Structural biology of mite allergens

Yubao Cui

Abstract Mite allergens contribute to a significant proportion of human allergic symptoms, including asthma and rhinitis. The development of therapies to treat and prevent these symptoms depends largely on our understanding of the properties of these allergens. Much effort has been devoted to determining the structure and organization of mite allergens, particularly of the house dust mites, toward understanding their activities and how they elicit immunological responses in humans. Here, we review the structural biology of the major allergens from two species of house dust mites, *Dermatophagoides farinae* and *D. pteronyssinus*, as well as allergens from a storage mite, *Blomia tropicalis*. The knowledge gained from the structural biology of these allergens will enable progress in producing novel, more effective treatments for mite allergies based on specific immunotherapy approaches.

Keywords Mite · Allergen · Structural biology

Introduction

Dust and storage mites produce allergens leading to widespread and common allergic diseases, including bronchial asthma, perennial rhinitis, and atopic dermatitis, that affect about 10–20 % of the global population [1, 2]. Two dust mite species, *Dermatophagoides farinae* and *D. pteronyssinus*, are the predominant sources of these allergens. Indeed, these species produce at least 23 groups of allergens [3]. The production of therapies against these allergens, to alleviate or eliminate allergic symptoms, hinges on our understanding of the individual allergens produced by dust mites as well as by storage mites like *Blomia tropicalis*. In fact, many studies have been devoted to understanding the *D. farinae* and *D. pteronyssinus* allergens from clinical, molecular, and immunological standpoints. In particular, analysis of allergen protein structure and organization has advanced the field so far as to allow the production of recombinant allergens. Importantly, allergen engineering offers the potential for large-scale production of more effective treatments, therefore such studies are indispensible in reducing allergic disease.

Studies of the structural biology of allergens, including *Dermatophagoides*, have established the foundations on which to explore the properties that determine allergenicity. Beginning with amino acid sequence homology, the higher-order structure and organization of the proteins can be derived to determine IgE binding capabilities or predict functions. Such information can also be used to understand potential cross-reactivity with other allergens. Thus, the contributions of structural biology approaches to studying mite allergens are crucial to advancing both clinical and immunological perspectives on these disease-causing proteins.

Structural biology: the means to an end

Structural biology studies seek to determine the physical properties of protein organization as a basis for imputing functions and interactions of a protein. These kinds of studies are commonly employed in allergen research because of the ease with which protein databases allow rapid classification and comparison of suspected allergens with known

Y. Cui (✉)
Department of Laboratory Medicine, Yancheng Health
Vocational & Technical College, Jiefangnan Road 263,
Yancheng 224006, Jiangsu Province, People's Republic of China
e-mail: ybcui1975@hotmail.com

proteins, particularly other allergens. Ascertaining how the protein folds allows researchers to deduce whether/to what extent the protein binds IgE antibodies; uncovering the size, stability, and solubility of an allergen aids in calculating its ability to cross mucosal barriers. Thus, structural information for allergens helps researchers gauge both the allergenicity—whether a protein induces a clinical or immunological allergic response—and cross-reactivity—ability to bind the same epitope as other proteins—of these widely-different proteins.

A variety of techniques are used to determine the structure of a protein. From sequence homology searches and computer modeling, to X-ray crystallography and nuclear magnetic resonance spectroscopy, the availability of both virtual and experimental approaches provides a range of possible methods to gain insight to protein organization and function. Indeed, the recent windfall of novel proteins—including allergens [4, 5]—identified through large-scale proteomics studies has placed structural biology techniques at the forefront of characterizing these new proteins.

The primary structure of a protein is determined by amino acid sequencing using standard molecular biology techniques, and often includes extrapolation from the nucleic acid sequence of a gene. Additionally, advanced mass spectrometry methods are commonly used to determine the amino acid composition of proteins. Often, though, bioinformatics analyses form the foundation for further characterization. Many freely-accessible databases (e.g., NCBI, PDB) are available for sequence alignments to identify similarities to proteins from other species or those with defined functions. Alignments also aid in predicting post-translational modifications that can alter a protein's conformation or activity.

