Published in final edited form as:

J Allergy Clin Immunol. 2013 December ; 132(6): . doi:10.1016/j.jaci.2013.06.014.

Novel structure of cockroach allergen Bla g 1 has implications for allergenicity and exposure assessment

Geoffrey A. Mueller, PhD^{a,*}, Lars C. Pedersen, PhD^{a,*}, Fred B. Lih, B.A.^a, Jill Glesner, BS^b, Andrea F. Moon, M.S.^a, Martin D. Chapman, PhD^b, Kenneth B. Tomer, PhD^a, Robert E. London, PhD^a, and Anna Pomés, PhD^b

^aLaboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, NC

^bINDOOR Biotechnologies, Inc. Charlottesville, VA

Abstract

Background—Sensitization to cockroach allergens is a major risk factor for asthma. The cockroach allergen Bla g 1 has multiple repeats of ~100 amino acids, but the fold of the protein and the biological function are unknown.

Objective—To determine the structure of Bla g 1, investigate the implications for allergic disease, and standardize cockroach exposure assays.

Methods—Natural Bla g 1 and recombinant constructs were compared by ELISA using specific murine IgG and human IgE. The structure of Bla g 1 was determined by X-ray crystallography. Mass spectrometry and NMR were utilized to examine ligand-binding properties of the allergen.

Results—The structure of a recombinant Bla g 1 construct with comparable IgE and IgG reactivity to the natural allergen was solved by X-ray crystallography. The Bla g 1 repeat forms a novel fold with 6 helices. Two repeats encapsulate a large and nearly spherical hydrophobic cavity, defining the basic structural unit. Lipids in the cavity varied depending on the allergen origin. Palmitic, oleic and stearic acids were associated with nBla g 1 from cockroach frass. One Unit of Bla g 1 was equivalent to 104 ng of allergen.

Conclusions—Bla g 1 has a novel fold with a capacity to bind various lipids, which suggests a digestive function associated with non-specific transport of lipid molecules in cockroaches. Defining the basic structural unit of Bla g 1 facilitates the standardization of assays in absolute units for the assessment of environmental allergen exposure.

Keywords

Allergen; asthma; Bla g 1; cockroach; structure; ligand-binding proteins; exposure assessment

Corresponding author: Geoffrey A. Mueller, National Institute of Environmental Health Sciences, 111 T.W. Alexander Drive, MD-MR-01, Research Triangle Park, NC 27709, T: 919-541-3872, F: 919-541-5707 Mueller3@niehs.nih.gov.

*GAM and LCP contributed equally to this article

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

The first report on cockroach allergy dates back to 1964, when Bernton and Brown described an association between sensitization to cockroach allergens and asthma morbidity in New York.¹ Subsequent studies have confirmed this original finding, especially among lower socio-economic populations, but also in suburban middle-class homes of asthmatic children.^{2–4} Rates of cockroach sensitization are highest in the Northeast U.S.A., reaching up to 81% in the Bronx, New York, with the highest allergen levels in high-rise apartments.³ A recent study showed that prenatal exposure to cockroach allergens was associated with a greater risk of developing sensitization.⁵ Allergic sensitization to cockroach is a risk factor for the development of asthma, as is sensitization to other indoor allergens from dust mite, mold, dog or cat.^{6–8}

A major breakthrough in understanding the role of cockroach allergy in asthma was the identification and measurement of cockroach allergens. Bla g 1 and Bla g 2 were the first two major cockroach allergens to be identified.⁹ Both are secreted in the cockroach digestive tract, and sensitization occurs by inhalation of fecal particles that carry the allergens and are released to the environment.¹⁰ Both cockroach allergens have been used as markers of allergen exposure. Exposure to low doses of Bla g 1, 2 U/g of dust or less, was determined to be a risk factor for sensitization, however the risk factor plateaus above 4 U/g of dust.^{2, 11} A seminal study by Rosenstreich et al. established 8U/g of dust as threshold levels of “morbidity due to asthma”, as part of a National Cooperative Inner-City Asthma Study.^{2, 11} However, the levels of Bla g 1 exposure in all these studies were expressed in arbitrary units, and should be converted to absolute protein values to make comparable assessments of exposure to different allergens.

Assays to measure cockroach allergen exposure were originally developed with monoclonal antibodies raised against cockroach extracts in the 1990s.¹² The cockroach extracts used for assay standardization were arbitrarily assigned a certain number of units (U) for each allergen (Bla g 1 and Bla g 2) in a fixed volume of extract.¹³ Subsequently, the molecular cloning of Bla g 2 revealed that this globular protein is stable with a well-defined structure, and this allowed the measurement of this allergen in absolute units based on amino acid analysis of purified Bla g 2 (1 Unit = ~ 80 ng).^{14, 15}

However, Bla g 1 is a more complex allergen.^{10, 13, 16, 17} Molecular cloning of Bla g 1 revealed an unusual primary structure containing ~100 amino acid repeats. The gene for Bla g 1 originated from a ~100 amino acid sequence that duplicated. The sequence of the second repeat diverged from the first, and this “duplex” subsequently multiplied to include many “duplexes” of the repeated motif. If we number the repeats consecutively, comparisons between consecutive repeats are 26–29% identical, whereas comparisons among odd numbered repeats or among even numbered repeats show 96–98% identity (Online Repository Figure E1).¹⁸ The number of repeats in clones reported for group 1 cockroach allergens are between four and fourteen.^{10, 19, 20} In addition to variability in length due to genetic diversity, the allergen may appear 33–37, 28, 25, or 6 kDa depending on the method used for purification or visualization.^{13, 16, 21} The different natural Bla g 1 molecular forms result from cleavage by trypsin-like enzymes present in the cockroach gut.^{10, 13, 16} Among insects, Bla g 1-homologous proteins are found with a similar genetic structure (Online Repository Figure E1).^{18, 22}

Given the complexity of this allergen, the Bla g 1 units remained arbitrary until the allergen would be better characterized. The determination of the structure of Bla g 1, presented in this paper, reveals a basic structural unit comprised of two consecutive amino acid repeats. Knowledge of the structure of Bla g 1 allows the standardization of assays to measure this

allergen in absolute instead of relative units. In addition, the structure provides novel information about the function of this protein in the cockroach gut, the nature of the repeats, and its allergenic potential.

