Molecular cloning of Per a 1 and definition of the cross-reactive Group 1 cockroach allergens

Erik Melén, BSc, Anna Pomés, PhD, Lisa D. Vailes, MS, L. Karla Arruda, MD, PhD, and Martin D. Chapman, PhD

Charlottesville, Va

Background: Sensitization to allergens produced by German and American cockroaches is strongly associated with the cause of asthma. Most of the cockroach allergens identified to date have been species specific.

Objective: The aim of this study was to identify and sequence cross-reactive cockroach allergens.

Methods: A Periplaneta americana cDNA library was screened with IgE antibody from patients in the United States who were allergic to cockroach and who were sensitized to Blattella germanica.

Results: A cDNA clone was isolated that contained an 870-bp sequence with a 695-bp open reading frame, encoding a 231 amino acid protein, molecular weight 26.2 kd. Plaque immunoassays using anti-Bla g 1 and anti-Per a 1 mAbs and a panel of human IgE antibodies showed that the protein expressed by these clones was Per a 1. Sequence homology searches showed that Per a 1 was homologous to 5 previously reported, but unidentified, sequences from B germanica and P americana. These sequences encoded proteins with multiple molecular sizes containing approximately 100 amino acid repeats. The Per a 1 sequence also showed 31% identity to a mosquito precursor protein, ANG12, which may be involved in digestion. The Per a 1 cDNA was expressed in Pichia pastoris to produce purified recombinant allergen (yield, 14 mg/L).

Conclusion: The results define the molecular structure and antigenic relationships between a new family of cross-reactive “Group 1” allergens produced by both P americana and B germanica. These recombinant allergens and specific mAbs will provide tools to improve the diagnosis and treatment of allergic diseases caused by cockroaches. (J Allergy Clin Immunol 1999;103:859-64.)

Key words: Allergen, amino acid sequence, American cockroach, asthma

Cockroaches produce multiple allergens that cause IgE antibody responses, and sensitization to these allergens is associated with allergic diseases, especially asthma.\textsuperscript{1-5} Epidemiologic studies have shown a high prevalence of sensitization to cockroach allergens among lower socioeconomic groups living in cockroach-infested housing, and they have also shown that sensitization to cockroaches is a risk factor for hospital admission with asthma.\textsuperscript{4-6} The National Cooperative Inner City Asthma Study recently reported that allergy to cockroaches, coupled with exposure to cockroach allergens, was responsible for increased asthma morbidity among inner-city children, as judged by hospitalization rates, days of wheezing, and loss of schooling.\textsuperscript{7}

In the United States, the principal domiciliary cockroach species are German and American cockroaches (Blattella germanica and Periplaneta americana). Over the past 10 years, there have been extensive studies to identify and clone cockroach allergens.\textsuperscript{8-17} Although there is cross-reactivity between German and American cockroach extracts on skin testing, all the cockroach allergens that have been cloned are species specific.\textsuperscript{11} These include the B germanica allergens, Bla g 2, Bla g 4, and Bla g 5, and the P americana allergen, Per a 3.\textsuperscript{12-14,17} The only cross-reactive allergens that have been defined to date are Bla g 1 and Per a 1, which were originally purified by immunochemical techniques and by the use of mAbs.\textsuperscript{9,10} The structure of these allergens had not been determined. A mAb immunoassay for Bla g 1 has been used as a marker for cockroach exposure, and patients allergic to cockroaches in the United States are primarily exposed to B germanica.\textsuperscript{7,10,18,19} However, in other parts of the world, including Japan, Taiwan, and Brazil, P americana is common, and there is high prevalence of IgE antibody to P americana (60%) among patients with asthma from those countries.\textsuperscript{20-22}

The aim of this study was to identify cross-reactive allergens from B germanica and P americana by screening a P americana cDNA library with IgE antibodies from patients allergic to cockroaches living in the United States, who were exposed to B germanica in their homes. Using this approach, we obtained the nucleotide and deduced amino acid sequence of Per a 1 and show that it is homologous to other previously unidentified cockroach allergen sequences and to a mosquito precursor protein (ANG12), which may be involved in digestion.
**IgE antibody**

<table>
<thead>
<tr>
<th>αBla g 1</th>
<th>αBla g 2</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>mAb 10A6</td>
<td>αBla g 1</td>
<td>mAb 8F4</td>
</tr>
<tr>
<td>α Bla g 2</td>
<td>α Bla g 2</td>
<td></td>
</tr>
</tbody>
</table>

**METHODS**

**Sera from patients allergic to cockroaches**

Twenty sera were obtained from patients with asthma who either were examined at the University of Virginia Allergy Clinic or had been enrolled in previous emergency department studies in Wilmington, Del, or Atlanta, Ga. The sera contained high levels of IgE antibody to *B germanica* and *P americana* by RAST (500 to 6000 U/mL; 1 U = approximately 0.1 ng of IgE antibody) and also contained known levels of IgE antibody to *Bla g 1* (4 to 10,000 U/mL) or *Bla g 2* (4 to 1000 units/mL). Control sera were obtained from 3 individuals who were not allergic to cockroaches or other allergens. Collection of sera for use in these studies was approved by the University of Virginia Human Subjects Investigation Committee.