The characterization of proteins beyond primary structure becomes gradually more complicated. Secondary structure, or the pattern of arrangement encompassing strands, helices, turns, and coils, while often predicted through computer modeling based on primary sequence, can be determined in solution using experimental approaches such as circular dichroism spectroscopy (CD). CD employs optical absorbance to determine the conformation of a protein in solution. Using this method, the arrangement of α-helices, β-sheets, turns, and coils can be determined. While low in resolution, the application of CD can quickly and reliably identify the secondary structure of a protein. For example, following primary structure determination through mass spectrometry, the secondary structure of the birch pollen allergen, Bet v 1, was deduced using CD [5].

Higher-order structures are determined by more advanced methods with higher resolution, particularly X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR). These techniques resolve protein structures at the atomic level, allowing the definition of tertiary and quaternary organization (in fact, primary and secondary structure can be determined at the same time). However, each of these approaches carries disadvantages. X-ray crystallography consumes time and materials and cannot offer insight to protein dynamics. NMR, while providing information on protein dynamics, is limited by the size of the protein being studied; larger macromolecules are precluded. Additionally, the comparability of crystal structures and solution-based structures (as in NMR) remains somewhat controversial. Despite these drawbacks, both approaches are widely used and continue to offer invaluable information toward the understanding of protein structure, function, and interactions. Further, these methods have enabled the classification of many different allergens and promoted a better understanding of allergic disease.

Finally, additional virtual methods are used to uncover biophysical information about proteins like allergens. Programs like Visual Molecular Dynamics (VMD; http://www.ks.uiuc.edu/Research/vmd), and the related NAMD, allow researchers to simulate the interactions of proteins, including their abilities to bind or cross membranes. The ability to model allergens in silico can lead to discoveries regarding cross-reactivity or allergenicity, as has been done for severity of response to peanut allergens [6].

Structural biology of mite allergens

The application of structural biology to the mite allergens has advanced both clinical and immunological understandings of how these proteins promote allergic disease. Allergens from *Dermatophagoides* are classified based on immune response into major groups 1 and 2 and middle groups 4, 5, and 7 [7]. However, there are at least 23 groups of allergens from the two most common dust mite species, *D. farinae* and *D. pteronyssinus*, all of which seem to elicit some degree of immune response [2, 3]. The importance of structural biology approaches to classifying and understanding these allergens has been invaluable—from outlining the basic properties of the proteins, to determining the ability to bind IgE and predicting cross-reactivity with related allergens, and, finally, aiding in the development of immunotherapy against the allergens [8]. Indeed, combining what is known about the major allergen groups with what can be derived from related proteins described in existing databases, much more information will be forthcoming regarding other allergen groups.

The spatial structure has been determined for several of the allergens of *Dermatophagoides* and *B. tropicalis*. Initial studies used in the late 1980s and early 90s used traditional molecular and immunological approaches to purify the allergen proteins (affinity chromatography), determine

their characteristics (SDS-PAGE), and quantify allergenicity (ELISA, radioimmunoassay) (e.g., [9–11]). Due to the advent of the many more advanced and specific techniques described above, a wide range of approaches have been employed to derive more detailed information about allergens. Below, we will review what is known about structures for the groups 1, 2, 5, 7, and 13 allergens, and how structural biology has contributed to the understanding of the allergenicity of these proteins.

Group 1 allergens

The group 1 allergens from *D. pteronyssinus* and *D. farinae*, Der p 1 and Der f 1, respectively, are major mite allergens, accounting for >50 % of IgE antibodies against total mite extract [12]. First isolated following cDNA cloning, these proteins are 82 % identical in sequence and belong to a family of papain-like cysteine proteases [13]. The crystal structures of both mature Der p 1 and Der f 1, as well as their recombinant prodomains, have been solved [14, 15]. Interestingly, despite extensive sequence homology, mature Der p 1 and Der f 1 do not exhibit identical structures. Der p 1 is monomeric and exhibits a cysteine protease fold and a magnesium binding site [15, 16]. Der f 1 is also monomeric, but lacks a metal binding site [15]. Structural differences are also found at the N-terminal and central loop regions, those sites that are exposed for surface binding. Additionally, Der f 1 appears to have fewer polymorphisms (five known) compared to Der p 1 (23 known) [7]. The prodomains of these two allergens differ from the mature forms with respect to secondary structure [14].