METHODS

Constructs

Natural Bla g 1 exists in multiple molecular forms and cloning of the gene revealed repeated amino acid sequences.^{10, 16} In order to determine the three-dimensional structure of Bla g 1, the largest repeated unit, containing a “duplex” from Bla g 1.0101 consisting of two consecutive repeats, was selected for expression in *E. coli* and crystallization, based on previous observations that proteolysis of the protein generated a similar sized molecule.¹⁷ GFP was fused to the N-terminus of the allergen to stabilize the construct and facilitate crystallization; the construct was named rBla g 1-GFP (see details in Online repository).^{23, 24}

For standardization purposes, three Bla g 1 preparations were tested containing Bla g 1 molecules of different lengths. Two were expressed by methanol induction in *Pichia pastoris* as previously described, and were named rBla g 1-PP.²⁵ These two preparations contained different molecular forms, resulting from expression of the same Bla g 1.0101 clone containing two duplexes (accession number AF072219), under different conditions (presence or absence of antibiotic zeocin for Lots 34074 and 33045, respectively). In addition, an isolated rBla g 1 construct of repeats 1 and 2 (rBla g 1-EC) was obtained by cleavage with TEV protease from a fusion with glutathione-S-transferase (GST) expressed in *E. coli*, (see details in Online repository). The reference standard containing natural Bla g 1 was a standard prepared from a *Blattella germanica* frass extract and used for ELISA that contains 10 U/mL of Bla g 1,¹³ (INDOOR Biotechnologies, Inc., Lot # 32023). The protein content of the three Bla g 1 preparations was quantified by amino acid analysis. Briefly, samples were hydrolyzed (6N HCl, 110°C, 24hr), and the resulting amino acids were separated on a strong cation exchange column, detected with a secondary reaction with ninhydrin and quantified against a known standard run in the same sequence.

For NMR and MS studies, the Bla g 1 constructs rBla g 1-EC and rBla g 1-PP Lot 34074 were used. Natural Bla g 1 (nBla g 1) was purified from cockroach frass (debris and feces produced by the cockroach) using the anti-Bla g 1 mAb 10A6 and a similar protocol to that described previously for purifying nBla g 2 from frass.^{10, 25}

Antibody binding to nBla g 1 and rBla g 1 constructs

IgE antibody binding to natural and recombinant Bla g 1-GFP used for crystallization was compared by ELISA using microtiter plates that were coated with anti-Bla g 1 mAb 10A6, as described previously.²⁵ Following incubation with the cockroach extract or the recombinant allergen, sera from cockroach allergic patients (n = 15) were added, and bound IgE was detected using biotinylated goat anti-human IgE. Details regarding the sera are in the Online Repository. Control sera were from a non-allergic patient and two mite allergic patients. An IgE standard curve was obtained with anti-Der p 2 mAb, natural Der p 2, and chimeric antibody 2B12-IgE, with a dynamic range of 0.5–250 ng IgE/mL.²⁶

For standardization purposes, a polyclonal rabbit anti-Bla g 1 was used for detection instead of IgE in an ELISA, as previously described.¹³ Dose-response curves (n = 4 per Bla g 1 preparation) were performed to compare three rBla g 1 preparations with the standard containing natural Bla g 1. Curves were fitted using the MATLAB and the equivalence between relative and absolute units at the level of the EC50 was determined.

Structural and Biochemical Characterization

Details of the crystallography, mass spectrometric analysis, and NMR procedures are given in the online repository.

RESULTS

Recombinant Bla g 1 is comparable to natural allergen

The rBla g 1 construct used for crystallization was compared to the natural allergen for human IgE antibody binding. There was an excellent quantitative correlation between IgE antibody binding to nBla g 1 and rBla g 1-GFP using sera from cockroach allergic patients ($n=15$, $r = 0.96$, $p<0.001$) (Figure 1). These results indicated that the recombinant and natural allergens were equally recognized by IgE antibodies and that the GFP molecule used in the rBla g 1-GFP construct for crystallography did not interfere with IgE antibody binding.

Bla g 1 has a novel three-dimensional structure

The structure of rBla g 1-GFP was determined by X-ray crystallography (PDB code: 4JRB, Figure E2 and Table E1 Online Repository). The structure of each 100 amino acid repeat of Bla g 1 is comprised of six short helices. The first five helices of each repeat form a planar pentagon, while the sixth helix is positioned above and across the pentagon face (Figure 2). Two consecutive pentagonal structures come together to form a duplex in which the interacting helices from each repeat run in opposite directions. The structures of the two repeats are remarkably similar (C α RMSD 1.4 Å over the first five helices) given only 26% identity of the amino acid sequences. Attempts to identify a similar fold in the protein database were unsuccessful using the program DALI.²⁷ The top 40 protein structures from the DALI search had low significance scores (5–7) but each was manually examined for similarities and none matched Bla g 1, indicating this is likely a new fold family.