**cDNA library screening**

A *P americana* cDNA expression library was prepared in the UniZAP XR expression vector (Stratagene, La Jolla, Calif) with 10 µg of mRNA isolated from adult cockroaches as previously described. The library was screened with a 1:2 dilution of serum (preabsorbed with *Escherichia coli* lysate) pooled from 5 patients who were sensitive to *B germanica*. Positive plaques were detected with alkaline phosphatase–labeled goat anti-human IgG (1:500). A positive cDNA clone (PA2) was identified and further screened by plaque immunoassay with 23 individual sera. Ten sera were obtained from patients having detectable IgE antibody to *Bla g 1* by RIA, 10 patients had positive RAST results to cockroaches but had no IgE antibody to *Bla g 2*, and 3 sera were from nonallergic control subjects. Clone PA2 was also screened against a panel of anti-cockroach mAbs, including anti-Bla g 1 mAb 10A6, anti-Bla g 2 mAb 8F4, and 7 anti-Per a 1 mAbs, designated Per a 1 (02 to 08; provided by Drs. H. Hemmar and C. Schou, ALK-ABELLO, Horsholm, Denmark). Two mAbs from a *Blomia tropicalis* fusion (5A5 and 5A7) served as negative controls. The mAbs were used at 1 to 20 µg/mL for screening; alkaline phosphatase–labeled goat anti-mouse IgG (1:1000 or 1:5000, Kirkaard & Perry Laboratories, Inc. Gaithersburg, Md) was used as secondary antibody for detection.

**In vivo excision, insert size measurement, and sequencing**

The p-Bluescript phagemid containing the cloned cDNA insert was excised from the LambdaZap II vector and plated with fresh *Escherichia coli* cells to produce colonies (ExAssist/SOLR System; Stratagene). Double-stranded cDNA was isolated and sequenced with the use of an ABI Prism 100 automated DNA sequencer (Model 377; Biomedical Research Facility, University of Virginia).

**Sequence analysis**

Protein and nucleotide sequences were compared with nonredundant GenBank, CDS translations, EMBL, DDBJ, DDBJ, PIR, SUpdate and SwissProt data bases, using Fasta (GCG, Oxford Molecular Group, Inc). Sequence alignments were made with the GCG program.

**Expression of Per a 1 in Pichia pastoris**

Clone PA2 cDNA was used as a template to generate a 731-bp PCR product containing 5′ Sna B1 and 3′ Not 1 restriction enzyme sites. The enzyme-digested PCR product was purified and ligated into the Sna B1 and Not 1 sites of the PIC9 *Pichia* expression vector as described previously. The pPIC9/PA2 construct was (20 µg) linearized and transformed by electroporation into the yeast genome of *Pichia* strains KM71 and GS115. Yeast clones incorporating PA2 DNA were identified by PCR and expressed in buffered minimal methanol media. Expression culture supernate was dialyzed against 50 mmol/L phosphate buffer containing 2 mol/L NaCl, pH 7.0, and passed over a phenyl sepharose CL-4B column. Protein was eluted stepwise in 0.5 mol/L NaCl buffer containing no salt or water only, and fractions were assayed for Per a 1 by ELISA. Recombinant Per a 1 (rPer a 1) was further purified over a Mono Q HR5/5 column (Pharmacia Biotech, Piscataway, NJ) with 20 mmol/L n-methyl piperazine, pH 4.5, in a 1 mol/L NaCl gradient over 20 minutes. The rPer a 1 eluting at 60% to 80% NaCl was dialyzed against PBS, concentrated, and assayed for purity by silver-stained SDS-PAGE (PhastSystem; Pharmacia Biotech).

