Title: Allergen ligands in the initiation of allergic sensitization

Running title: Allergen ligands

Author: Wayne R Thomas PhD

Affiliation: Centre for Child Health Research, University of Western Australia, Telethon Institute for Child Health Research.

Street address: Telethon Institute for Child Health Research, 100 Roberts Road, Subiaco 6008, Western Australia, Australia

Email: wayne@ichr.uwa.edu.au

Statement: The author has no conflicts of interest with the article and the article does not contain studies with humans or animals by the author.
Abstract

As investigations into the innate immune responses that lead to allergic sensitization become better defined there is a need to determine how allergens could interact with pattern recognition receptors that bind non-proteinaceous moieties. Many important allergens are not covalently bound to lipid or carbohydrate but have structures belonging to lipid, glycan and glycolipid-binding families. These include ML-domain proteins, lipopolysaccharide binding/cell permeability increasing proteins, von Ebner gland lipocalins, salivary lipocalins/major urinary proteins, plant pathogenesis related proteins PR-5 &10, uteroglobins, non-specific lipid transfer proteins, large lipid transfer proteins and proteins with chitin and other carbohydrate-binding modules. The binding expected is overviewed with regard to importance of the allergens and their ability to elicit responses proposed from experimental models. The evidence compiled showing that allergens from the same source sensitize for different types of adaptive immune responses supports the concept that individual allergens within these sources have their own distinctive interactions with innate immunity.

Keywords: Allergen, Ligand, Innate, Cytokine, IgE, Pathogen-associated molecular pattern, PAMP, Pattern recognition receptors, PRR, Lipopolysaccharide, LPS, lipid, glycolipid, chitin, carbohydrate, lipocalin, Dermatophagoides, House dust mite, Blomia, Birch, Grass, Cat, Dog, TLR4, C-type lectin.
Introduction to innate pathways

The importance of allergens signaling via the pattern receptors (PRR) of innate immunity and the inflammatory genes activated by the different receptor and signaling pathways have been expounded elsewhere [1, 2, 3] and analysed within the context of known properties of important allergens [4]. The engagement of these PRR by pathogen associated microbial pattern (PAMP)-like moieties on allergens would enhance endocytosis, drive adaptive responses and determine a cytokine bias. There are many types of soluble and cell-associated PRR but most investigations have focussed on the cell-associated Toll like receptors (TLR), especially TLR4, and cellular C-type lectin receptors (CTLR). Allergens could signal via either directly conjugated or associated PAMPs or the signaling could be instigated by co-presented PAMPs or by tissue damage acting via by IL-25, IL-33 and TSLP. Signaling that is not intimately associated with an allergen should have similar effects on responses to all proteins from the same allergen source while signaling by conjoint PAMPs could differentially affect responses to different allergens.

TLR4 signals via two pathways known by the acronyms MyoD88 and TRIF. MyoD88 signaling activates inflammatory responses via the NFκB and AP-1 transcription factors and TRIF signaling via NFκB and interferon regulatory factor 3 (IRF3) to activate type 1 interferons. The degree to which the signaling is directed down the TRIF and MyoD88 pathways can be instigated by the nature of the ligand as shown for lipopolysaccharide (LPS) derivatives and by interactions with signaling from other PRR. An important mechanism for TLR4 is that it signals from the membrane via MyoD88 and via TRIF when translocated to the endosomes with the activation of type 1 interferons that can block Th2 responses by suppressing GATA-3 and inducing IL-12. In mice TLR4 signaling by low doses of LPS induce Th2 responses while higher doses promote Th1. The doses that can be reckoned to be deposited in the lungs of asthmatics (with deposition of 50% of the particles from the 18m³ air inhaled per day) would deliver Th2 signals. TLR signaling also promotes DC trafficking to lymph nodes.

Members of the large CTLR family include the mannose receptor that has broad ligand specificity and the main function of transporting antigen for T-cell presentation. Other CTLR instigate or modify PRR signaling. DC-SIGN, mainly on tissue macrophages, modifies TLR signaling to promote IL-10 production and down regulate IL-12 establishing Th2-type milieux. Dectin-1 that binds β-glucans and other fungal and microbial carbohydrates transduces signals for dendritic cell (DC) maturation, endocytosis and Th17 with lesser Th1 cytokine production. It is found on DCs, macrophages and neutrophils particularly in the lung and intestine. Dectin-2 on macrophages, neutrophils and DC subsets binds mannose rich carbohydrates from fungi and bacteria inducing Th1 and Th17 cytokines but also cysteinyl leukotrienes. These can activate epithelial leukocytes called group 2 innate lymphoid cells to promote Th2 responses. DC-SIGN and Dectin-2 would favour Th2 responses and Dectin-1 could attenuate Th1 responses.