Defining the allergenic unit

Based on the structure and the data in Figure 1, a duplex of two repeated sequences appears to define the minimum structural unit. The recombinant allergens expressed in *E. coli* (rBla g 1-EC) and *P. pastoris* (rBla g 1-PP, lots 34075 and 33045) were compared for murine/ rabbit IgG antibody binding to the ELISA standard containing natural Bla g 1 in extract. Note that rBla g 1-PP has two of the structural units defined in Figure 2, rBla g 1-EC has one structural unit, and the nBla g 1 has been reported as differently sized molecular forms.^{10, 12, 16, 21} Parallel and overlapping dose-response curves were obtained for the three allergen preparations by ELISA, indicating equivalent antibody binding on a g/ml basis (Figure 3A). The equivalence between absolute units and relative units of the natural Bla g 1 in the standard extract was determined to be 1 Unit equals 104 ± 1 ng ($n = 4$ per preparation; 3 preparations). Figure 3B shows an SDS-PAGE analysis of the samples used for standardization. The antibodies equivalently recognized the different preparations despite seemingly different fragmentation patterns. Mass spectrometry analysis showed that the fragmentation patterns of rBla g 1-PP resulted from partial proteolysis between and within the duplexes in the unstructured loops between the repeats (Online repository Table 2, Figure E5).

Lipid Binding of Bla g 1

Another defining feature of the Bla g 1 molecule is that the duplex of repeats 1 and 2 enclose a large spherical cavity that is lined exclusively with hydrophobic residues. The cavity is approximately $3,750 \text{ Å}^3$ (Figure 2). For comparison, in a survey of almost 600 enzymes the mean active site volume was 1072 Å^3 .²⁸ Within the cavity exists electron

density consistent with long chain fatty acids. This fragmented density was modeled with six carbon chains of variable length as well as one generic phospholipid.

In order to confirm the ligand content in the cavity, Bla g 1 molecules from various sources were analyzed for lipid content by mass spectrometry and NMR. In the *E. coli* expressed Bla g 1 constructs (rBla g 1-EC), phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) head groups with various lengths of saturated and unsaturated diacyl chains were found in abundance by MS (Online Repository, Figure E3). Based on the total ion count, the Bla g 1-EC sample contained 60% PG and 40% PE, while the *E. coli* lysate contained 36% PG and 64% PE. rBla g 1-EC had fewer of the cyclopropyl acyl chains than are commonly found in *E. coli* membranes. Figure 4A shows the ^{31}P NMR spectra of the phospholipids derived from the insoluble fraction of *E. coli* lysate, from rBla g 1-EC, and from GST as a negative control (see Online Repository Methods). The ^{31}P NMR spectrum of rBla g 1-EC showed a stronger peak intensity at the shift of PG versus PE, which is the opposite of the normal distribution in *E. coli* (Figure 4B). The MS and NMR data indicate that the Bla g 1 has a relatively greater affinity for the anionic PG relative to the other lipid species present in *E. coli*, although it is not so selective that it binds to only this type of lipid. As a negative control, the GST sample did not show any significant phospholipid peaks (Figure 4C). In summary, the data shows that Bla g 1 has an affinity for saturated or nearly saturated acyl chains and preferentially binds PG over PE from *E. coli*.

A similar MS and NMR analysis of rBla g 1 expressed in *Pichia pastoris* (rBla g 1-PP) showed that rBla g 1-PP retains phosphatidylinositol (PI), phosphatidylserine (PS), and phosphatidylcholine (PC). An NMR comparison of the relative abundance of phospholipids in an extract of *P. pastoris* (Figure 4D) and the yeast-expressed allergen (Figure 4E) showed that rBla g 1-PP preferentially retains PC and PI and does not bind phosphatidic acid (PA) nor PE. To determine if phospholipid binding was relevant in the cockroach, natural Bla g 1 purified from cockroach frass was analyzed by mass spectrometry. Surprisingly, no phospholipids were detected in nBla g 1. Instead, the excreted allergen contains primarily the non-phosphorylated fatty acids: palmitate (C16:0), stearate (C18:0) and oleate (C18:1), data not shown.

Comparisons with Per a 1 and *P. rapae* NSP protein

Bla g 1 is the first structure in a family of proteins that has been identified in cockroaches (Per a 1, 70–72% identity), butterflies (*P. rapae* NSP, 20–30% identity), and other insects.²⁹ Based on the sequence alignments in the Online Repository Figure E1, residues on the molecular surface of Bla g 1 were categorized according to their similarity with Per a 1 and *P. rapae* NSP (Figure 5). This analysis predicts that the Per a 1 and Bla g 1 surfaces will be nearly identical (Figures 5A & B). Using the Bla g 1 structure, calculations of solvent accessible area showed that 87% of the surface of cockroach Group 1 allergens is likely to be identical or similar (Figure 5C). The most notable area of sequence differences forming a contiguous surface is located in the segment from helix 4 to helix 6 (Figure 5B). In contrast, a similar comparison between Bla g 1 and a distantly related protein from *P. rapae* showed that only 39% of the surface area was predicted to be identical or similar (Figure 5D). The sequence differences between Bla g 1 and the other proteins were not limited solely to the external surfaces, but were equally distributed among buried and surface exposed residues in Bla g 1.