The structural differences between Der p 1 and Der f 1 likely result in differences in IgE binding. While these allergens exhibit high cross-reactivity in humans [17], antibody binding demonstrates a high degree of species specificity [9, 10, 17]: monoclonal antibodies raised against either species are ∼95 % species-specific. Four surface patches differ between Der f 1 and Der p 1, and areas these proteins have in common are the putative basis for cross-reactivity between the two allergens. The Der f 1 and Der p 1 structures show the presence of a conserved water binding site at one end and in the catalytic site at the other. The binding of water molecules could be biologically significant and may serve to catalyze the cysteine protease activity of the allergen [14].

The proforms of both Der p 1 and Der f 1 exhibit less allergenicty than do the mature and recombinant forms. The sites where the prosegments are anchored likely reflect the major IgE epitopes of the two allergens [14]. Based on the crystal structures of other cysteine proteases, molecular modeling of pro-Der f 1 was performed to demonstrate information on the location of IgE-binding sites blocked by the propeptides. To be more exact, the Der f 1 prosegment consists of an N-terminal domain-

structure followed by a C-terminal extended loop and is tightly anchored at two regions of the mature enzyme. The prosegment binding loop, the substrate binding cleft surrounding catalytic residue Cys35, and Asn53, the potential *N-glycosylation* site in mature Der f 1, is located on the surface distant from the catalytic site and the prosegment. These two regions on the molecular surface of the mature portion were blocked by the prosegments as major conformational IgE epitopes. The allergenicities of the proforms are less than those of the mature forms in a broad population of patients. Further, recombinant mature Der p 1 and Der f 1 are similar to the natural types in their structures and allergenicities and retain proteolytic activities [14].

Group 2 allergens

The Der p 2 and Der f 2 allergens belong to the ML domain lipid binding protein family, which also includes the Neimann-Pick type C2 (NPC2) proteins [7]. The crystal structures of these proteins have been determined, revealing anti-parallel beta strands and a hydrophobic cavity [18, 19]. Similarities with the structures of NPC2 proteins suggest a potential role in lipid binding or transport.

The overall tertiary fold of Der f 2 is that of two anti-parallel β-pleated sheets overlying each other. This fold is characteristic of the immunoglobulin superfamily. There are three disulfide bonds in Der f 2, forming covalent bonds between residues 8 and 119, 21 and 27, and 73 and 78 [20, 21], which are critical to the IgE-binding capacity of Der f 2. Indeed, Takai et al. [20] extended initial findings of the importance of these disulfide bonds to the allergenicity of Der f 2 [22–24] by studying engineered recombinant proteins with mutations in these bonds. Alterations in the disulfide bonds of Der f 2 led to conformational changes that reduced its IgE binding capacity. Therefore, loosening of the rigid tertiary structure by elimination of key intramolecular interactions in the allergen molecule would be an effective strategy in allergen engineering for safe and effective allergen-specific immunotherapy [20].

The similarities between Der f 2 and Der p 2 explain the extensive cross-reactivity between these allergens [11]. Indeed, the majority of the residues [86 % homology between Der f 2 and Der p 2 [21] on the binding surface of the group 2 allergens is conserved, allowing for similar IgE binding capacities [19]. NMR and structural modeling studies indicate two surface patches representing the IgE binding epitopes [25].

Group 5 allergens

The group 5 allergens from *D. pteronyssinus* and *D. farinae* show similarity to the group 5 allergen from another house dust mite species, *B. tropicalis*. Der p 5 and Blo t 5 are 42 % identical in amino acid sequence [26]. Additionally,

the crystal structure of Der p 5 was recently solved, revealing a monomeric protein that appears similar to the solution structure of Blo t 5 [27]. Further, Der p 5 has a helical structure that seems to interact with itself in an anti-parallel manner and a hydrophobic cavity. However, the function and activity of Der p 5 and other group 5 allergens remain unknown.