Different responses to different allergens and antigens
IgE antibody responses to many allergen sources are preferentially directed to one or a few allergens (Table 1). There could be a large contribution from the abundance since, with the possible exception of Der p 23 [15], important allergens are present in high amounts [4] but the reverse is not true. For the house dust mite (HDM) the highly-abundant fatty acid binding protein is [4,16] is only occasionally allergenic [4, 17] and likewise no or low IgE antibody titers are induced by the abundant HDM arginine kinase [17,18] and ferritin [19,20].

Early experiments comparing peripheral blood mononuclear cells (PBMC) responses to the HDM allergens Der p 1 and Der p 7 [21], and HDM ferritin and Der p 2 [19], showed independent patterns of cytokine induction. Similar results have now been obtained with a global-epitope approach where T-cells were stimulated with peptides representing the T-cell epitopes of all the allergens from the one source. For the grass pollen *Phleum pratense*, the allergens Phl p 1 &11 only induced IL-5 release while Phl p 4&5 induced IL-5, IL-10, IL-17 and IFN-γ and Phl p 2, 3&13 only induced IL-5 and IFN-γ release [22•]. Supporting results were found for cockroach [23]. The same approach also found that non-allergenic pollen proteins, as found for HDM ferritin, elicited Th2 cytokines that were more prominent for responses of non-allergic subjects [24]. This indicates that the overall pattern of responses to the allergens are important as shown by the differences in the IL-10 production to Der p 2 and ferritin [19].

**Covalent lipid and carbohydrate modification of allergens**

The important allergens from common sources are not known to be covalently bound to lipid and only some are glycosylated (Table 1). It is noteworthy that the Der p 2 sequence has no N- or O-glycosylation amino acid sequence motifs and by mass spectrometry natural Der p 2 is not glycosylated ([25]). Its presentation by mannose receptors [26] and binding to lectins [27] needs explanation. Other important allergens that are not non-glycosylated are Bet v 1 [28], group 5 grass allergens [29], Amb a 1 [30] and Ara h 2 [31], with new mass spectrometry data for Ara h 2 revising observation made with a partially purified protein. From their primary and tertiary structures however many allergens would be expected, to be receptors for ligands with PAMP-like properties (Table 2). The following overview compiles the PAMP-binding that might be expected from different allergens in relationship to their ability to sensitize allergic subjects.

**Allergens with ligand-binding properties outlined in Table 2**

**ML-domain lipid binding proteins**

The group 2 HDM allergens are homologues of the ML-domain protein MD-2 that binds LPS to load TLR4 for innate immune signaling. Investigations with LPS-induced IL-8 induction by HEK293 cells showed that cells transfected with Der p 2 did not require MD-2 to respond and that Der p 2 made an MD-2-like association with TLR4 [32]. The IL-8 induction was however low but increased 30 fold when Der p 2 was transfected into MD-2 replete cells, and since this was twice that released by cells with only the MD-2 an additive effect was possible. The importance of the MD-2 mimicry in allergic sensitization was partially supported in a mouse model. Der p 2 administered with LPS could sensitize MD-2 mice but only with 10x higher doses
than those used in the controls. Since Der p 2 can stimulate non-TLR4 mediated pro-inflammatory cytokine release from epithelial cells [33] including GM-CSF further investigation is required. GM-CSF independently promotes sensitization to inhaled proteins [34]. There are thus issues that need to resolved to determine if the mimicry occurs in physiological settings and then if it does occur humans. The sensitization in the mouse model also needs to be distinguished from transient sensitization that precedes the development of mucosal tolerance [35].