DISCUSSION

In this study we present an unusual example of how the determination of the molecular structure of an allergen, Bla g 1, allows progress in the assessment of allergen exposure associated with respiratory symptoms. Over the years, measurement of environmental

exposure to cockroach allergens has been essential to risk assessment for inner city asthma studies.^{2-4, 11, 30} Allergen measurements have also been important for integrative pest management to reduce allergen exposure and improve asthma.³¹⁻³³ Regarding diagnosis, a high variability in Bla g 1 and Bla g 2 content has been found in commercial cockroach skin test extracts, with up to 200-fold difference in Bla g 1 levels (4.7 to 1085 U/ml; n = 6 commercial extracts).³⁴ However, in all the previous studies, the levels of allergen exposure were expressed in relative units.

A process of assay standardization in absolute units is presented for Bla g 1. This usually requires the comparison of the Bla g 1 standard used in the environmental assays with a known amount of purified allergen using a reliable protein quantification technique. Natural Bla g 1 is a very unusual allergen that exists in multiple molecular forms, resulting from trypsin cleavage of the protein (Online Repository).¹⁰ A molecule that breaks down into fragments would be typically difficult or impossible to standardize due to the consequent abrogation of antibody epitopes. The importance of the new structural motif for the quantitation of Bla g 1 is that it provides a rationale to understand why in solution the molecule preserves antibody epitopes and yet appears fragmented on a denaturing gel. The 3D structure of repeats 1 & 2 determined by X-ray crystallography is novel, and the basic structural unit has the shape of a capsule or bead with a large internal cavity that accommodates lipid ligands. This suggests that natural Bla g 1 might resemble beads on a string, attached by flexible linkers (Figure E5). The multiple molecular forms of Bla g 1 observed by SDS-PAGE resemble fragments of a broken necklace, each with different numbers of beads or half beads since the long flexible linker between the repeats is also vulnerable to proteolysis (Figure E5). The crystallographic and antibody binding results suggest that the capsule structure is likely maintained in solution even with limited proteolysis. This scenario is very different from a globular molecule that degrades into different fragments with a consequent loss of epitopes, which makes standardization impossible.

We showed by ELISA dose-response curves, that the behavior of different Bla g 1 preparations was very similar regarding antibody binding and comparable to the natural allergen despite a different molecular fragment composition. The basic structural unit showed a good correlation with the natural allergen regarding IgE antibody binding as also reported for the *Pichia* expressed allergen.¹⁷ In addition, the four bands of the *Pichia* expressed Bla g 1 were shown to bind IgE antibodies,¹⁷ which is consistent with the fact that allergic sensitization occurs by inhalation of fragmented natural Bla g 1.¹⁰ In summary, a similar fragmentation pattern occurs in natural and recombinant allergens leading to molecules that are relevant for patient IgE and monoclonal antibody binding.

One arbitrary unit (U) of Bla g 1 was calculated to be equivalent to 104 ng, as absolute units. Bla g 2, with a similar equivalence of 80 ng per unit, has been reported to be very potent, inducing IgE sensitization at exposure levels that are 10–100 fold lower than other common indoor allergens, such as dust mite and cat.³⁵ The threshold levels for risk of sensitization and asthma morbidity (2 and 8 Units/g of dust, respectively) would be equivalent to 208 and 832 ng/g of dust for Bla g 1. These values were comparable to levels of 160 and 640 ng/g of dust for Bla g 2, but lower than values reported for dust mite and cat, confirming the potency of cockroach allergens. In summary, these results indicate that Bla g 1 and Bla g 2 have a similar allergenic potential, which is high compared to other indoor allergens, and allow direct comparison of allergen exposure measurements for different allergens.⁶

The structure of Bla g 1 is the first to be described for this group of insect proteins. Electron density in the cavity of Bla g 1 resembled phospholipids, and this in turn led us to investigate the internal ligands in Bla g 1 by MS and NMR. Both methods indicated

phospholipids with various head groups could be retained by Bla g 1. MS analysis primarily identified saturated or monounsaturated acyl chains, as well as a few cyclopropyl acyl chains. Bla g 1.01 appears partially selective for certain phospholipid head groups including PC, PI, and PG. Promiscuous binding of lipids would be consistent with the idea of Bla g 1 being important for general nutrient uptake in cockroach: 1) the allergen is expressed in the midgut and excreted in fecal particles;^{10, 36} 2) the expression of Bla g 1 increases with food intake;³⁷ 3) knockdown of Bla g 1 expression using RNAi blocked nutrient absorption;³⁸ 4) homologous proteins in mosquitoes (ANG12 and AEG12) are up-regulated in the midgut after a blood meal,³⁹ and 5) AEG12 was localized by immuno-gold staining to microvilli in the gut.⁴⁰ Based on our results it is conceivable that Bla g 1 may be involved in the uptake of dietary lipids. The binding of certain lipids (palmitate, stearate and oleate) and not others to Bla g 1 from cockroach frass can be interpreted as a selective removal and excretion of these lipids into the feces, to facilitate retention and absorption of other, more nutritive compounds. Alternatively, homologous butterfly proteins allow the larvae to excrete toxic nitriles as an adaptation against the major chemical defense of their host plants.^{22, 41} In either case, the structure and data shown here suggests this protein fold could accommodate a variety of hydrophobic ligands for various physiological purposes.

