Cloning of Cockroach Allergen, Bla g 4, Identifies Ligand Binding Proteins (or Calycins) as a Cause of IgE Antibody Responses*

L. Karla Arruda‡, Lisa D. Vailes‡, Mary L. Hayden‡, David C. Benjamin§, and Martin D. Chapman‡¶

From the Asthma and Allergic Diseases Center, Departments of ‡Internal Medicine and §Microbiology, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908

An allergen cloned from a *Blattella germanica* (German cockroach) cDNA library, encoded a 182-amino acid protein of 20,904 Da. This protein, designated *B. germanica* allergen 4 (Bla g 4), was expressed as a glutathione *S*-transferase fusion protein in *Escherichia coli* and purified by affinity chromatography and high-performance liquid chromatography. The prevalence of serum IgE antibody to recombinant Bla g 4 in 73 cockroach allergic patients with asthma ranged from 40% (antigen binding radioimmunoassay) to 60% (plaque immunoassay). Cockroach allergic patients gave positive intradermal skin tests to recombinant Bla g 4 at concentrations of 10^{-3}–10^{-5} μg/ml, whereas non-allergic controls, or cockroach allergic patients with no detectable serum IgE antibody to Bla g 4, gave negative skin tests to 1 μg/ml. Polymerase chain reaction and Southern analysis identified a 523-base pair DNA encoding Bla g 4 in both *B. germanica* and *Periplaneta americana* (American cockroach). However, Northern analysis showed that mRNA encoding Bla g 4 was transcribed in *B. germanica* but not in *P. americana*, suggesting that allergen expression was species specific. Sequence similarity searches showed that Bla g 4 was a ligand binding protein or calycin and unexpectedly revealed that this family contained several important allergens: β-lactoglobulin, from cow milk, and rat and mouse urinary proteins. Although the overall sequence homology between these proteins was low (~20%), macromolecular modeling techniques were used to generate two models of the tertiary structure of Bla g 4, based on comparisons with the x-ray crystal coordinates of bilin binding protein and rodent urinary proteins. The results show that members of the calycin protein family can cause IgE antibody responses by inhalation or ingestion and are associated with asthma and food hypersensitivity.

Inhalation of environmental allergens, derived from pollens, dust mites, animal danders, insects, and fungi, is the most common cause of IgE antibody responses in humans. Cockroaches (CRs)[1] secrete potent allergens that induce IgE antibody responses and subsequent development of asthma in atopic individuals who are chronically exposed to these allergens. Indeed, although CRs are known to harbor viral and bacterial pathogens, the only specific disease associated with CR-infested housing is bronchial asthma. Sensitization occurs through inhalation of allergens either in CR-infested houses or following occupational exposure, *e.g.* among entomologists and laboratory personnel who work with CR (1–4). In urban or "inner city" areas, up to 60% of patients with asthma have IgE antibodies to CR allergens, as demonstrated by skin tests, bronchial challenge tests, and serum IgE antibody assays (1–4). Delayed bronchial reactions involving eosinophil recruitment, a characteristic immunopathologic feature of asthma, also occur following inhalation of CR extract (5). Asthma mortality and morbidity is increasing in the United States, and recent epidemiologic studies have shown that IgE-mediated sensitivity to CR allergens is a major risk factor for emergency room admission with asthma (6–8).

The principal domiciliary CR species are *Blattella germanica* (German cockroach) and *Periplaneta americana* (American cockroach). The molecular nature of the allergens produced by either species is poorly understood. Moreover, despite the extensive use of CR in biological research, there have been few fundamental studies of CR proteins. Several low molecular mass (10–70 kDa) protein bands have been identified as allergens using IgE antibodies, and two allergens, Bla g 1 and Bla g 2, have been defined using protein purification techniques and monoclonal antibodies (2, 9–13). We have used molecular cloning techniques to identify and sequence cDNA clones encoding *B. germanica* allergens. Here, we report that a 21-kDa CR allergen, Bla g 4, is a ligand binding protein, or calycin, and show that this family contains several allergens that are a common cause of IgE antibody responses, including mouse and rat urinary proteins, and β-lactoglobulin. Recombinant Bla g 4 expressed in *Escherichia coli* was strongly reactive on skin testing and in serum IgE antibody assays, suggesting that this recombinant allergen will be useful for diagnostic and therapeutic purposes.

