or storage (Vincentelli et al., 2004). There are indications this is happening as the Western Blot has a band at the appropriate weight of slightly more than 25 kDa (Figure 5, lane 4). This band is not

readily visible on the SDS-PAGE gel likely due to the differing detectable limits of Coomassie stain and Clarity™ Western ECL substrate; Coomassie stain has a limit in the microgram region, whereas the substrate can detect femtograms of protein. Tyr p 2 also has a higher theoretical pI than the other Group 2 allergens at 8.1. It was because of this that Tyr p 2 was initially purified using buffers at pH 9.0; however, very little protein was detected in the elution from the nickel column, and what little protein was present aggregated during purification on the weak anion exchange column. When using buffers at pH 8.0, Tyr p 2 was successfully purified and run on both the weak anion exchange column as well as the gel filtration column. This batch of protein was used to check for thermal stability and folding by NMR. What was interesting was that Tyr p 2 showed the highest thermal stability of all the Group 2 proteins but the NMR spectra indicated it was not folded. This may be due to aggregation as the protein was at a significantly higher concentration for NMR experiments than DSF. Small amounts of amphiphilic compounds (glycerol) or reducing agents may need to be added to the buffer when purifying Tyr p 2 to help prevent aggregation and encourage proper disulfide bond formation (Vincentelli et al., 2004).

Der p 2 was co-expressed with PDI by a previous grad student from *E. coli* BL21(DE3) cells (Kapingidza, 2020). Current attempts at expression using the same construct and same methodology have been successful but not consistently so. Sometimes expression and purification yield protein and sometimes not. However, enough protein was obtained to run a DSF experiment (Figure 6).

2.8 Summary and Future Directions

Proteins that require proper disulfide bond formation for stability are notoriously difficult to produce in *E. coli.* Because of this, one of the most common expression systems for Group 2 allergens is yeast, as they have the machinery required for proper disulfide bond formation (Ma et al., 2020). The protein expression and purification methods in this experiment used an *E. coli* strain with an isomerase that promotes correct formation of disulfide bonds, eliminating the need for denaturants. This method was shown to be successful for Der f 2 and Blo t 2 while confirmation of folding is still needed for Eur m 2 and Der p 2. And while Tyr p 2 was found in the soluble fraction of

the cell lysate, NMR suggests it is not folded and functional and further optimization is required.

Moving forward, this method will be tested on Group 2 allergens from SMs *Lepidoglyphus destructor* and *Glycyphagus domesticus,* and HDM *Dermatophagoides maynei.* Any proteins that can be demonstrated to have proper folding will be used for crystallographic studies and immunological studies.

CHAPTER 3

SMALL MOLECULE LIGAND BINDING TO GROUP 2 ALLERGENS

3.1 Introduction to Ligand Binding Studies

As previously mentioned, the physiological role that Group 2 allergens play for mites is unknown. However, they have been shown to bind cholesterol, LPS, and other lipids (Reginald & Chew, 2019) and we hypothesize that discovering the natural ligands of these proteins will provide insight into how they invoke allergy in humans. The two methods employed in this study to ascertain what molecules may bind in the hydrophobic pocket of Group 2 allergens were molecular docking experiments using AutoDock Vina and ChimeraX (Goddard et al., 2018; Pettersen et al., 2021; Trott & Olson, 2010), and competition-based assays which use a fluorescent probe to determine binding.

Molecular docking is usually performed between a small molecule and a macromolecule with the goal of understanding and predicting their interactions both structurally and energetically (Morris & Lim-Wilby, 2008). It can be used to screen large libraries of molecules against the target macromolecule to find appropriate targets for further experimentation. There are many tools available for docking experiments but for this study AutoDock Vina was used. AutoDock Vina is designed to predict noncovalent binding of macromolecules (receptors) and small molecules (ligands) using a scoring function to approximate the standard chemical potentials of the system with a more negative score equating to a stronger binding affinity (Trott & Olson, 2010). Docking experiments were performed in this study to find ligands that may aid in crystallographic studies, as well as potential targets for further investigation in fluorescence-based competition assays.

