The novel structure of the 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, BA,[a] Jill Glesner, BS,[b] Andrea F. Moon, MS,[a] Martin D. Chapman, PhD,[b] Kenneth B. Tomer, PhD,[a] Robert E. London, PhD,[a] and Anna Pomés, PhD[b] *Research Triangle Park, NC, and Charlottesville, Va*

Background: Sensitization to cockroach allergens is a major risk factor for asthma. The cockroach allergen Bla g 1 has multiple repeats of approximately 100 amino acids, but the fold of the protein and its biological function are unknown.
Objective: We sought to determine the structure of Bla g 1, investigate the implications for allergic disease, and standardize cockroach exposure assays.
Methods: nBla g 1 and recombinant constructs were compared by using ELISA with specific murine IgG and human IgE. The structure of Bla g 1 was determined by x-ray crystallography. Mass spectrometry and nuclear magnetic resonance spectroscopy were used to examine the ligand-binding properties of the allergen.
Results: The structure of an rBla 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 nonspecific 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. (J Allergy Clin Immunol 2013;132:1420-6.)

Key words: *Allergen, asthma, Bla g 1, cockroach, structure, ligand-binding proteins, exposure assessment*

The first report on cockroach allergy dates back to 1964, when Bernton and Brown[1] described an association between sensitization to cockroach allergens and asthma morbidity in New York. Subsequent studies have confirmed this original finding, especially among lower socioeconomic populations but also in suburban middle-class homes of asthmatic children.[2-4] Rates of cockroach sensitization are highest in the northeastern United States, 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 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 2 major allergens to be identified.[9] Both are secreted in the cockroach digestive tract, and sensitization occurs through 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 ($\leq$2 U/g of dust) was determined to be a risk factor for sensitization; however, the risk factor plateaus at greater than 4 U/g of dust.[2,11] A seminal study by Rosenstreich et al[2,11] established 8 U/g of dust as threshold levels of "morbidity due to asthma," as part of a National Cooperative Inner-City Asthma Study. 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 mAbs raised against cockroach extracts in the 1990s.[12] The cockroach extracts used for assay standardization were arbitrarily assigned a certain number of units 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 = approximately 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 repeats of approximately 100 amino acids. The gene for

From [a]the Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, and [b]Indoor Biotechnologies, Charlottesville.
*These authors contributed equally to this work.
Supported in part by Research Project no. Z01- ES102885-01 to R.E.L., Z01-ES50161 to K.B.T., and ZIA-ES102645 to L.C.P. 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 no. R01AI077653 (to A.P. and M.D.C.). 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.
Disclosure of potential conflict of interest: A. Pomés and M. D. Chapman have received a grant from the National Institute of Allergy and Infectious Disease (NIAID) and are employed by Indoor Biotechnologies, Inc. M. D. Chapman is co-owner and shareholder of Indoor Biotechnologies, Inc, and receives royalties from the University of Virginia. The rest of the authors declare that they have no relevant conflicts of interest.
Received for publication January 30, 2013; revised May 28, 2013; accepted for publication June 19, 2013.
Available online July 31, 2013.
Corresponding author: Geoffrey A. Mueller, PhD, National Institute of Environmental Health Sciences, 111 T.W. Alexander Dr, MD-MR-01, Research Triangle Park, NC 27709. E-mail: Mueller3@niehs.nih.gov.
0091-6749
http://dx.doi.org/10.1016/j.jaci.2013.06.014

Abbreviations used

GFP: Green fluorescent protein
GST: Glutathione-S-transferase
MS: Mass spectrometry
NMR: Nuclear magnetic resonance spectroscopy
NSP: Nitrile-specifier protein
PC: Phosphatidylcholine
PE: Phosphatidylethanolamine
PG: Phosphatidylglycerol
PI: Phosphatidylinositol
rBla g 1–EC: Recombinant Bla g 1 expressed in *Escherichia coli*
rBla g 1–GFP: Recombinant Bla g 1 with an N-terminal GFP
rBla g 1–PP: Recombinant Bla g 1 expressed in *Pichia pastoris*
TEV: Tobacco etch virus
TLR: Toll-like receptor