LPS binding experiments initially only detected weak binding by recombinant Der p 2 [36] although without addressing the discrepancy others, using the same methodology, found binding with Der f 2 with nanogram affinity and 1:1 stoichiometry [37]. Accompanying NMR perturbation studies showed that the residues that bound LPS were to those equivalent to the LPS-binding residues of MD-2 and residues expected to shift to create a second TLR4 binding region were also perturbed. An indication of the importance of the LPS binding is that Blo t 2 from Blomia tropicalis, which is not an important allergen [17] lacks two key TLR-4-binding residue that are conserved in other group 2 allergens [4].

Endogenous lipid ligands now found for MD-2 [38] include glycolipids so perhaps similar ligands for HDM allergens provide the ligands for the CTLLR. A further interesting ligand for MD-2 is morphine that induces Th2 microglial mediated neuroinflammation [39] via MD-2 and TLR4. Since morphine is known to promote Th2 type immune responses [40] this might occur in other antigen-presenting cells.

Lipopolysaccharide binding protein/bacterial permeability-increasing family binding

Equ c 4, Der p 7 and Fel d 8 are LPS binding protein/bacterial permeability-increasing proteins (LBP/BPI). BPI neutralizes LPS while LBP transfers LPS to TLR4 [41]. LBP/BPI have one or two matching ligand-binding hydrophobic cavities (42,43) and have diverse roles beside LPS binding [44]. Der p 7 [45] and Der f 7 [46] have hydrophobic binding cavities while the latherin Equ c 4 does not [47] indicating a primary surfactant function. Der p 7 has been shown to bind the lipid-tailed cyclic peptide polymyxin B but not LPS. Der p 7 sensitizes half of HDM-allergic subjects with IgE titers approaching those to major allergens although, for individuals, in proportion to their Der p 1&2 titers. Fel d 8 is the highest IgE inducing allergen for a significant number of cat-allergic subjects while Equ c 4 binds IgE in 72% of sera from horse allergic subjects, often strongly.

von Ebner gland/tear-like lipocalins

The salivary von Ebner gland lipocalin allergens resemble tear lipocalin and β-lactoglobulin, an occasional milk allergen. Can f 1 is amongst the highest IgE binding specificities produced by dogs [48] and Fel d 7 induces similar IgE antibody titers in cat allergy [49,50]. It is one of the main allergens for the 35% of cat-allergic subjects who do not have Fel d 1 as their highest IgE binding specificity [50] consistent with it independently driving Th2 responses. An innate ability to stimulate Th2 responses is indicated by the fact that only a few of the many abundant proteins in saliva [51] have been identified as allergens.
Lipocalins have a characteristic fold enclosing a hydrophobic ligand-binding site. By analogy to tear lipocalin the von Ebner gland-like lipocalins could bind many large lipids within a wide hydrophobic cavity with a flexible entrance [52,53]. Ligands such as retinoic acid and 18-carbon chain fatty acids bind within the cavity [53,54] while longer chain and complex lipids, especially with polar groups, must be extruded from the cavity. They however bind with the same high affinity. They include phosphocholines, sphingomyelin, lysophosphatidylcholines, glycolipids and oxidised phospholipids (53, 55). A scavenger function for tear lipocalin has been proposed from its ability to bind poorly soluble lipids and oxidised phospholipids [55]. Many of lipids have immunomodulating activities. Oxidised phospholipids inhibit DC maturation [56] and bind the scavenger receptors C-reactive protein and CD36 [57] for increased phagocytosis and cytokine production. CD36 has a role as a TLR2/6 signaling [58] with CD36 knockout mice showing exaggerated inflammatory responses [59]. Phosphocholines especially those produced by helminths and bacteria modulate immune responses by multiple mechanisms to reduce Th1 responses by interfering with TLR4 signaling [60]. Retinoids block the maturation of DCs and promote mucosal migration [61]. Binding to the abundant glycolipids present in saliva [62] could also deliver the lipocalins to CTLRs.

Lipocalins also bind directly to cellular receptors. The type 1 lipocalin receptor binds a range of lipocalins including β-lactoglobulin [63] mediating its endocytosis. Northern blots show its expression in intestine, trachea and peripheral blood leukocytes with abundant expression in lymph node [64]. The lipocalin 2 receptor, or megalin, binds both tear and odorant binding lipocalins [53]. It expressed in T-cells, myeloid cells and epithelial cells, including alveolar type II cells [65] that have been shown to orchestrate DC function in murine viral pneumonia [66].