A common feature of many allergens is the ability to bind hydrophobic ligands, which are potent stimulators of the innate immune system.⁴² The lipids that were found by MS in the cockroach frass are fatty acids that are known to activate TLR-4 and TLR-2.⁴³ Low-level stimulation of the TLR-4 pathway by LPS was shown to drive the immune response towards Th2,^{44, 45} and TLR-2 stimulation in the presence of allergen enhances Th2 responses.^{46, 47} Lipid adjuvants, like those found here associated with Bla g 1, can strongly influence the innate immune response and subsequently skew the adaptive response towards allergy. However, this remains to be demonstrated for nBla g 1 and fatty acids.

There is currently very little literature on specific epitopes of group 1 cockroach allergens except for two small scale studies with peptides of Per a 1.^{48, 49} Based on the ~70% sequence identity between Per a 1 and Bla g 1, the homologous allergen from the American cockroach is expected to have a similar fold and similar ligand specificity, and is unlikely to be a membrane spanning protein as was previously suggested.⁵⁰ Looking at the unique fold of Bla g 1, there are very few long stretches of amino acids that are significantly buried because there isn't a canonical protein hydrophobic core. The hollow spherical shape exposes many more residues as potential IgE epitopes compared with a globular protein of similar mass, hence almost the entire sequence is potentially antigenic. The structure presented here should be very useful for future analysis of the antigenic surface and the design of Bla g 1 molecules for safer and more effective immunotherapeutics. It also provides insight into the digestive role of Bla g 1, with the potential adjuvant effects of lipids in allergenicity. The Bla g 1 structure facilitated measurement of the conversion factor from the previous arbitrary units to absolute units for the standardization of assays to assess allergen environmental exposure and allergen content in extracts used for diagnosis. These measurements are essential for understanding allergen levels associated with risk of sensitization and asthma morbidity, and for reducing allergen exposure and managing disease.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by Research Project Number Z01- ES102885-01 to REL, Z01-ES50161 to KBT, and ZIA- ES102645 to LCP in the Intramural Research Program of the National Institute of Environmental Health Sciences, National Institutes of Health. Research reported in this publication was supported in part by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Award Number R01AI077653 (AP and MDC). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Use of the Advanced Photon Source, Argonne National Laboratory, was supported by the US department of Energy, Basic Energy Sciences, Office of Science, under Contract No. W-31-109-Eng-38. The authors appreciate the assistance of Jason Williams with mass spectrometry.

Abbreviations

MS	Mass Spectrometry
GFP	Green fluorescent protein
GST	Glutathione-S-transferase
TEV	Tobacco Etch Virus
rBla g 1-GFP	recominant Bla g 1 with an N-terminal GFP
rBla g 1-EC	recombinant Bla g 1 expressed in <i>E. coli</i>
rBla g 1-PP	recombinant Bla g 1 expressed in <i>Pichia pastoris</i>
PA	Phosphatidic acid
PG	Phosphatidylglycerol
PE	Phosphatidylethanolamine
PC	Phosphatidylcholine
PS	Phosphatidylserine
PI	Phosphatidylinositol
LPC	Lsophosphatidylcholine
CL	Cardiolipin
NSP	nitrile specifier protein
PBS	phosphate buffered saline
TLR	Toll-like receptor

References

1. Bernton HS, Brown H. Insect Allergy-Preliminary Studies of Cockroach. *Journal of Allergy*. 1964; 35:506–13. [PubMed: 14226309]
2. Rosenstreich DL, Eggleston P, Kattan M, Baker D, Slavin RG, Gergen P, et al. The role of cockroach allergy and exposure to cockroach allergen in causing morbidity among inner-city children with asthma. *New England Journal of Medicine*. 1997; 336:1356–63. [PubMed: 9134876]
3. Gruchalla RS, Pongracic J, Plaut M, Evans R, Visness CM, Walter M, et al. Inner City Asthma Study: Relationships among sensitivity, allergen exposure, and asthma morbidity. *Journal of Allergy and Clinical Immunology*. 2005; 115:478–85. [PubMed: 15753892]
4. Matsui EC, Wood RA, Rand C, Kanchanaraksa S, Swartz L, Curtin-Brosnan J, et al. Cockroach allergen exposure and sensitization in suburban middle-class children with asthma. *Journal of Allergy and Clinical Immunology*. 2003; 112:87–92. [PubMed: 12847484]
5. Perzanowski MS, Chew GL, Divjan A, Jung KH, Ridder R, Tang DL, et al. Early-life cockroach allergen and polycyclic aromatic hydrocarbon exposures predict cockroach sensitization among