Competition binding assays are a common method for measuring the binding affinity of a ligand to a receptor. These assays often use ligands that have been labelled with a fluorescent or radioactive tag to measure initial binding to a protein, and to track their displacement in the presence of varying concentrations of non-labelled ligands (Hu et al., 2013). The competition assay employed in this study used a similar approach but instead of using a labelled ligand, a fluorescent probe, 11-(Dansylamino) undecanoic acid (DAUDA), was used. DAUDA is a fluorescent fatty acid probe the contains an 11-

carbon acyl chain and a polarity-sensitive dansyl-type fluorophore that has up to a 60-fold increase in fluorescence when binding certain proteins (Wang et al., 2011). It has been used by other researchers to determine the affinity of compounds with long carbon chains (like fatty acids) for proteins by measuring the fluorescence of the protein, probe, ligand complex at excitation and emission wavelengths of 345 nm and 543 nm respectively (Wang et al., 2011). The goal of our studies was to determine what fatty acids or other small hydrophobic molecules may bind in the pocket of Group 2 allergens, and if so, do any display higher affinities than others.

3.2 Materials and Methods

3.2.1 *In Silico* Studies

Using Autodock Vina (Trott & Olson, 2010), arachidonic acid, butyric acid, cholesterol sulfate, lauric acid, myristic acid, oleic acid, palmitic acid, pinolenic acid, and stearic acid were docked onto Der p 2 and Blo t 2 AlphaFold (Varadi et al., 2022) generated structure. All ligands and proteins were prepared for docking using AutoDock Tools 1.5.7. For proteins all waters were removed and Kollman charges were added as well as non-polar hydrogens. For each ligand the proper torsion parameters were added according to the software protocol. Blind docking was performed for each protein ligand pair using grid box parameters 40×40×40 Å centered on the coordinates x = 1.269, y = 4.175, and z = -0.292 which encompassed most of the binding pocket and some external residues. All experiments were run in triplicate with an energy range of 4 and an exhaustiveness of 8. Appropriateness of the grid box parameters was assessed by comparing the x-ray crystal structure of human NPC2 bound to cholesterol sulfate (PDB 5KWY), against a docked version of the same complex using the above described methods. All results were analyzed using ChimeraX (Goddard et al., 2018; Pettersen et al., 2021).

3.2.2 DAUDA Binding Assay

Blo t 2 and Der p 2 binding to 11-(Dansylamino) undecanoic acid (DAUDA) (Caymen Chemical Company) was assessed using fluorescence measurement with a Biotek Synergy Neo microplate reader (Agilent) located in MSU's Assay Development and Drug Repurposing core. The excitation wavelength was 345 nm and 380-600 nm sweeping emission filters and a gain of 115. 10 μM of protein was added to 2 μM

DAUDA in a total reaction volume of 50 µL with a final DMSO concentration of 0.1%. The buffer used to bring the reaction volume to 50 µL was 50 mM Tris-HCl, 150 mM NaCl at pH 8.0. The reaction mixture was incubated for 15 minutes at room temperature in a 96 well black, flat bottom plate (Nalge NUNC International Corporation) and all reactions were performed in triplicate (Min et al., 2023). A series of measurements with DUADA alone were run in the same plate to establish baseline fluorescence, as well as a series of measurements with protein alone to establish any intrinsic fluorescence of the proteins. Initial data plotting was performed using BioTek Gen5 software and final analysis was performed using Microsoft Excel.

3.2.3 Ligand Screening

Der p 2 and Blo t 2 ligand binding were assessed using DAUDA displacement assay with various fatty acids (arachidonic, butyric, lauric, myristic, oleic, palmitic, stearic (Sigma Aldrich); pinoleic (Cayman Chemical Company)), and cholesterol (Sigma Aldrich). To a 96 well black, flat bottom plate (Nalge NUNC International Corporation) was added 10 µM of protein and 2 µM of DAUDA, then 10 µM of ligand. The reaction mixture was incubated at room temperature for 1 hour. Data was collected with Biotek Synergy Neo microplate reader (Agilent) with 345 nm excitation and 380-600 nm sweeping emission filters and a gain of 115. All reactions were run in triplicate. Initial data plotting was performed using Biotek Gen5 software and final analysis was performed using Microsoft Excel.

3.3 Results

3.3.1 *In Silico* Studies

The results of the *in silico* studies are based on the scoring function used by AutoDock Vina which is inspired by X-score, but goes beyond linear regression in its tuning and ranks conformations not only by the intermolecular contributions but also by intramolecular ones (Trott & Olson, 2010). Results are given in kcal/mol with a more negative score being equivalent to the tightest binding and best fit. When selecting which conformations to include in our results we first looked at if the ligand was in the binding pocket, second where it ranked, and third the upper bound root mean square deviation from the best fit. Based on these parameters, the docking results indicate that only Der p 2 is capable of binding DAUDA and it displayed binding to all fatty acids

except cervonic acid and cholesterol. Blo t 2 displayed binding to all fatty acids except arachidonic acid, cholesterol, and oleic acid. (Table 1.) The results suggest both proteins are capable of binding a variety of fatty acids with no obvious preference for one kind over another.