Salivary lipocalins/major urinary proteins

The salivary lipocalins Equ c 1 [48], Fel d 4 [49,50], and Can f 6 [67] and the structurally related major urinary proteins (MUPS) such as the mouse Mus m 1 [68] induce similar IgE antibody responses to those induced by the von Ebner gland-type lipocalins. They however only bind small volatile compounds like odorants, pheromones and sex steroids [69] entirely within their small hydrophobic cavities [70]. It follows that the report of Can f 6 enhancing TLR4 dependent LPS signaling [71] is not due to canonical ligand binding. Sex steroids however can affect immune responses [72] so perhaps the co-delivery of small mediators might be important. Also members of this type of lipocalin have been shown to directly mediate communication between animals without their ligands. Fel d 4 is a kairomone that delivers avoidance signals to mice by exciting vomeronasal sensory neurons [73] and MUPS deliver pheromone-independent signals to G-protein receptors [74]. Having been evolved to be transferred from animal to animal by inhalation might be a useful for an allergen. Salivary lipocalins also bind the lipocalin 2 receptor [53].

Insect lipocalin

Cockroach Bla g 4 is a male pheromone-carrying lipocalin shown to bind juvenile hormone [75]. Analogous to the salivary lipocalins the ligand-binding cavity of Bla g 4 is buried within the center of the protein so the docked ligand is not exposed to the solvent and the entrance is closed. Ligand-mediated effects on allergenicity are
accordingly not that obvious and Bla g 4 induces less prevalent and smaller IgE responses than the non-proteolytic aspartic-protease-homologue Bla g 2 and the glutathione-S-transferase Bla g 5 [76].

Bet v 1 PR-10

Bet v 1 from the PR-10 pathogenesis family is the main allergen for birch pollen. PR-10 proteins are also the main allergens for other Fagales trees such as Cor a 1 and Aln g 1 from hazel and alder and food allergens of plants, for example apple Mal d 1. The PR-10 protein fold creates a large hydrophobic pocket that could carry or store many ligands including flavonoids, cytokinins and fatty acids. The natural ligand for the main isoform for Bet v 1a (Bet v 0101) is now known to be the flavonoid glycoside Q3OS (quercetin-3-O-sophoroside) [77] that requires both the glycan and the lipid for binding.

Despite being readily found in pollen there are several isoforms of Bet v 1 that bind very IgE antibody [78]. They have 95% amino acid sequence identity with Bet v 1a and almost identical tertiary structures. One key difference is that they, as determined with model ligands, have different binding profiles [79]. This therefore could be important for natural ligands. Experiments with the non-allergenic isoforms without ligands have however shown that when endocytosed by human DCs they have a poor ability to activate T helper cells suggesting the importance of other properties [78].

Uteroglobin

Fel d 1, a uteroglobin, has small internal cavities that could only bind small molecules [80] so its reported ability to help LPS induce inflammatory cytokines [71] is not due to an expected ligand interaction. Given the very high concentration of Fel d 1 required to mediate the reported augmentation of cytokine release by non-genetically-engineered cells further fractionation of the protein preparation used might be informative. Low doses of LPS have been demonstrated to enhance Fel d 1-induced Th2 cytokine release by PBMC of cat-allergic subjects and Th1 cytokine by PBMC of non-allergic subjects [81] but almost identical effects have been reported for grass pollen [82]. Fel d 1 could have other immuno-enhancing properties such as the suggested calcium-ion-mediated activation of phospholipase A2 [80] or by delivering lipid mediators. Uteroglobins have many anti-inflammatory effects thought to be either mediated by pharmacologically active ligands or by blocking the receptors of mediators and uteroglobin-knockout mice show heightened Th2 responses [83].

Non-specific lipid transfer proteins

Non-specific lipid transport proteins are important food allergens for peach (Pru p 3), vegetables and nuts (Cas s 8) and aeroallergens, reported for Par j 8 from Parietaria judaica pollen, olive Ole e 7 and latex Hev b 12 [84]. They are associated with severe Rosaceae spp. fruit allergy in the Mediterranean area but their importance otherwise is uncertain. They have a tunnel-like cavity that can accommodate phospholipids glycolipids and fatty acids but while their transfer of phospholipids and glycolipids between plant organelles has been demonstrated [85] there is paucity of information on ligands and other, anti-microbial, functions have
been proposed [84]. The fruit allergens are unglycosylated but have the important properties for food allergens of resistance to heat and proteolysis.