- inner-city children. *Journal of Allergy and Clinical Immunology*. 2013; 131:886–93. [PubMed: 23391330]
6. Platts-Mills TA, Vervloet D, Thomas WR, Aalberse RC, Chapman MD. Indoor allergens and asthma: report of the Third International Workshop. *The Journal of allergy and clinical immunology*. 1997; 100:S2–24. [PubMed: 9438476]
 7. Salo PM, Arbes SJ, Crockett PW, Thorne PS, Cohn RD, Zeldin DC. Exposure to multiple indoor allergens in US homes and its relationship to asthma. *Journal of Allergy and Clinical Immunology*. 2008; 121:678–84. [PubMed: 18255132]
 8. Matsui EC, Sampson HA, Bahnson HT, Gruchalla RS, Pongratic JA, Teach SJ, et al. Allergen-specific IgE as a biomarker of exposure plus sensitization in inner-city adolescents with asthma. *Allergy*. 2010; 65:1414–22. [PubMed: 20560910]
 9. Arruda LK, Vailes LD, Ferriani VPL, Santos BR, Pomes A, Chapman MD. Cockroach allergens and asthma. *Journal of Allergy and Clinical Immunology*. 2001; 107:419–28. [PubMed: 11240940]
 10. Pomes A, Melen E, Vailes LD, Retief JD, Arruda LK, Chapman MD. Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. *Journal of Biological Chemistry*. 1998; 273:30801–7. [PubMed: 9804858]
 11. Eggleston PA, Rosenstreich D, Lynn M, Gergen P, Baker D, Kattan M, et al. Relationship of indoor allergen exposure to skin test sensitivity in inner-city children with asthma. *Journal of Allergy and Clinical Immunology*. 1998; 102:563–70. [PubMed: 9802363]
 12. Pollart SM, Smith TF, Morris EC, Gelber LE, Plattsmills TAE, Chapman MD. Environmental Exposure to Cockroach Allergens - Analysis with Monoclonal Antibody-Based Enzyme Immunoassays. *Journal of Allergy and Clinical Immunology*. 1991; 87:505–10. [PubMed: 1993810]
 13. Pollart SM, Mullins DE, Vailes LD, Hayden ML, Plattsmills TAE, Sutherland WM, et al. Identification, Quantitation, and Purification of Cockroach Allergens Using Monoclonal-Antibodies. *Journal of Allergy and Clinical Immunology*. 1991; 87:511–21. [PubMed: 1993811]
 14. Arruda LK, Vailes LD, Mann BJ, Shannon J, Fox JW, Vedvick TS, et al. Molecular-Cloning of a Major Cockroach (*Blattella-Germanica*) Allergen, Bla-G-2 - Sequence Homology to the Aspartic Proteases. *Journal of Biological Chemistry*. 1995; 270:19563–8. [PubMed: 7642642]
 15. Gustchina A, Li M, Wunschmann S, Chapman MD, Pomes A, Wlodawer A. Crystal structure of cockroach allergen Bla g 2, an unusual zinc binding aspartic protease with a novel mode of self-inhibition. *Journal of Molecular Biology*. 2005; 348:433–44. [PubMed: 15811379]
 16. Schou C, Lind P, Fernandezcaldas E, Lockey RF, Lowenstein H. Identification and Purification of an Important Cross-Reactive Allergen from American (*Periplaneta-Americana*) and German (*Blattella-Germanica*) Cockroach. *Journal of Allergy and Clinical Immunology*. 1990; 86:935–46. [PubMed: 2262648]
 17. Pomes A, Vailes LD, Helm RM, Chapman MD. IgE reactivity of tandem repeats derived from cockroach allergen, Bla g 1. *European Journal of Biochemistry*. 2002; 269:3086–92. [PubMed: 12071974]
 18. Pomes A, Wunschmann S, Hindley J, Vailes LD, Chapman MD. Cockroach allergens: Function, structure and allergenicity. *Protein and Peptide Letters*. 2007; 14:960–9. [PubMed: 18220993]
 19. Yang CY, Wu JD, Wu CH. Sequence analysis of the first complete cDNA clone encoding an American cockroach Per a 1 allergen. *Biochimica Et Biophysica Acta-Gene Structure and Expression*. 2000; 1517:153–8.
 20. Helm R, Cockrell G, Stanley JS, Brenner RJ, Burks W, Bannon GA. Isolation and characterization of a clone encoding a major allergen (Bla g Bd90K) involved in IgE-mediated cockroach hypersensitivity. *Journal of Allergy and Clinical Immunology*. 1996; 98:172–80. [PubMed: 8765832]
 21. Twarog FJ, Picone FJ, Strunk RS, So J, Colten HR. Immediate Hypersensitivity to Cockroach - Isolation and Purification of Major Antigens. *Journal of Allergy and Clinical Immunology*. 1977; 59:154–60. [PubMed: 64484]
 22. Fischer HM, Wheat CW, Heckel DG, Vogel H. Evolutionary origins of a novel host plant detoxification gene in butterflies. *Molecular Biology and Evolution*. 2008; 25:809–20. [PubMed: 18296701]