**RESULTS**

**Identification of cDNA clones encoding Per a 1**

The aim of these studies was to identify cross-reactive allergens produced by *B germanica* and *P americana*, by screening a *P americana* cDNA with IgE antibodies in a serum pool from patients allergic to cockroaches living in the United States, who were primarily sensitized to *B germanica*. A cDNA clone (PA2) was obtained and rescreened by plaque immunoassay against a panel of 10 individual sera from patients allergic to cockroaches, with known levels of IgE antibody to *Bla g 1* or *Bla g 2*. The results showed that 7 of 10 sera from patients with IgE antibodies to *Bla g 1* reacted strongly with the cDNA clone, whereas sera from 10 patients with IgE antibodies to *Bla g 2*, but not to *Bla g 1*, showed no reactivity (Fig 1). Similarly, an anti-*Bla g 1* mAb (10A6) bound strongly to allergen expressed by the clone, and anti-*Bla g 2* mAb (8F4) gave negative results (Fig 1). These results suggested that the allergen encoded by the cDNA clone...
Homology between Per a 1 sequences and mosquito protein (ANG12)*

<table>
<thead>
<tr>
<th></th>
<th>Per a 1.01</th>
<th>Cr-PII</th>
<th>ANG12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>EFQSIIS</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>277</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Sequence comparisons are for Per a 1.01 (GenBank Accession No. AF72222), Cr-PII (L69261) and ANG12 (Z29925) (25).

**FIG 3.** Alignment between Per a 1, Cr-PII, and mosquito protein ANG12. Amino acids in bold indicate homology with ANG12 (overall, 31%). The 14 amino acids that differ between Per a 1 and Cr-PII are indicated by asterisks.
was Per a 1, the homologous allergen to Bla g 1 produced by *P americana*.9 The findings were confirmed by plaque immunoassay with a panel of 7 anti-Per a 1 mAbs provided by Helene Henmar and Dr Carsten Schou. Strong positive binding was obtained with 5 of 7 anti-Per a 1 mAbs, as compared with anti-Bla g 2 mAb and 2 mAbs raised against mite, *Blomia tropicalis*, used as negative controls (Fig 2).

**Nucleotide and deduced amino acid sequence of Per a 1**

The Per a 1 cDNA clone was fully sequenced and contained an 870 bp and a 695-bp open reading frame, encoding a 231 amino acid protein with a molecular weight of 26.2 kd. This clone, Genbank Accession Number AF72222, was designated as Per a 1.01 in accordance with the WHO/IUIS allergen nomenclature.24 The Per a 1.01 cDNA had a stop codon, TGA, at position 696, followed by a polyA signal and a polyA tail further downstream.

**Homology between Per a 1, other cockroach allergen sequences, and mosquito ANG12 gene**

When the nucleotide sequence of Per a 1.01 was compared with other sequences in the GenBank, we were surprised to find extensive homology to a previously reported German cockroach allergen sequence, termed Bla g Bd90K.15 The cDNA encoding Bla g Bd90K was 4058 bp long and contained 7 highly homologous (approximately 90%) tandem repeats of approximately 576 bp. The nucleotide sequence of 1 of these repeats was published previously, and no significant homology had been found between the partial Bla g Bd90K sequence and other proteins.15 The Per a 1.01 sequence showed 70% homology to Bla g Bd90K, and homology was found in all 3 reading frames of the Bla g Bd90K translation.

Additional database searches revealed that Per a 1.01 was almost identical to 4 *P americana* allergen sequences recently reported by Wu and colleagues:25 Cr-PII; Cr-PII2; Cr-PII3; and Cr-PII4 (Fig 3). For example, Per a 1.01 differed from Cr-PII clones by only 14 amino acids, indicating that the Cr-PII clones are variants of Per a 1 (Fig 3). The Cr-PII clones encoded proteins of 274, 446, 395, and 228 amino acids, with estimated molecular weights of 31, 51, 45, and 26 kd, respectively, and had no cysteine residues or potential N-glycosylation sites.

The Per a 1 (and Cr-PII) sequences also showed homology to a mosquito (*Anopheles gambiae*) precursor protein, named ANG12. The homology with ANG12 reaches over 40% in parts (for example, the GVDVDH motif is completely preserved) and the overall homology is 31% (Fig 3).

**Expression of Per a 1 in Pichia pastoris**

Three of 15 clones tested by PCR had Per a 1 DNA incorporated into the *Pichia pastoris* strain, and 1 was a GS115 His*<sup>+</sup> Mut*<sup>+</sup> phenotype. All clones secreted rPer a 1 into the medium when grown in shake flasks under methanol and yielded 14 mg rPer a 1 per liter of culture after partial purification over phenyl sepharose. Further purification by anion exchange liquid chromatography resulted in a single nonglycosylated 20-kd band on silver-stained SDS-PAGE (Fig 4). The *Pichia* expressed rPer a 1 showed reactivity in the mAb based ELISA for Bla g 1 (data not shown).