Large lipid transfer proteins

This family includes vitellogenin, insect apolipophorins and the mammalian apolipoprotein B-100, each binding many types of phospholipids and for invertebrates even to binding and neutralizing PAMPs encountered during bacterial and fungal infections [86]. The HDM Der 14 was defined following studies in Japan which discovered a protein designated M-177 that bound IgE in nearly all HDM-allergic subjects although investigations elsewhere with Der p 14 and Blo t 14 and Der p 14 [17, 88] found very modest IgE binding using fragments reported for Der f 14 to bind IgE [87]. Recently however vitellogenins have been identified as Api m 12 and Ves v 6 allergens of bee and wasp venom binding IgE in 40% of sera [89] and vitellogenins are the main allergens of fish roe [90] from which they can induce severe reactions in subjects not allergic to fish flesh.

Der p 5-like dimers

The group 5 and 21 HDM allergens are related proteins that form bundles of coiled coils. For Dermatophagoides spp. they are significant but not top-tier IgE binders [17,18, 88] but they are the main allergens of B. tropicalis [17]. Der p 5 forms structures that dimerize to create a large hydrophobic cavity [91]. This has however not been found for Blo 5 [92] and Blo t 21[93] and modelling of the known group 5 and 21 sequences suggest that this might only occur for Der p 5 and Der f 5 [94]. It thus appears that it is the less, not the more, allergenic molecules that have the ligand binding structure. Blo t 5 does have an N-glycosylation motif but there are none for other group 5/21 allergens.

Non-chitin carbohydrate binding modules

The major group 1& 2/3 grass pollen allergens have a carbohydrate-binding module constituted by a β fold with a concave surface containing a groove with surface exposed aromatic residues for carbohydrate attachment [95]. They are β-expansin-family proteins that facilitate pollen tube penetration by non-enzymatically destabilizing extracellular matrix polysaccharides.

Ole e 9&10 from olive pollen share another type of carbohydrate binding module without a β-fold but with two grooves containing clusters of surface exposed aromatic residues. Ole e 9 binds and hydrolyses 1,3-beta-glucans during pollen germination [96] while Ole e 10 binds the glucose polymer laminarin an activity also associated with pollen tube formation [97]. They are not the most prevalent IgE binding proteins of olive pollen but might have important in severe allergy [98].

A carbohydrate binding domain was been implicated in the binding oxidised cellulose particles by Der p 2 [99] but while this ML-domain protein has a β-fold it does not have surface grooves and exposed aromatic residues. Oxidised cellulose binds proteins without carbohydrate binding domains via polyuronic acid carboxylic groups and, as found for Der p 2 [99], the highest binding is found with preparations with the highest carboxylic content [100]. Indeed oxidized cellulose
bound all detectable the proteins in HDM extracts [99]. It is likely, that as discussed for ML-domain proteins, the carbohydrate bound by Der p 2 will be glycolipid.

Ole e 1-like

Ole e 1 from olive and Che a 1 and Pla l 1 from the chenopod and English plantain are very important allergens [98]. The biochemical properties are unknown but by BLAST they have distant amino sequence to the expansin although without the critical prolines. Ole is variably but highly glycosylated [98].

Thaumatin proteins

The thaumatin food allergens that include Pru p 2 from peach, Act d 2 from kiwi fruit and Mal d 2 from apple are PR-5 pathogen-response proteins that bind β-glucans for fungal resistance [101]. Pru p 2 binds IgE in over 50% of allergic patients indicating it can be important [102]. Cup a 3 and Jun a 3 from cypress and cedar pollens are also thaumatins along with pollen proteins from mugwort, birch and plane trees. They do not show prevalent IgE binding but microarray analysis of patients in different geographical region has indicated an association of the exposure to pollen allergen and severity of the food allergy to its corresponding fruit [102].