23. Mueller GA, Edwards LL, Aloor JJ, Fessler MB, Glesner J, Pomes A, et al. The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. *Journal of Allergy and Clinical Immunology*. 2010; 125:909–17. [PubMed: 20226507]
24. Mueller GA, Gosavi RA, Pomes A, Wunschmann S, Moon AF, London RE, et al. Ara h 2: crystal structure and IgE binding distinguish two subpopulations of peanut allergic patients by epitope diversity. *Allergy*. 2011; 66:878–85. [PubMed: 21255036]
25. Pomes A, Chapman MD, Vailes LD, Blundell TL, Dhanaraj V. Cockroach allergen Bla g 2 - Structure, function, and implications for allergic sensitization. *American Journal of Respiratory and Critical Care Medicine*. 2002; 165:391–7. [PubMed: 11818327]
26. Schuurman J, Perdok GJ, Lourens TE, Parren PWHI, Chapman MD, Aalberse RC. Production of a mouse/human chimeric IgE monoclonal antibody to the house dust mite allergen Der p 2 and its use for the absolute quantification of allergen-specific IgE. *Journal of Allergy and Clinical Immunology*. 1997; 99:545–50. [PubMed: 9111501]
27. Holm L, Sander C. The FSSP database: fold classification based on structure-structure alignment of proteins. *Nucleic Acids Research*. 1996; 24:206–9. [PubMed: 8594580]
28. Jimenez-Morales D, Liang J, Eisenberg B. Ionizable side chains at catalytic active sites of enzymes. *European Biophysics Journal with Biophysics Letters*. 2012; 41:449–60. [PubMed: 22484856]
29. Dostalova A, Votypka J, Favreau AJ, Barbian KD, Volf P, Valenzuela JG, et al. The midgut transcriptome of *Phlebotomus (Larroussius) perniciosus*, a vector of *Leishmania infantum*: comparison of sugar fed and blood fed sand flies. *Bmc Genomics*. 2011;12. [PubMed: 21211037]
30. deBlay F, Sanchez J, Hedelin G, PerezInfante A, Verot A, Chapman M, et al. Dust and airborne exposure to allergens derived from cockroach (*Blattella germanica*) in low-cost public housing in Strasbourg (France). *Journal of Allergy and Clinical Immunology*. 1997; 99:107–12. [PubMed: 9003218]
31. Arbes SJ, Sever M, Mehta J, Gore JC, Schal C, Vaughn B, et al. Abatement of cockroach allergens (Bla g 1 and Bla g 2) in low-income, urban housing: Month 12 continuation results. *Journal of Allergy and Clinical Immunology*. 2004; 113:109–14. [PubMed: 14713915]
32. Gore JC, Schal C. Cockroach allergen biology and mitigation in the indoor environment. *Annual Review of Entomology*. 2007; 52:439–63.
33. Morgan WJ, Crain EF, Gruchalla RS, O'Connor GT, Kattan M, Evans RI, et al. Results of a home-based environmental intervention among urban children with asthma. *New England Journal of Medicine*. 2004; 351:1068–80. [PubMed: 15356304]
34. Patterson ML, Slater JE. Characterization and comparison of commercially available German and American cockroach allergen extracts. *Clinical and Experimental Allergy*. 2002; 32:721–7. [PubMed: 11994096]
35. Sporik R, Squillace SP, Ingram JM, Rakes G, Honsinger RW, Platts-Mills TAE. Mite, cat, and cockroach exposure, allergen sensitisation, and asthma in children: a case-control study of three schools. *Thorax*. 1999; 54:675–80. [PubMed: 10413718]
36. Gore JC, Schal C. Gene expression and tissue distribution of the major human allergen Bla g 1 in the German cockroach, *Blattella germanica* L. (Dictyoptera: Blattellidae). *Journal of Medical Entomology*. 2004; 41:953–60. [PubMed: 15535627]
37. Gore JC, Schal C. Expression, production and excretion of Bla g 1, a major human allergen, in relation to food intake in the German cockroach, *Blattella germanica*. *Medical and Veterinary Entomology*. 2005; 19:127–34. [PubMed: 15958021]
38. Suazo A, Gore C, Schal C. RNA interference-mediated knock-down of Bla g 1 in the German cockroach, *Blattella germanica* L. implicates this allergen-encoding gene in digestion and nutrient absorption. *Insect Molecular Biology*. 2009; 18:727–36. [PubMed: 19758414]
39. Nolan T, Petris E, Muller HM, Cronin A, Catteruccia F, Crisanti A. Analysis of two novel midgut-specific promoters driving transgene expression in *Anopheles stephensi* mosquitoes. *PloS one*. 2011; 6:e16471. [PubMed: 21326609]
40. Shao L, Devenport M, Fujioka H, Ghosh A, Jacobs-Lorena M. Identification and characterization of a novel peritrophic matrix protein, Ae-Aper50, and the microvillar membrane protein, AEG12,

- from the mosquito, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology*. 2005; 35:947–59. [PubMed: 15978997]
41. Wittstock U, Agerbirk N, Stauber EJ, Olsen CE, Hippler M, Mitchell-Olds T, et al. Successful herbivore attack due to metabolic diversion of a plant chemical defense. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:4859–64. [PubMed: 15051878]
 42. Karp CL. Guilt by intimate association: what makes an allergen an allergen? *The Journal of allergy and clinical immunology*. 2010; 125:955–60. [PubMed: 20381850]
 43. Huang SR, Rutkowski JM, Snodgrass RG, Ono-Moore KD, Schneider DA, Newman JW, et al. Saturated fatty acids activate TLR-mediated proinflammatory signaling pathways. *Journal of Lipid Research*. 2012; 53:2002–13. [PubMed: 22766885]
 44. Eisenbarth SC, Piggott DA, Huleatt JW, Visintin I, Herrick CA, Bottomly K. Lipopolysaccharide-enhanced, toll-like receptor 4-dependent T helper cell type 2 responses to inhaled antigen. *The Journal of experimental medicine*. 2002; 196:1645–51. [PubMed: 12486107]
 45. Piggott DA, Eisenbarth SC, Xu L, Constant SL, Huleatt JW, Herrick CA, et al. MyD88-dependent induction of allergic Th2 responses to intranasal antigen. *Journal of Clinical Investigation*. 2005; 115:459–67. [PubMed: 15650773]
 46. Chisholm D, Libet L, Hayashi T, Horner AA. Airway peptidoglycan and immunostimulatory DNA exposures have divergent effects on the development of airway allergen hypersensitivities. *Journal of Allergy and Clinical Immunology*. 2004; 113:448–54. [PubMed: 15007346]
 47. Redecke V, Hacker H, Datta SK, Fermin A, Pitha PM, Broide DH, et al. Cutting edge: activation of Toll-like receptor 2 induces a Th2 immune response and promotes experimental asthma. *Journal of Immunology*. 2004; 172:2739–43.
 48. Tsai WJ, Liu CH, Chen ST, Yang CY. Identification of the antigenic determinants of the American cockroach allergen Per a 1 by error-prone PCR. *Journal of Immunological Methods*. 2003; 276:163–74. [PubMed: 12738370]
 49. Wu CH, Lee MF, Yang JS, Tseng CY. IgE-binding epitopes of the American cockroach Per a 1 allergen. *Molecular Immunology*. 2002; 39:459–64. [PubMed: 12413697]
 50. Diraphat P, Sookrung N, Chaicumpa W, Pumhirun P, Vichyanond P, Tapchaisri P, et al. Recombinant American cockroach component, Per a 1, reactive to IgE of allergic Thai patients. *Asian Pacific Journal of Allergy and Immunology*. 2003; 21:11–20. [PubMed: 12931746]

Clinical Implications

- Bla g 1 like many allergens binds hydrophobic ligands. It is thought that these ligands may contribute to allergenicity as adjuvants.
- The structurally defined basic allergenic unit provides absolute standardization for cockroach exposure assessment in asthma.