Table 1. Summary of *In silico* experiments. Each number represents the number of theoretical binding conformations that that ligand had in the pocket of the respective protein. Conformations were accepted by looking first at the conformation with the strongest affinity – this became the reference conformation. Then any conformations with the ligand in the binding pocket that had an upper bound root mean square deviation (RMSD) <2 Å from the reference were also accepted.

	Der p 2	**Blo t 2**
Fatty Acid or Probe	# of Theoretical Binding Conformations	# of Theoretical Binding Conformations
DAUDA	3	-
Arachidonic Acid (20:4)	3	-
Butyric Acid (4:0)	3	5
Cervonic Acid (22:6)	-	1
Cholesterol Sulfate	-	-
Lauric Acid (12:0)	3	2
Linoleic Acid (18:2)	2	2
Myristic Acid (14:0)	2	1
Oleic Acid (18:1)	1	-
Palmitic Acid (16:0)	3	1
Pinolenic Acid (18:3)	3	2

3.3.2 DAUDA Assay and Ligand Displacements Assays

Initial results of both the DAUDA assay and ligand displacement assay showed good binding of the probe and protein, as well as binding of both oleic and pinolenic acid to the protein. This was evidenced by an increase in fluorescence when the probe and protein were incubated together, and a subsequent decrease in fluorescence after the ligand was added as seen in Figure 9.

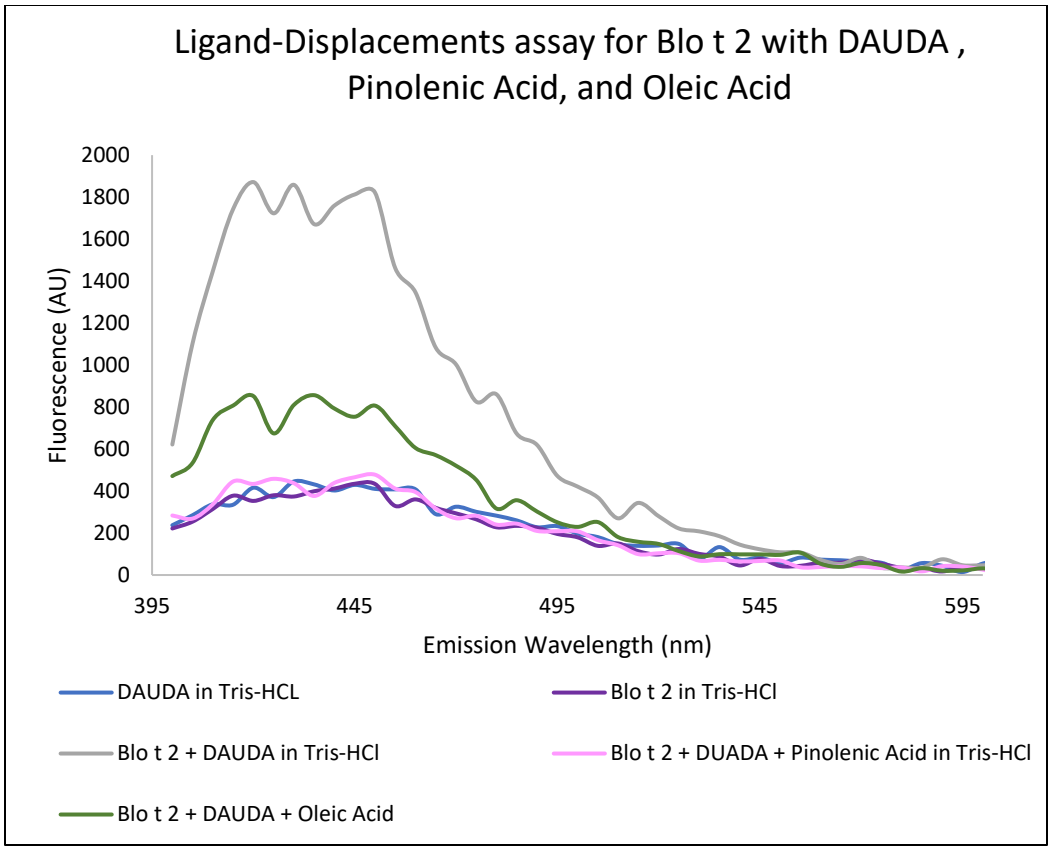


Figure 9. Fatty acid binding to Blo t 2. Binding of Blo t 2 to oleic acid and pinolenic acid was screened using DAUDA-displacement assay. The reduction of fluorescence in the presence of the fatty acid indicates binding of the ligand to the protein.