Chitin binding allergens

Chitin-binding proteins are a subset of the carbohydrate binding module proteins. They include chitinases, non-catalytic chitinase-like proteins and proteins with chitin-binding domains. They have evolved by the fusion of different combinations of genes encoding chitin binding, catalytic and glycosylation domains for defence against chitin-coated fungi and for tissue remodelling of invertebrates and for adding function to their chitinous structures. Adjuvant effects of chitin depend on polymer size [103]. Large chitin is inert, intermediate-sized stimulates TNF-α via TLR2 and NFkB while small chitin (2–10 µm) stimulates cytokine especially IL-10 via dectin-1 with lesser contributions from TLR-2. Chitin can be an adjuvant for immune responses induced by ovalbumin injections, sensitizing mice for Th2-type pulmonary eosinophilia by TLR2 and MyD88-dependent mechanisms and, in keeping with a possible role for Dectin-1, IL-17 with little IFN-γ [104].

Arthropods typically produce family 18 glycosyl hydrolase chitinases and chitinase-like lectins and multiple peritrophin-like proteins with type 2 chitin binding domains (ChrBD2), also called carbohydrate binding module 4 with a characteristic 6 cysteine organisation [105]. ChrBD1 are eight-cysteine domains characteristic of family 19 chitinases and chitin-binding proteins of plants where they are often referred to as a hevein domain after the rubber latex hevein protein. Class 1 of seven classes of plant family 19 chitinases are known for their association with the latex-fruit syndrome. Hev b 6.02 from rubber latex is a 43 amino acid peptide corresponding to the ChBD1 domain naturally cleaved from the larger prohevein, Hev b 6.01 and as found for ChBD1-containing wheat germ agglutinins might be produced for resistance to insect larvae [106]. Invertebrates use chitin for their skeletons for protecting mucosal surfaces. The peritrophins attach to chitin matrices that line the gut of insects and surround their food being released into
environment in dung. They can be readily eluted from these matrices [107]. Chitinases have been shown to remain bound to insoluble chitin [105]. Natural associations of plant heveins with chitin might not be so common but they bind to insoluble chitin and fungi are abundant commensals of the gut, skin and nose.

Hev b 5 and Hev v 6 are the major latex allergens [108] and for Hev b 6 the hevein ChBD1 domain, binds most of the IgE [109]. Homologous chitinases of avocado (Pra a 1), bananas (Mus a 1), chestnuts (Cas s 5) and other foods are important allergens with almost complete cross reactivity. The hevein domain is not only the important IgE-binding determinant but also resistant to digestion. Hev b 11 is a leaf chitinase similar to Hev b v 6 is less cross reactive with fruits and not as allergenic [108].

The group 15 HDM allergens are family 18 chitinases with extensive O-glycosylation constituting half their mass. Der f 15 has been detected in the gut but is not in the dung ball. It is a major IgE-binding protein for HDM-allergic dogs [110] but as related elsewhere [111•] most of the IgE binding was found to be anti-carbohydrate. Anti-polypeptide antibodies to Der p 15 are however found in 30-50% of HDM-allergic humans but usually at low titer [111•,112]. Relevant to the possible adjuvanticity of chitin, the anti-Der p 15 titers of individuals correlate very strongly with their titers to the non-cross reactive chitinase-like lectin Der p 18 but not with titers to Der p 1, 2, 5&7 all of correlate [111•]. The group 18 allergens are chitinase-like lectins that lack critical chitinase catalytic residues and O-glycosylation sites [112, 113]. The sequence identity of Der p 18 and Der p 15 is only 29% and they do not cross-react serologically. Der p 23 is a 69 amino acid peritrophin mostly comprised by a ChrBD2 [15]. It is produced by gut epithelial cells and can, only just, be detected in the faecal pellet. It induces prevalent and high IgE titers in HDM-allergic subjects that have an overall correlation with those to Der p 1&2 but some subjects show high responses to Der p 23 with no or little IgE binding to Der p 1&2. The results point to Der p 23 being a hitherto unrecognized major mite allergen with an ability to drive allergic responses.

The Blo t 12 allergens contain a ChBD2. Blo t 12.0101 from Colombia binds IgE in 35% of B. tropicalis-allergic subjects and Blo t 12.0102 binds in 23%[114]. The higher responses to Blo t 12.0101 is consistent with Blo t 12.0102 lacking a cysteine in its ChBD and with experiments showing that chitin enhanced IgE responses to Blo t 12.0101 injected into mice with alum but not those to Blo t 12.0101[114]. Both however are considerably less allergenic than Blo t 5&21 [17].