EXPERIMENTAL PROCEDURES

cDNA Cloning—Total RNA was isolated from 6-g adult *B. germanica*, of mixed sexes, using 5 M guanidinium isothiocyanate, 0.01 M EDTA, 5% 2-mercaptoethanol, 0.05 M Tris-HCl, pH 7.5 (14). Messenger RNA was isolated using a FastTrack kit (Invitrogen). *B. germanica* cDNA library was prepared from 10 μg of mRNA in the UniZAP-XR phagemid expression vector (Stratagene). The *B. germanica* cDNA library was plated on NZY agar and screened using IgE antibodies.

[1] The abbreviations used are: CR, cockroach; PCR, polymerase chain reaction; HPLC, high performance liquid chromatography; rBla g 4, recombinant Bla g 4; ab, antibody; bp, base pair(s); PAGE, polyacrylamide gel electrophoresis; SCR, structurally conserved regions.

Protein expression was induced for 3 h using nitrocellulose filters soaked in 10 mM isopropyl-1-thio-β-D-galactopyranoside. Filters were incubated with 1% dried milk, 0.2% bovine serum albumin, 0.4% goat serum, 0.03% gelatin for 1 h, followed by a 1:2 dilution of IgE serum pool from eight CR-allergic patients (which had been pre-absorbed with *E. coli* lysate: 4 mg of lysate/ml of cyanogen bromide-activated Sepharose). IgE antibody was detected using alkaline phosphatase-labeled anti-IgE and 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium substrate (KPL, Gaithersburg, MD). Six positive plaques were identified after screening ~150,000 plaques. Selected cDNA clones were screened against individual sera from 20 CR-allergic patients or 4 non-allergic controls by plaque immunoassay. Clone *bg*12A was screened against a further 53 sera from CR-allergic patients. Double-stranded sequencing of cDNA clone *bg*12A was carried out by dideoxynucleotide chain termination using a Sequenase kit (U. S. Biochemical Corp.) (15). This sequence was designated Bla g 4 in accordance with allergen nomenclature (16).

Expression of Recombinant Bla g 4 in E. coli—Bla g 4 plasmid DNA (50 ng) was used as template to generate a 546-bp PCR product containing *Bam*HI and *Xho*I restriction enzyme sites to allow unidirectional subcloning into the pGEX-4T1 expression vector (Pharmacia Biotech Inc.). Primers for PCR were synthesized as follows: 5′-CGCG-GATCCACAGATACATTGGCGAA-3′ (sense) and 5′-CCGCTCGAGT-TAGTGACATGTGGAGTG-3′ (antisense). PCR incubations were 1 min at 94 °C, 1 min at 37 °C, and 3 min at 72 °C for 30 cycles in a 50-μl volume. An initial 5-min incubation step at 95 °C was performed, and each reaction was terminated for 15 min at 72 °C. The 546-bp PCR-amplified DNA was ligated into *Bam*HI-*Xho*I-digested pGEX-4T1. DNA ligation and transformation of competent *E. coli* strain TOP10F′ (Invitrogen) was as described (17). Expression of Bla g 4 as a fusion protein with glutathione *S*-transferase was induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside, and recombinant protein was purified from cell lysates by chromatography over glutathione-agarose (18). Digestion with thrombin (10 units/mg of protein for 18 h at room temperature) released the 21-kDa Bla g 4 protein, which was recovered in the the flow-through following further purification over glutathione-agarose. To assess purity, rBla g 4 was analyzed by silver-stained SDS-PAGE and by size exclusion HPLC over a Superdex 75 HR 10/20 column (Pharmacia). Recombinant Bla g 4 eluted as a single HPLC peak, and the amino acid sequence of the 5 NH$_2$-terminal residues was confirmed by Edman degradation. The final yield was 250 μg of purified rBla g 4 per liter of culture.

Immunoassay for IgE Antibodies to rBla g 4—IgE anti-Bla g 4 ab were measured in sera from 73 CR-allergic asthmatic patients, using an antigen-binding radioimmunoassay (19). Briefly, 9 μg of rBla g 4 was radiolabeled with 0.5 mCi of ^{125}I, using the chloramine-T technique (specific activity, 3 μCi/μg). Serum dilutions of 1:2 and 1:10 were incubated with ^{125}I-rBla g 4 (~100,000 cpm added) for 4 h at room temperature and precipitated overnight at 4 °C with 50 μl of sheep anti-human IgE (The Binding Site, San Diego, CA). IgE myeloma serum (patient P.S.) diluted 1:200 was used as carrier. Precipitates were washed with BBS and counted in a γ-counter. The assay was quantitated using a control curve, constructed with patient SW serum, assigned to contain 10,000 units/ml IgE antibody.