3.4 Discussion

Molecular docking and competitive binding assays are common tools employed to assess protein-ligand interactions. In this study AutoDock Vina was used to dock nine fatty acids of varying length and saturation, cholesterol, and DAUDA to Der p 2 and Blo t 2. Before docking could begin, both the protein and ligand of interest had to be prepared as well as a docking site selected. These were accomplished according to the protocol described in 3.2.1. The appropriateness of the grid box, or docking site, was assessed by comparing the X-ray crystal structure of cholesterol sulfate bound human NPC2 (a homologous protein to Group 2 allergens) with a docked version of the same complex. As seen in figure 10, AutoDock Vina placed cholesterol sulfate in the same region of NPC2 as in the crystal structure and in a near identical orientation. As the grid box encompassed most of what is known to be the binding site of Blo t 2 and Der p 2,

and the docking results of cholesterol sulfate to NPC2 were satisfactory, the parameters for docking studies were left at 40x40x40 Å centered on the coordinates x = 1.269, y = 4.175, and z = -0.292.

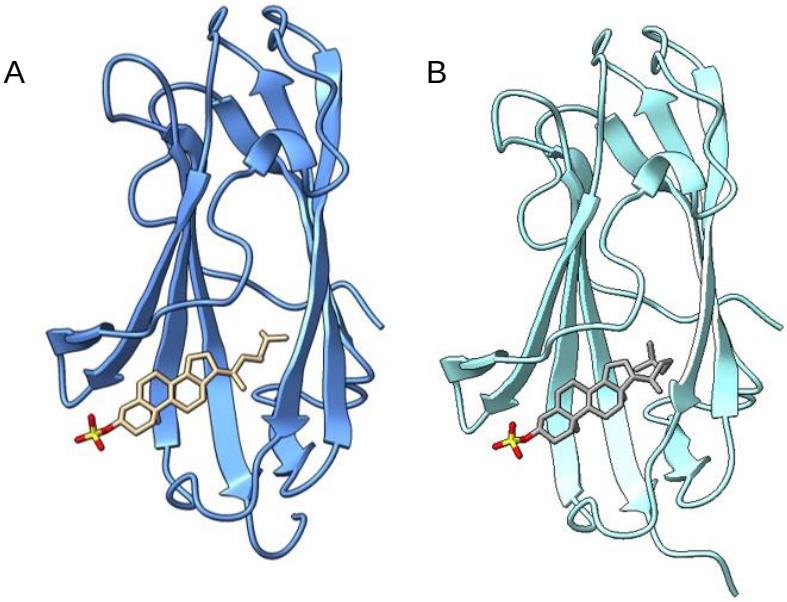


Figure 10. Cholesterol sulfate in the binding pocket of human NPC2. (A) X-ray crystal structure of the complex. (B) AutoDock Vina results of the same complex.

All results from AutoDock Vina were assessed using the docking analysis tool in ChimeraX. As each experiment was run in triplicate, so there were three data sets for each protein/ligand pair. ChimeraX allowed for the visualization and assessment of all three data sets for each pair at the same time. For each data set, the conformation with the strongest binding affinity had the most negative score and became the reference point for determining the root mean square deviation (RMSD) for all other conformations in that data set. Scores in ChimeraX are equivalent to the binding affinity determined by AutoDock Vina which is measured in kcal/mol. Conformations accepted as plausible were those where the ligand was in the binding pocket of the protein and had either the strongest binding affinity (most negative score) or had an upper bound RMSD <2 Å from the reference conformation. The upper bound RMSD is an atom-atom comparison of conformations meaning the same atom in both conformations are compared (https://www.cgl.ucsf.edu/chimerax/docs/user/commands/rmsd.html). The results, as seen in Table 1, indicate that DAUDA should not bind Blo t 2 at all but should have

some binding to Der p 2. And, although there is some difference in which fatty acids would bind each protein, the most interesting results were those for cholesterol. As one study conducted a pull-down assay and found that cholesterol had a higher affinity for the binding pocket of Der p 2 than other lipids and using liquid chromatography-mass spectrometry “confirmed that cholesterol is the natural ligand of Der p 2 (Reginald & Chew, 2019).”