Figure 1.
Correlation between IgE antibody binding to rBla g 1-GFP and natural Bla g 1, using sera from cockroach allergic patients measured by ELISA.

Figure 2.

Bla g 1 Structure. A) "Top View" shows the first 5 helices forming a pentagon. The dashed red arrow indicates residues with missing electron density between helix 6 and 1. The hydrophobic cavity is indicated by red mesh. B) Topology diagram showing the interactions of the helices. Electron density in the cavity was modeled with PA (panel C) and short acyl chains, rendered as sticks in panel D. D) The "Side View" shows how the helices interact between the repeats.

Figure 3.

Standardization of nBla g 1 Units. A) Binding curves with rabbit polyclonal antibodies to rBla g 1-EC (green circles, n=4), rBla g 1-PP Lot 34074 (blue circles, n=4), rBla g 1-PP Lot 33045 (magenta circles, n=4), and nBla g 1 from the standard cockroach extract (Lot # 32023, red circles, n=4). Fits to the curves to determine the EC₅₀ are shown for nBla g 1 (red line), and the recombinant constructs, which were fit as one experimental data set, n=12 (black line). B) Silver stained SDS-PAGE of protein samples used in the standardization. Note that the concentration loaded on the gel was ~1,000 fold greater than the highest data point in the ELISA. A small amount of uncleaved (GST)-Bla g 1 is apparent in lane 4 at 50 kDa.

Figure 4.

Comparison of the phospholipid content using ^{31}P NMR of A) *E. coli* extract (positive control), B) rBla g 1-EC, C) GST expressed in *E. coli* (negative control), D) *P. pastoris* extract, and E) rBla g 1-PP. The ^{31}P NMR peaks were assigned based on literature values, relative abundance, and the MS data in this manuscript.

Figure 5.

Surface residue comparison between Bla g 1, Per a 1 and *P. rapae* NSP. A) and B) show the solvent accessible surface of Bla g 1 in two different orientations colored by residue identity (white), similarity (cyan) and differences (hot pink) with Per a 1. Visible helices are labeled by number beneath the semi-transparent surface. C) A pie chart of the total solvent accessible surface categorized similarly. D) A similar comparison with the distantly related *P. rapae* NSP.

Table I

Crystallographic data statistics		
data set	rBla g 1-NGFP (high)	rBla g 1-NGFP (low)
unit cell	a=91.77 Å, b=141.92 Å, c=93.47 Å; $\alpha = \beta = \gamma = 90^\circ$	a=92.67 Å, b=141.62 Å, c=93.53 Å; $\alpha = \beta = 90^\circ$, $\gamma = 90.4^\circ$
Space Group	C222 ₁	C2
Resolution (Å)	50.0 – 2.4	50–2.8
# of observations	165,823	103,638
unique reflections	23,505	29,002
Rsym(%) (last shell) ¹	8.9 (45.8)	6.8 (25.1)
I/ σ I (last shell)	14.9 (2.5)	11.6 (2.9)
Mosaicity range	0.69–0.93	0.27–0.58
completeness(%) (last shell)	99.3 (94.0)	97.8 (79.4)
<u>Refinement statistics</u>		
Rcryst(%) ²	19.8	
Rfree(%) ³	24.6	
# of waters	100	
Overall Mean B (Å ²)	58.2	
Average for: GFP	51.4	
Bla g 1	67.6	
ligands	75.3	
water	51.8	
<u>r.m.s. deviation from ideal values</u>		
bond length (Å)	0.007	
bond angle (°)	1.1	
dihedral angle (°)	16.8	
<u>Ramachandran Statistics⁴</u>		
Residues in:		
favored (98%) regions (%)	99.2	
allowed (>99.8%) regions (%)	100	

¹ $R_{\text{sym}} = \frac{\sum_i |I_i - \langle I \rangle|}{\sum_i I_i}$ where I_i is the intensity of the i th observation and $\langle I \rangle$ is the mean intensity of the reflection.

² $R_{\text{cryst}} = \frac{\| |F_o| - |F_c| \|}{\| |F_o| \|}$ calculated from working data set.

³ R_{free} was calculated from 5% of data randomly chosen not to be included in refinement.

⁴Ramachandran results were determined by MolProbity.

Table 2

Mass Spectrometry Analysis of rBla g 1-PP

Mass (Da)	Repeat(s)	Residues (Bla g 1.0101)
9811.4	1 or 4	23–112, or 318–403
9967.8	1	38–128
10099.0	4	322–411
10237.1	2	108–198
10284.2	2	133–225
10374.0	1,3, or 4	38–131, 221–315, or 321–412
10505.0	2 or 3	129–221, or 218–313
10666.3	3	224–320
10965.0	1 or 4	12–111, or 316–412
11260.8	1	19–121