The group 5 allergens represent an important allergen group because of high responsiveness of patients to dust mite extracts containing these proteins. However, while some cross-reactivity exists among the group 5 allergens, in general IgE binding appears to be species-specific [26]. The Blo t 5 allergen appears to elicit IgE binding in 90 % of asthmatic and allergic patients [28]. However, further work is needed to understand the function and activity of this group of allergens.

Group 7 allergens

Group 7 allergens are fairly abundant in mite extracts. Indeed, these are the 44th most abundant protein produced by mites, ranking not far behind Der p 2 (41st) [29]. These allergens share sequence similarity to the prenylcysteine lyase proteins predicted to exist in insects [30]. Analyses also predict at least three isoforms exist for Der p 7; these may represent different levels of glycosylation [31]. A more recent characterization of Der p 7 indicates a structure comprising anti-parallel β sheets. Additionally, this allergen appears to bind a bacterial lipid product and resembles proteins in the Toll-like receptor pathway [27].

Some earlier work demonstrated that 50 % of dust mite-allergic patients react to Der p 7. This protein also appears to elicit a high T cell proliferation response [32, 33]. This specificity may be related to the similarity of Der p 7 to Toll-like receptors, which are involved in innate immunity [34]. However, further work is necessary to uncover the reason for this specificity.

Group 13 allergens

Group 13 allergens are considered fatty acid binding proteins (FABP) and are fairly abundant in mite extracts [7]. NMR studies of Der f 13 revealed anti-parallel beta strands preceeded by a helix-turn-helix motif [35]. Several surface residues are conserved between group 13 allergens and other proteins involved in the transport of fatty acids, with the highest homology to human brain FABP.

Epitope mapping and mutational analysis indicate that the surface residues are critical to IgE binding. However, despite high abundance in mite extracts and good capacity to bing IgE, the group 13 allergens do not elicit a severe immune response; only some patients produce IgE against these allergens [35].

Immunotherapy against mite allergens

Given the high allergenicity of the groups 1 and 2 allergens from house dust mites, effort has been devoted to developing specific immunotherapies (SIT) against these proteins. Additionally, advances in generating large amounts of recombinant proteins have led to animal-model studies of potential vaccines. For example, the changes induced in IgE binding capacity by altering the disulfide bonds of Der f 2 suggests potential avenues for immunotherapy against the group 2 allergens. Studies of human cells in vitro and of immunized mice indicate that the production of a hypoallergen—an allergenic protein less capable of binding IgE but retaining potent T cell reactivity—through genetic techniques may produce immunity to group 2 allergens [36]. A vaccination was recently developed by creating hypoallergens using recombinant Der p 1 and Der p 2; testing in rabbits demonstrated good levels of efficacy and tolerability [37]. Recent work also suggests that expressing the dominant T cell epitopes of Der p 1 in transgenic rice seed may produce a viable vaccination against group 1 allergens [38].

Several clinical trials have been conducted in recent years to test the efficacy and safety of SIT against the dust mite allergens, albeit with mixed results. Two types of SIT have been used: subcutaneous (SCIT) and sublingual (SLIT). Some have used purified mite extracts; others have used mite extracts combined with adjuvant therapies. In general, it appears that these SIT against mite allergens are effective, but the magnitude of effect varies between studies and types of SIT [39–41]. A search of government-sponsored clinical trials in the U.S. (clinicaltrials.gov; August 2012) returned 14 active trials investigating the efficacy of house dust mite SIT. A windfall of findings from these studies—and many more worldwide—can be expected within the next several years.

While SIT still have far to go before widespread use in humans, they offer some hope for eventual reduction of allergic disease. Continued advances in the field will rely heavily on the information generated by structural biology studies of each allergen. The work reviewed here has established strong foundations for progress in both understanding the structural biology of allergens and developing immunotherapies against them.

Conclusions

Proteins produced by mite species, especially the *Dermatophagoides*, are responsible for allergic disease worldwide. Knowledge gained from structural biology studies of these allergens has established a foundation for development of new therapies to treat and prevent allergic disease. Such studies will continue to contribute greatly to clinical

advances, particularly in this era of high-throughput genomics and proteomics.