**DISCUSSION**

Natural Bla g 1 and Per a 1 were previously identified with immunochemical techniques and shown to elicit IgE responses in 30% to 50% of patients allergic to cockroaches by skin testing and in vitro tests.9,10 These allergens are important because they are the only cockroach allergens discovered to date that are produced by both *B germanica* and *P americana*. The results clearly show that we have cloned Per a 1, based on immunologic identification of the protein expressed by the cDNA clone with mAbs and human IgE antibodies. This protein reacted strongly with anti-Per a 1 and anti-Bla g 1 mAbs and with IgE antibodies in sera from patients with known levels of IgE antibodies to Bla g 1, but not with control mAbs, or sera containing IgE antibodies to Bla g 2, but not Bla g 1. We recently screened a *B germanica* cDNA library with anti-Bla g 1 mAb (10A6) and isolated 3 cDNA clones encoding Bla g 1, which showed 70% to 72% sequence identity to Per a 1.26 Together, these results demonstrate that Per a 1 and Bla g 1 constitute a family of structurally and antigenically related “Group 1” allergens produced by *B germanica* and *P americana*.

**FIG 4.** SDS-PAGE analysis of recombinant Per a 1 expressed in *Pichia pastoris*. Lane 1, Molecular weight markers; lane 2, culture supernate from a rPer a 1 *Pichia pastoris* clone 3-1#3; lane 3, rPer a 1 purified from culture supernate which was passed over a Phenyl Sepharose column and eluted with water; lane 4, Phenyl Sepharose rPer a 1 further purified over a MonoQ H5/5 HPLC column.
The identification of the Per a 1 and Bla g 1 cDNA clones was dependent on the use of mAbs with defined specificities. A striking feature of the sequence homology searches was that Per a 1 and Bla g 1 showed extensive identity (60% to 95%) to the previously reported Cr-PII sequences and to Bla g Bd90K. Wu et al25 showed that these sequences were related and raised mAbs that bound to the expressed recombinant Cr-PII allergens. However, the specificity of their mAb was not known, and the relationship of the sequences to other cockroach allergens could not be established. Taking our data in conjunction with that of Wu et al25 and Helm et al,15 it is clear that the Cr-PII sequences and Bla g Bd90K belong to the same family of Group 1 allergens. These allergens occur in multiple molecular forms with different molecular weights, which are the result of replication of approximately 100 amino acid repeats in the sequences. A detailed analysis of the Bla g 1 repeats and of the homology between these sequences has recently been published.26 A comparison of the cDNAs, their predicted protein size, GenBank accession numbers, and WHO/IUIS allergen nomenclature is shown in Table I. The nomenclature was developed by considering 1 cDNA as the reference clone. If 2 clones shared 67% or greater amino acid sequence identity, they were considered isoallergens, whereas if the amino acid sequence identity was higher than 93%, the clones were considered variants (isoforms) of the same isoallergen. The Bla g 1.01, Bla g 1.02, and Per a 1.01 allergens were designated as such because they were the first clones that were definitively identified. Of the 3 Bla g 1 cDNA clones, 2 were 100% identical and were designated Bla g 1.0101 (Table I). The third clone showed 75% amino acid identity to Bla g 1.0101 and was designated Bla g 1.02. The Bla g 1 clones showed homology with Bla g Bd90K.15 Because this nucleotide sequence was nontranslatable, it was designated Bla g 1.0102 based on extensive (93%) nucleotide homology with Bla g 1.0101.

The Per a 1 sequences were named based on their homologic features to the reference clone Per a 1.01 (AF72222). Three of the 4 Cr-PII allergens reported by Wu et al25 showed more than 93% amino acid identity to Per a 1.01 and were designated Per a 1.0102, Per a 1.0103, and Per a 1.0104 (Table I). The other Cr-PII clone was an isoallergen showing 78.2% identity with Per a 1.01; this clone was designated Per a 1.02.

A panel of recombinant P americana allergens is now available for clinical and immunologic studies. Recombinant Per a 1 has been expressed in Pichia pastoris and in the bacterial pET system (Cr-PII clones). The immunologic reactivity of these recombinant allergens needs to be assessed in a multicenter study, including other cloned P americana allergens, such as Per a 3. Tropomyosin has also recently been identified as an important P americana allergen and shows a high degree of sequence homology with dust mite and shrimp tropomyosins. With this panel of recombinant P americana allergens, it will be possible to develop improved reagents for diagnostic use and for incorporation into allergen-specific treatment strategies. These sequences should facilitate further investigation of the role of P americana in causing asthma.

We thank Helene Henmar and Dr Carsten Schou of ALK-ABELLO for providing anti-Per a 1 mAb and samples of natural Per a 1; Glaxo-Wellcome for a travel grant to enable EM to present the data at the AAAAI Annual Meeting in March 1998; and Wanda Harvey for secretarial assistance.

REFERENCES