Conclusions

It is important to know how PRRs signal to drive the adaptive responses to allergens. The signals and the PRR engaged can be different for different allergens from the same source and for allergens from different sources. While the intricate immunological investigations into PRR signaling have rightly attracted considerable interest [3, 32] they have not been well-married to the knowledge of the PAMPS attached or associated with the allergens and to the likelihood that the allergens investigated are important drivers of immune responses to themselves and co-presented antigens. Critically the cellular experiments have pointed to the need to define the glycosylation and post-translational modifications of allergens and their associations with potential PAMP ligands. The nature of the glycan carried by Der p
2, as an example, is unknown and, like investigations begun for Bet v 1, it should be possible to determine if there is a diversity of ligands and if the binding by variants known to have different allergenicity[115] is different [. Most allergens do not covalently bind lipid and many of the main allergens are not glycosylated, so their potential to bind PAMP-like ligands overviewed here can be important for their ability to activate PRRs. With experiments begun on non-allergenic proteins from sources of allergens and with the quantitation of IgE responses, thanks to microarrays, becoming more commonplace there is considerable potential to explore these structure function relationships.

References


38. Kondo Y, Ikeda K, Tokuda N, et al. TLR4-MD-2 complex is negatively regulated by an endogenous ligand, globotetraosylceramide. Proc Natl Acad Sci U S A. 2013;110(12):4714-9. This article describes endogenous non-LPS ligands for MD-2 including glycans that should be considered there relevance to natural Der p 2 ligands.


50. Hales, BJ, Chai LY, Hazell L, et al. IgE and IgG binding patterns and T-cell recognition of Fel d 1 and non-Fel d 1 cat allergens. J Allergy Clin Immunol: In Practice. 2013;1(6):656-65. This report compares the IgE responses to different cat allergens made by cat-allergic subjects and shows that for many Fe d 1 is not the main allergen so the lipocalin allergens and the putative lipopolysaccharide-binding allergen must (in the absence of another mechanisms) must initiate their own adaptive responses.


111. Hales BJ, Elliot CE, Chai LY, et al. Quantitation of IgE binding to the chitinase and chitinase-like house dust mite allergens Der p 15 and Der p 18 compared to the major and mid-range allergens. Int Arch Allergy Immunol. 2013;160(3):233-40. The paper shows that IgE antibodies to the chitin binding allergens of HDM correlate with each other but not to Der p 1,2,57 7 that all correlate with each other. IgE responses to groups of allergens from the same source accordingly must be being independently regulated.


<table>
<thead>
<tr>
<th>Allergen source</th>
<th>Allergen</th>
<th>Biochemical</th>
<th>Glycosylation</th>
<th>Glycosylation motif</th>
</tr>
</thead>
<tbody>
<tr>
<td>House dust mite</td>
<td>Der p 1</td>
<td>Cysteine protease</td>
<td>Present</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td></td>
<td>Der p 2</td>
<td>Lipid transport</td>
<td>Absent</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Der p 23</td>
<td>Peritrophin</td>
<td>Unknown</td>
<td>O-glycosylation</td>
</tr>
<tr>
<td>Blomia tropicalis</td>
<td>Blo t 5</td>
<td>Unknown</td>
<td>Unknown</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td></td>
<td>Blo t 21</td>
<td>Unknown</td>
<td>Unknown</td>
<td>None</td>
</tr>
<tr>
<td>Cat</td>
<td>Fel d 1</td>
<td>Secretoglobin</td>
<td>Present</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td>Dog</td>
<td>Can f 1</td>
<td>von Ebner lipocalin</td>
<td>Unknown</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td></td>
<td>Can f 4</td>
<td>Odorant lipocalin</td>
<td>Unknown</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Can f 5</td>
<td>Prostate kallikrein</td>
<td>Unknown</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Can f 6</td>
<td>Salivary lipocalin</td>
<td>Unknown</td>
<td>None</td>
</tr>
<tr>
<td>Cockroach</td>
<td>Bla g 2</td>
<td>Aspartic protease like</td>
<td>Present</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td></td>
<td>Bla g 5</td>
<td>Glutathione-S-transferase</td>
<td>Unknown</td>
<td>None</td>
</tr>
<tr>
<td>Grass</td>
<td>Phl p 1</td>
<td>Expansin</td>
<td>Present</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Phl p 5</td>
<td>Unknown</td>
<td>Absent</td>
<td>None</td>
</tr>
<tr>
<td>Birch</td>
<td>Bet v 1</td>
<td>Lipid transport</td>
<td>Absent</td>
<td>None</td>
</tr>
<tr>
<td>Olive</td>
<td>Ole e 1</td>
<td>Pollen tube formation</td>
<td>Present</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td>Peanut</td>
<td>Ara h 1</td>
<td>Storage protein</td>
<td>Present</td>
<td>N-glycosylation</td>
</tr>
<tr>
<td></td>
<td>Ara h 2</td>
<td>Storage protein</td>
<td>Absent</td>
<td>N-glycosylation*</td>
</tr>
<tr>
<td>Milk</td>
<td>Bos d 9-12</td>
<td>Caseins</td>
<td>Present</td>
<td>O-glycosylation</td>
</tr>
</tbody>
</table>