Bla g 1 originated from an approximately 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, consecutive repeats are 26% to 29% identical, whereas comparisons among odd-numbered repeats or among even-numbered repeats show 96% to 98% identity (see Fig E1 in this article's Online Repository at www.jacionline.org).[18] The number of repeats in clones reported for group 1 cockroach allergens is between 4 and 14.[10,19,20] In addition to variability in length due to genetic diversity, the allergen can appear to be 33 to 37, 28, 25, or 6 kDa depending on the method used for purification or visualization.[13,16,21] The different nBla 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 (see Fig E1).[18,22]

Given the complexity of this allergen, the Bla g 1 units remained arbitrary until the allergen could be better characterized. The determination of the structure of Bla g 1, as presented in this article, reveals a basic structural unit comprised of 2 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

nBla g 1 exists in multiple molecular forms, and cloning of the gene revealed repeated amino acid sequences.[10,16] The largest repeated unit, containing a "duplex" from Bla g 1.0101 consisting of 2 consecutive repeats, was selected for expression in *Escherichia coli* and crystallization (based on previous observations that proteolysis of the protein generated a similarly sized molecule) to determine the 3-dimensional structure of Bla g 1.[17] Green fluorescent protein (GFP) was fused to the N-terminus of the allergen to stabilize the construct and facilitate crystallization; the construct was named rBla g 1 with an N-terminal GFP (rBla g 1–GFP; see Fig E2 and details in the Methods section in this article's Online Repository at www.jacionline.org).[23,24]

For standardization purposes, 3 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 recombinant Bla g 1 expressed in *P pastoris* (rBla g 1–PP).[25] These 2 preparations contained different molecular forms, resulting from expression of the same Bla g 1.0101 clone containing 2 duplexes (accession no. 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 expressed in *E coli* [rBla g 1-EC]) was obtained by means of cleavage with tobacco etch virus (TEV) protease from a fusion with glutathione-S-transferase (GST) expressed in *E coli* (see details in the Methods section in this article's Online Repository). The reference standard containing nBla g 1 was a standard prepared from a *Blattella germanica* frass extract and used for ELISA that contains 10 U/mL Bla g 1 (lot no. 32023; INDOOR Biotechnologies, Charlottesville, Va).[13] The protein content of the 3 Bla g 1 preparations was quantified by using amino acid analysis. Briefly, samples were hydrolyzed (6N HCl, 110°C, 24 hours), 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 nuclear magnetic resonance spectroscopic (NMR) and mass spectrometric (MS) studies, the Bla g 1 constructs rBla g 1–EC and rBla g 1–PP (lot 34074) were used. nBla g 1 was purified from cockroach frass (debris and feces produced by the cockroach) by 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 using ELISA with microtiter plates that were coated with anti–Bla g 1 mAb 10A6, as described previously.[25] After incubation with the cockroach extract or the recombinant allergen, sera from patients with cockroach allergy (n = 15) were added, and bound IgE was detected by using biotinylated goat anti-human IgE. Details regarding the sera are shown in the Methods section in this article's Online Repository. Control sera were from a nonallergic patient and 2 patients with mite allergy. An IgE standard curve was obtained with anti–Der p 2 mAb, nDer p 2, and the chimeric antibody 2B12-IgE, with a dynamic range of 0.5 to 250 ng IgE/mL.[26]

For standardization purposes, a polyclonal rabbit anti–Bla g 1 antibody 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 3 rBla g 1 preparations with the standard containing nBla g 1. Curves were fitted by using MATLAB, and the equivalence between relative and absolute units at the half maximal effective concentration was determined.

Structural and biochemical characterization

Details of the crystallography, MS analysis, and NMR procedures are presented in the Methods section in this article's Online Repository.

RESULTS

rBla g 1 is comparable with natural allergen

The rBla g 1 construct used for crystallization was compared with 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 by using sera from patients with cockroach allergy (n = 15, $r = 0.96$, $P < .001$, Fig 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 3-dimensional structure

The structure of rBla g 1–GFP was determined by using x-ray crystallography (PDB code: 4JRB, see Fig E3 and Table E1 in this