References

1. Platts-Mills TA, Vaughan JW, Carter MC, Woodfolk JA (2000) The role of intervention in established allergy: avoidance of indoor allergens in the treatment of chronic allergic disease. J Allergy Clin Immunol 106:787–804
2. Platts-Mills TA (2002) The future of allergy and clinical immunology lies in evaluation, treatment, and research on allergic disease. J Allergy Clin Immunol 110:565–566
3. Bessot JC, Pauli G (2011) Mite allergens: an overview. Eur Ann Allergy Clin Immunol 43:141–156
4. Verdino P (2006) Structural characterization of pollen allergens. Clin Rev Allergy Immunol 30:73–95
5. Bollen MA, Garcia A, Cordewener JH et al (2007) Purification and characterization of natural Bet v 1 from birch pollen and related allergens from carrot and celery. Mol Nutr Food Res 51:1527–1536
6. Lehmann K, Schweimer K, Reese G et al (2006) Structure and stability of 2S albumin-type peanut allergens: implications for the severity of peanut allergic reactions. Biochem J 395:463–472
7. Thomas WR, Heinrich TK, Smith WA et al (2007) Pyroglyphid house dust mite allergens. Protein Pept Lett 14:943–953
8. Thomas WR, Hales BJ, Smith WA (2010) House dust mite allergens in asthma and allergy. Trends Mol Med 16(7):321–328
9. Heymann PW, Chapman MD, Platts-Mills TA (1986) Antigen Der f I from the dust mite *Dermatophagoides farinae*: structural comparison with Der p I from *Dermatophagoides pteronyssinus* and epitope specificity of murine IgG and human IgE antibodies. J Immunol 137:2841–2847
10. Chapman MD, Heymann PW, Wilkins SR et al (1987) Monoclonal immunoassays for major dust mite (*Dermatophagoides*) allergens, Der p I and Der f I, and quantitative analysis of the allergen content of mite and house dust extracts. J. Allergy Clin Immunol 80:184–194
11. Heymann PW, Chapman MD, Aalberse RC et al (1989) Antigenic and structural analysis of group II allergens (Der f II and Der p II) from house dust mites (*Dermatophagoides* spp). J. Allergy Clin Immunol 83:1055–1067
12. Thomas WR, Smith WA, Hales BJ et al (2002) Characterization and immunobiology of house dust mite allergens. Int Arch Allergy Immunol 129:1–18
13. Chua KY, Stewart GA, Thomas WR et al (1988) Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1. Homology with cysteine proteases. J Exp Med 167:175–182
14. Takai T, Kato T, Yasueda H et al (2005) Analysis of the structure and allergenicity of recombinant pro- and mature Der p 1 and Der f 1: major conformational IgE epitopes blocked by prodomains. J Allergy Clin Immunol 115:555–563
15. Chruszcz M, Chapman MD, Vailes LD et al (2009) Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. J Mol Biol 386:520–530
16. de Halleux S, Stura E, VanderElst L et al (2006) Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. J Allergy Clin Immunol 117:571–576
17. Lind P, Hansen OC, Horn N (1988) The binding of mouse hybridoma and human IgE antibodies to the major fecal allergen, Der p I, of *Dermatophagoides pteronyssinus*. Relative binding site location and species specificity studied by solid-phase inhibition assays with radiolabeled antigen. J Immunol 140:4256–4262
18. Derewenda U, Li J, Derewenda Z et al (2002) The crystal structure of a major dust mite allergen Der p 2, and its biological implications. J Mol Biol 318:189–197
19. Johannessen BR, Skov LK, Kastrup JS et al (2005) Structure of the house dust mite allergen Der f 2: implications for function and molecular basis of IgE cross-reactivity. FEBS Lett 579:1208–1212
20. Takai T, Ichikawa S, Yokota T et al (2000) Unlocking the allergenic structure of the major house dust mite allergen der f 2 by elimination of key intramolecular interactions. FEBS Lett 484:102–107
21. Suzuki M, Tanaka Y, Korematsu S et al (2006) Crystal structure and some properties of a major house dust mite allergen, Derf 2. Biochem Biophys Res Commun 339:679–686