*Table 1.* Glycosylation of main allergens from common sources allergens showing whether the presence or the absence of glycosylation has been described (Glycosylation column) or is unknown and whether the amino acid sequences contain N-glycosylation or possible O-glycosylation motifs.

* Ara h 2 sequence contains an N-glycosylation motif but is not glycosylated (see text).
<table>
<thead>
<tr>
<th>Ligand binding protein type</th>
<th>Representative allergens</th>
<th>Proposed Ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML-domain</td>
<td>Der p 2, Der f 2, Blo t 2</td>
<td>Lipopolysaccharide, glycolipid, opioids</td>
</tr>
<tr>
<td>LPS binding/Cell permeability increasing protein</td>
<td>Der p 7, Fel d 8, Equ c 4</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>von Ebner gland/tear-like lipocalins</td>
<td>Can f 1, Fel d 7, Bos d 2</td>
<td>Phospholipids, glycolipids, lipids</td>
</tr>
<tr>
<td>Salivary lipocalins/major urinary proteins</td>
<td>Fel d 4, Can f 6, Equ c 1, Mus m 1</td>
<td>Small volatile lipids/lipid hormones/pheromones</td>
</tr>
<tr>
<td>Invertebrate lipocalin</td>
<td>Bla g 4</td>
<td>Pheromone</td>
</tr>
<tr>
<td>Bet v 1 like (PR-10)</td>
<td>Bet v 1, Cor a 1, Aln g 1, Mal d 1</td>
<td>Bet v 1a binds glycolipid</td>
</tr>
<tr>
<td>Uteroglobin</td>
<td>Fel d 1</td>
<td>lipid hormones/mediators</td>
</tr>
<tr>
<td>Non-specific lipid transfer proteins</td>
<td>Pru p 3, Act d 10, Par j 8, Cas s 8</td>
<td>phospholipids, glycolipids, fatty acids,</td>
</tr>
<tr>
<td>Large lipid transfer proteins</td>
<td>Der p 14, Blo t 14, Ves v 6, Api m 12, Fish roe</td>
<td>Large lipids, phospholipids, retinoids</td>
</tr>
<tr>
<td>Der p 5 like dimers</td>
<td>Der p 5 with Der p 21, Blo t 5 &amp; 21 possible</td>
<td>Hydrophobic unknown</td>
</tr>
<tr>
<td>Non chitin carbohydrate binding module containing</td>
<td>Phl p 1 &amp; 2/3, Zea m 1, Ole 9 &amp; 10</td>
<td>Extracellular matrix plant carbohydrate, 1,3-beta-glucans</td>
</tr>
<tr>
<td>Ole e 1 like</td>
<td>Ole e 1, Che a 1, Pla l 1</td>
<td>Extracellular plant matrix proposed</td>
</tr>
<tr>
<td>Thaumatin (PR-5)</td>
<td>Pru p 2, Act d 2, Mal d 2, Cup a 3, Jun a 3</td>
<td>β-glucans</td>
</tr>
<tr>
<td>Chitin binding</td>
<td>Der p 15, 18, m 23; Blo t 12, Hev b 6 &amp; 11, Prs a 1, Mus a 2, Cas s 5,</td>
<td>Chitin</td>
</tr>
</tbody>
</table>

**Table 2**

Grouping of allergens into proposed types of ligand binding proteins. A synopsis for each ligand binding protein type is given in the order of the table in the main text following the heading "Allergens with ligand-binding properties outlined in Table 2". The sources of the allergens in the table are also given therein.