Next, we wanted to see how our docking results compared to competitive binding assays; would the same lipids bind, would there be any discernible preferential binding for one kind of lipid over another for each protein, and would cholesterol bind? DAUDA was chosen as the fluorescent probe as it had been used by our collaborators in Bangkok who had done similar assays with Der p 2, and the first fatty acids tested were oleic acid and pinolenic acid. Oleic acid had also been successfully used by our collaborators at Chulalongkorn University, and pinolenic acid was chosen because it has the same number of carbons but has two more double bonds. The controls for the first experiment were DAUDA in the buffer and the protein of interest in the buffer both to measure and background fluorescence. The results of the first assay as seen in Figure 9 show increased fluorescence when DAUDA and Blo t 2 are combined from either alone indicated binding. A subsequent decrease in fluorescence was noted when either fatty acid ligand was added to the reaction indicated the displacement of DAUDA. What was of particular interest was that the presence of pinolenic acid produced a greater decrease in fluorescence than oleic acid, indicating a potential higher binding affinity of Blo t 2 for pinolenic acid than oleic acid.

We wanted to investigate further and see if Group 2 allergens really displayed a preference for lipids with more rigidity over those with less and if the length of the carbon chain played a role in binding. So, 7 more fatty acids were selected (arachidonic, butyric, lauric, linoleic, myristic, palmitic, and stearic acid) that have carbon chain lengths varying from 4-22 and saturation levels from 0-6, as well as cholesterol. (Group 2 allergens are found in the guts of mites so most of the fatty acids selected are ones found in the food of both HDMs and SMs.) We also wanted to investigate any possible differences in HDM Group 2 allergen binding vs. SM allergens, as well as the impact of changes in the pH of the buffer. Unfortunately, we were not able to obtain any

consistent data from these assays. We began to observe non-specific binding of DAUDA to Blo t 2 and Der p 2, as well as non-specific binding of DAUDA to some of the fatty acids (not shown). A different batch of Blo t 2 was used for the second round of assays than the first so the protein used may not have been as well folded, but since DAUDA showed increased binding in the presence of some of the fatty acids when no protein was present a different probe will likely need to be found.

3.5 Summary and Future Directions

Molecular docking and competitive binding assays were utilized to study the binding of small molecule ligands to Group 2 allergens Blo t 2 and Der p 2. *In silico* studies indicate that neither bind cholesterol, but that they will bind various fatty acids without any clear indication of preference for one kind of fatty acid over another. Initial competitive binding assays showed deviation from the *in silico* studies in that Blo t 2 did bind DAUDA as well as oleic acid. The initial assay also indicated that Blo t 2 may preferentially bind pinolenic acid over oleic acid as seen by the greater decrease in fluorescence when pinolenic acid was present. As further attempts at competitive binding assays did not yield consistent results and displayed non-specific binding of DAUDA to both the proteins and the fatty acid ligands, further optimization is required before any conclusions can be drawn.

In the future we want to quantitatively measure the binding of fatty acids to Group 2 allergens not only through competitive binding assays, but also by utilizing thermophoresis techniques. We also want to investigate what residues are involved in small molecule binding through the use of 2-D NMR and X-ray crystallographic studies. In terms of immunological studies we want to test the binding affinity of ligand bound allergens to their known IgE antibody as compared to the apo form bound to the same IgE.

CHAPTER 4

CONCLUSION

Mites are found in nearly every ecosystem on earth and are so tiny they generally go unnoticed (Bensoussan et al., 2016; Coddington & Colwell, 2001). However, some can cause major health problems for humans. Domestic mites, which include house dust mites and storage mites are one of the leading sources of indoor aeroallergens around the globe (Jeffrey D. Miller, 2019). It is not the mites themselves that cause allergy, but various proteins found in their fecal pellets, saliva, skin sheds, and eggs that induce allergy (Bessot & Pauli, 2011; Khatri et al., 2023) Of the 128 proteins produced by mites that are known to cause allergery, proteins belonging to Group 1 and 2 are considered major allergens. These proteins induce a Type 1 hypersensitization response which occurs in 2 phases: a sensitization phase and an effect phase (Abbas et al., 2023; Calderon et al., 2015) More recent studies indicate that it is not the proteins alone that are the cause of allergy, but small molecules that bind these proteins likely play a role in the sensitization phase and/or the effect phase.

This research focused on Group 2 allergens and their interactions with small molecule ligands. The goal was to structurally characterize Group 2 allergens from both house dust mites and storage mites using X-ray crystallography techniques as only Der p 2 and Der f 2 from *D. pteronyssinus* and *D. farinae* respectively, have structures deposited in PDB. In order to conduct crystallographic studies protein was required in large quantities, so methods had to be developed for the production of these proteins in *E. coli.* Der f 2, Tyr p 2, Blo t 2, and Der p 2 were successfully expressed and purified from the soluble fraction of the cell lysate when using *E. coli* T7 SHuffle express cells. Blo t 2 was plated with several different commercial crystallization screens and crystals were collected and sent for diffraction. Some of the crystals turned out to be salt crystals or other small molecule crystals, but some did not diffract. Further optimization of crystallization conditions is required, and there are plans to set plates for the other Group 2 allergens purified.