22. Nishiyama C, Fukada M, Usui Y et al (1995) Analysis of the IgE-epitope of Der f 2, a major mite allergen, by in vitro mutagenesis. Mol Immunol 32:1021–1029
23. Takai T, Yokota T, Yasue M et al (1997) Engineering of the major house dust mite allergen Der f 2 for allergen-specific immunotherapy. Nat Biotechnol 15:754–758
24. Takai T, Yuuki T, Okumura Y et al (1997) Determination of the N- and C-terminal sequences required to bind human IgE of the major house dust mite allergen Der f 2 and epitope mapping for monoclonal antibodies. Mol Immunol 34:255–261
25. Ichikawa S, Takai T, Inoue T et al (2005) NMR study on the major mite allergen Der f 2: its refined tertiary structure, epitopes for monoclonal antibodies and characteristics shared by ML protein group members. J Biochem 137:255–263
26. Arruda LK, Vailes LD, Platts-Mills TA et al (1997) Sensitization to *Blomia tropicalis* in patients with asthma and identification of allergen Blo t 5. Am J Respir Crit Care Med 155:343–350
27. Mueller GA, Gosavi RA, Krahn JM et al (2010) Der p 5 crystal structure provides insight into the group 5 dust mite allergens. J Biol Chem 285:25394–25401
28. Naik MT, Chang CF, Kuo IC et al (2008) Roles of structure and structural dynamics in the antibody recognition of the allergen proteins: an NMR study on *Blomia tropicalis* major allergen. Structure 16:125–136
29. Batard T, Hrabina A, Bi XZ et al (2006) Production and proteomic characterization of pharmaceutical-grade *Dermatophagoides pteronyssinus* and *Dermatophagoides farinae* extracts for allergy vaccines. Int Arch Allergy Immunol 140:295–305
30. Thomas WR, Hales BJ, Smith WA (2005) Structural biology of allergens. Curr Allergy Asthma Rep 5:388–393
31. Shen HD, Chua KY, Lin KL et al (1993) Molecular cloning of a house dust mite allergen with common antibody binding specificities with multiple components in mite extracts. Clin Exp Allergy 23:934–940
32. Shen HD, Chua KY, Lin WL et al (1996) IgE and monoclonal antibody binding by the mite allergen Der p 7. Clin Exp Allergy 26:308–315
33. Hales BJ, Shen HD, Thomas WR (2000) Cross-reactivity of T-cell responses to *Dermatophagoides pteronyssinus* and *D. farinae*. Studies with group 1 and 7 allergens. Clin Exp Allergy 30:927–933
34. Mueller GA, Edwards LL, Alloor II et al (2010) The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. J Allergy Clin Immunol 125:909–917
35. Chan SL, Ong ST, Ong SY et al (2006) Nuclear magnetic resonance structure-based epitope mapping and modulation of dust mite group 13 allergen as a hypoallergen. J Immunol 176:4852–4860
36. Gafvelin G, Parmley S, Neimert-Andersson T et al (2006) Hypoallergens for allergen-specific immunotherapy by directed molecular evolution of mite group 2 allergens. J Biol Chem 282(6):3778–3787

37. Chen KW, Blatt K, Thomas WR et al (2012) Hypoallergenic Der p 1/Der p 2 combination vaccines for immunotherapy of house dust mite allergy. J Allergy Clin Immunol 130(2):435–443

38. Suzuki K, Kaminuma O, Yang L et al (2011) Prevention of allergic asthma by vaccination with transgenic rice seedexpressing mite allergen: induction of allergen-specific oral tolerance without bystander suppression. Plant Biotechnol J 9:982–990

39. Senti G, Johansen P, Haug S et al (2009) Use of A-type CpG oligodeoxynucleotides as an adjuvant in allergen-specific immunotherapy in humans: a phase I/IIa clinical trial. Clin Exp Allergy 39(4):562–570

40. Swamy RS, Reshamwala N, Hunter T et al (2012) Epigenetic modifications and improved regulatory T-cell function in subjects undergoing dual sublingual immunotherapy. J Allergy Clin Immunol 130(1):215

41. Nahm DH, Kim ME (2012) Treatment of severe atopic dermatitis with a combination of subcutaneous allergen immunotherapy and cyclosporin. Yonsei Med J 53(1):158–163