The second part of the project was to investigate the binding of Group 2 allergens to small molecule ligands as knowing what they bind will provide insights into their ability to cause allergic disease. The methods employed were molecular docking

using AutoDock Vina and ChimeraX, and competitive binding assays using DAUDA, a fluorescent probe. Molecular docking experiments showed very little difference in fatty acid binding to Der p 2 vs. Blo t 2, but interestingly, it indicated that DAUDA would not bind Blo t 2 and that neither protein would bind cholesterol sulfate. When comparing these results to competitive binding assays, the initial results showed good binding of Blo t 2 to DAUDA and possible preferential binding to pinolenic acid over oleic acid. Unfortunately, further experiments gave no other insights as they showed large variability in binding and DAUDA displayed non-specific binding to some of the fatty acids. Further optimization is needed.

REFERENCES

Abbas, M., Moussa, M., & Akel, H. (2023). *Type 1 Hypersensitivity Reaction*. Stat Pearls.

Andersen, K. R., Leksa, N. C., & Schwartz, T. U. (2013). Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. *Proteins*, *81*(11), 1857-1861. https://doi.org/10.1002/prot.24364

Arlian, L. G., Neal, J. S., Morgan, M. S., Vyszenski-Moher, D. L., Rapp, C. M., & Alexander, A. K. (2001). Reducing relative humidity is a practical way to control dust mites and their allergens in homes in temperate climates. *J Allergy Clin Immunol*, *107*(1), 99-104. https://doi.org/10.1067/mai.2001.112119

Azad, A. F. (1986). *Vector Control Series: Mites Training and Information Guide*. World Health Organization.

Bensoussan, N., Santamaria, M. E., Zhurov, V., Diaz, I., Grbic, M., & Grbic, V. (2016). Plant-Herbivore Interaction: Dissection of the Cellular Pattern of Tetranychus urticae Feeding on the Host Plant. *Front Plant Sci*, *7*, 1105. https://doi.org/10.3389/fpls.2016.01105

Bertani, B., & Ruiz, N. (2018). Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus*, *8*(1). https://doi.org/10.1128/ecosalplus.ESP-0001-2018

Bessot, J., & Pauli, G. (2011). Mite allergens: an overview. *European annals of allergy and clinical immunology*, *43*(5), 141.

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X. Z., & Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Front Bioeng Biotechnol*, *9*, 630551. https://doi.org/10.3389/fbioe.2021.630551

Brondyk, W. H. (2009). Selecting an appropriate method for expressing a recombinant protein. *Methods in enzymology*, *463*, 131-147. https://doi.org/10.1016/s0076-6879(09)63011-1

Calderon, M. A., Linneberg, A., Kleine-Tebbe, J., De Blay, F., Hernandez Fernandez de Rojas, D., Virchow, J. C., & Demoly, P. (2015). Respiratory allergy caused by house dust mites: What do we really know? *J Allergy Clin Immunol*, *136*(1), 38-48. https://doi.org/10.1016/j.jaci.2014.10.012

Chruszcz, M., Chew, F. T., Hoffmann-Sommergruber, K., Hurlburt, B. K., Mueller, G. A., Pomes, A., Rouvinen, J., Villalba, M., Wohrl, B. M., & Breiteneder, H. (2021). Allergens and their associated small molecule ligands-their dual role in sensitization. *Allergy*, *76*(8), 2367-2382. https://doi.org/10.1111/all.14861

Coddington, J. A., & Colwell, R. K. (2001). Arachnids. In S. A. Levin (Ed.), *Encylopedia of Biodiversity*.

Durham, S. R., & Shamji, M. H. (2023). Allergen immunotherapy: past, present and future. *Nat Rev Immunol*, *23*(5), 317-328. https://doi.org/10.1038/s41577-022-00786-1

Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci*, *27*(1), 14-25. https://doi.org/10.1002/pro.3235

Gour, N., & Wills-Karp, M. (2015). IL-4 and IL-13 signaling in allergic airway disease. *Cytokine*, *75*(1), 68-78. https://doi.org/10.1016/j.cyto.2015.05.014

Hofmann, L., Gulati, S., Sears, A., Stewart, P. L., & Palczewski, K. (2016). An effective thiol-reactive probe for differential scanning fluorimetry with a standard real-time polymerase chain reaction device. *Anal Biochem*, *499*, 63-65. https://doi.org/10.1016/j.ab.2016.01.016

Hu, L., Yang, L., Lipchik, A. M., Geahlen, R. L., Parker, L. L., & Tao, W. A. (2013). A quantitative proteomics-based competition binding assay to characterize pITAM-protein interactions. *Anal Chem*, *85*(10), 5071-5077. https://doi.org/10.1021/ac400359t

Ichikawa, S., Hatanaka, H., Yuuki, T., Iwamoto, N., & Kojima, S. (1998). Solution Structure of Der f 2, the Major Mite Allergen for Atopic

Diseases. *Journal of Biological Chemistry*, *273*(1), 356-360.

Kapingidza, A. B. (2020). *Structural, Functional, and Immunological Studies of Group 1 and 2 House Dust Mite and Profilin Allergens for the Development of Allergen Specific Diagnostics and Immunotherapy* [Doctoral Dissertation, University of South Carolina].

Khatri, K., O'Malley, A., Linn, C., Kowal, K., & Chruszcz, M. (2023). Role of Small Molecule Ligands in IgE-Mediated Allergy. *Curr Allergy Asthma Rep*, *23*(9), 497-508. https://doi.org/10.1007/s11882-023-01100-2

Kim, I. S. (2023). New approaches to immunotherapy in house dust mite allergy. *Clin Exp Pediatr*, *66*(4), 161-168. https://doi.org/10.3345/cep.2022.00479

Krantz, G. W. (1978). *A Manual of Acarology* (2 ed.). Oregon State University Book Stores, Inc.

Lackman, R. L., Jamieson, A. M., Griffith, J. M., Geuze, H., & Cresswell, P. (2007). Innate immune recognition triggers secretion of lysosomal enzymes by macrophages. *Traffic*, *8*(9), 1179-1189. https://doi.org/10.1111/j.1600-0854.2007.00600.x

Larche, M., Akdis, C. A., & Valenta, R. (2006). Immunological mechanisms of allergen-specific immunotherapy. *Nat Rev Immunol*, *6*(10), 761-771. https://doi.org/10.1038/nri1934

Ma, Y., Lee, C. J., & Park, J. S. (2020). Strategies for Optimizing the Production of Proteins and Peptides with Multiple Disulfide Bonds. *Antibiotics (Basel)*, *9*(9). https://doi.org/10.3390/antibiotics9090541

Miller, J. D. (2019). The Role of Dust Mites in Allergy. *Clinical Reviews in Allergy & Immunology*, *57*(3), 312-329. https://doi.org/10.1007/s12016-018-8693-0

Miller, J. D. (2019). The Role of Dust Mites in Allergy. *Clin Rev Allergy Immunol*, *57*(3), 312-329. https://doi.org/10.1007/s12016-018-8693-0

Min, J., Foo, A. C. Y., Gabel, S. A., Perera, L., DeRose, E. F., Pomés, A., Pedersen, L. C., & Mueller, G. A. (2023). Structural and ligand binding analysis of the pet allergens Can f 1 and Fel d 7 [Original Research]. *Frontiers in Allergy*, *4*. https://doi.org/10.3389/falgy.2023.1133412

Morris, G. M., & Lim-Wilby, M. (2008). Molecular docking. *Methods Mol Biol*, *443*, 365-382. https://doi.org/10.1007/978-1-59745-177-2_19

Page, R., Peti, W., Wilson, I. A., Stevens, R. C., & Wüthrich, K. (2005). NMR screening and crystal quality of bacterially expressed prokaryotic and eukaryotic proteins in a structural genomics pipeline. *Proceedings of the National Academy of Sciences*, *102*(6), 1901-1905. https://doi.org/doi:10.1073/pnas.0408490102

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci*, *30*(1), 70-82. https://doi.org/10.1002/pro.3943

Pomes, A., Mueller, G. A., & Chruszcz, M. (2020). Structural Aspects of the Allergen-Antibody Interaction. *Front Immunol*, *11*, 2067. https://doi.org/10.3389/fimmu.2020.02067

Poulsen, F. M. (2003). A brief introduction to NMR spectroscopy of proteins.

Reginald, K., & Chew, F. T. (2019). The major allergen Der p 2 is a cholesterol binding protein. *Sci Rep*, *9*(1), 1556. https://doi.org/10.1038/s41598-018-38313-9

Sanchez-Borges, M., Martin, B. L., Muraro, A. M., Wood, R. A., Agache, I. O., Ansotegui, I. J., Casale, T. B., Fleisher, T. A., Hellings, P. W., Papadopoulos, N. G., Peden, D. B., Sublett, J. L., Tilles, S. A., & Rosenwasser, L. (2018). The importance of allergic disease in public health: an iCAALL statement. *World Allergy Organ J*, *11*(1), 8. https://doi.org/10.1186/s40413-018-0187-2

Scheiner, O. (1993). Molecular and functional characterization of allergens: fundamental and practical aspects. *Wien Klin Wochenschr.*, *105*, 653-658.

Stavnezer, J., & Schrader, C. E. (2014). IgH chain class switch recombination: mechanism and regulation. *J Immunol*, *193*(11), 5370-5378. https://doi.org/10.4049/jimmunol.1401849

Strobl, M. R., Demir, H., Sanchez Acosta, G., Drescher, A., Kitzmuller, C., Mobs, C., Pfutzner, W., & Bohle, B. (2023). The role of IgG(1) and IgG(4) as dominant IgE-blocking antibodies shifts during allergen immunotherapy. *J Allergy Clin Immunol*, *151*(5), 1371-1378 e1375. https://doi.org/10.1016/j.jaci.2023.01.005

Thomas, W. R. (2015). Hierarchy and molecular properties of house dust mite allergens. *Allergol Int*, *64*(4), 304-311. https://doi.org/10.1016/j.alit.2015.05.004

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*, *31*(2), 455-461. https://doi.org/10.1002/jcc.21334

Tscheppe, A., & Breiteneder, H. (2017). Recombinant Allergens in Structural Biology, Diagnosis, and Immunotherapy. *Int Arch Allergy Immunol*, *172*(4), 187-202. https://doi.org/10.1159/000464104

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Zidek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., . . . Velankar, S. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*, *50*(D1), D439-D444. https://doi.org/10.1093/nar/gkab1061

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., & Bignon, C. (2004). High-throughput automated refolding screening of inclusion bodies. *Protein Sci*, *13*(10), 2782-2792. https://doi.org/10.1110/ps.04806004

Vitte, J., Vibhushan, S., Bratti, M., Montero-Hernandez, J. E., & Blank, U. (2022). Allergy, Anaphylaxis, and Nonallergic Hypersensitivity: IgE, Mast Cells, and Beyond. *Med Princ Pract*, *31*(6), 501-515. https://doi.org/10.1159/000527481

Wanderlingh, U., Branca, C., Crupi, C., Conti Nibali, V., La Rosa, G., Rifici, S., Ollivier, J., & D'Angelo, G. (2017). Molecular Dynamics of POPC Phospholipid Bilayers through the Gel to Fluid Phase Transition: An Incoherent Quasi-Elastic Neutron Scattering Study. *Journal of Chemistry*, *2017*, 1-8. https://doi.org/10.1155/2017/3654237

Wang, Y., Luo, Z., Shi, X., Wang, H., Nie, L., & Huang, M. (2011). A fluorescent fatty acid probe, DAUDA, selectively displaces two myristates bound in human serum albumin. *Protein Sci*, *20*(12), 2095-2101. https://doi.org/10.1002/pro.749

Xu, H., Shu, H., Zhu, J., & Song, J. (2019). Inhibition of TLR4 inhibits allergic responses in murine allergic rhinitis by regulating the NF-kappaB pathway. *Exp Ther Med*, *18*(1), 761-768. https://doi.org/10.3892/etm.2019.7631

Xu, Y., Zhang, Q., Tan, L., Xie, X., & Zhao, Y. (2019). The characteristics and biological significance of NPC2: Mutation and disease. *Mutat Res Rev Mutat Res*, *782*, 108284. https://doi.org/10.1016/j.mrrev.2019.108284

Yu, Z., Eaton, J. W., & Persson, H. L. (2003). The radioprotective agent, amifostine, suppresses the reactivity of intralysosomal iron. *Redox Rep*, *8*(6), 347-355. https://doi.org/10.1179/135100003225003384

Yuriev, S., Rodinkova, V., Mokin, V., Varchuk, I., Sharikadze, O., Marushko, Y., Halushko, B., & Kurchenko, A. (2023). Molecular sensitization pattern to house dust mites is formed from the first years of life and includes group 1, 2, Der p 23, Der p 5, Der p 7 and Der p 21 allergens. *Clin Mol Allergy*, *21*(1), 1. https://doi.org/10.1186/s12948-022-00182-z