Immediate Skin Testing—Quantitative intradermal skin tests were performed using serial 10-fold dilutions of *B. germanica* extract (1/20 w/v, Allergy Laboratories of Ohio, Columbus, OH) or purified recombinant Bla g 4 from 10^0–10^{-5} μg/ml, as described previously (12, 19). A wheal of >8 × 8 mm in diameter 15–20 min after injection was regarded as a positive skin test. Immediate skin testing using rBla g 4 was approved by the Human Investigation Committee of the University of Virginia.

Human Sera—Sera from CR-allergic patients had been collected from patients enrolled in previous studies on emergency room asthma. All patients had asthma and had serum IgE antibodies to CR detectable by radioallergosorbent test (values 40–6,000 units/ml, 1 unit = ~0.1 ng of IgE) (7, 8). Collection of sera for use in these studies was approved by the Human Investigation Committee of the University of Virginia.

Sequence Analyses and Molecular Modeling—Protein sequences were compared using the FASTA program to search the NBRF, GenBank, and Swiss-Prot sequence data bases (20). Homology-based molecular modeling was carried out using a combination of knowledge-based and *ab initio* methods (21). The SYBYL modeling package (Tripos Inc., St. Louis, MO) was used to determine homology with known structures from the Brookhaven Protein Data base and for loop searches, mutations, and graphic visualization. The CHARMM-based program CONGEN was used for energy minimizations, simulated annealing, molecular dynamics, conformational searches, and free energy
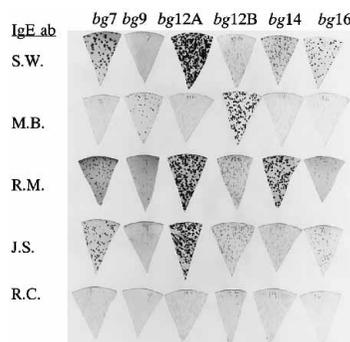
FIG. 1. **Human IgE antibody binding to *B. germanica* cDNA clones.** Six clones were obtained after screening a *B. germanica* cDNA library with pooled IgE antibodies, and each clone was screened against individual sera from 20 CR-allergic patients or 4 non-allergic controls by plaque immunoassay. Selected results on four CR-allergic patients (SW, MB, RM, JS) and one control (RC) are shown. The complete nucleotide sequences of clones *bg*12A, *bg*12B, and *bg*16 have been determined, and partial sequences have been obtained for the other clones. Sequence data on clones other than *bg*12A will be published elsewhere.

calculations (22, 23).

Northern Analysis of CR mRNA—*B. germanica* or *P. americana* mRNA was obtained from total RNA using the Poly(A) Tract mRNA isolation system (Promega). Samples containing 2 μg of *B. germanica* or *P. americana* mRNA were electrophoresed in formaldehyde-denaturing 1% agarose gel, and RNA was transferred to a nylon membrane (Zetabind, Cuno, Meridien, CT). A 650-bp *Sma*I-*Kpn*I-digested pBLUE-SCRIPT plasmid DNA fragment containing Bla g 4 cDNA and an 8.6-kilobase *Neurospora crassa* DNA (pRW528), coding for highly conserved ribosomal sequences, were labeled with [α-^{32}P]dCTP by random priming. Hybridization was performed at 37 °C, as described previously (24).

PCR and Southern Analysis—*B. germanica* or *P. americana* genomic DNA was extracted from 0.1 g ground cockroach tissue using a blood and cell culture DNA kit (QIAGEN, Chatsworth, CA). 100 ng of CR genomic DNA was amplified by PCR using *Taq* polymerase (GeneAmp kit; Perkin Elmer Corp.). The following oligonucleotide primers, derived from the nucleotide sequence of Bla g 4 DNA, were used to amplify a 523-bp fragment: 5′-ACAGATACATTGGCGAA-3′ (sense) and 5′-GA-CATGTGGAGTGTAAG-3′ (antisense). PCR conditions were as described above using an annealing temperature of 42 °C. PCR products were electrophoresed in 1% agarose gel, and DNA was transferred to nylon membrane. The Southern blot was hybridized with [α-^{32}P]dCTP-labeled 650-bp Bla g 4 cDNA probe at 37 °C and autoradiographed following a 3-h exposure to Kodak XAR film.

RESULTS

Molecular Cloning of B. germanica Allergens—A unidirectional *B. germanica* cDNA library was screened using pooled IgE antibodies from CR-allergic patients with asthma. Six positive plaques were cloned and rescreened against a panel of sera from CR-allergic patients by plaque immunoassay. Selected results on four CR-allergic patients and one control are shown in Fig. 1. Most patients had IgE antibody to two or more clones (*e.g.* SW, *bg*7, *bg*12A; RM, *bg*7, *bg*12A, *bg*14) and showed different patterns of IgE antibody binding, suggesting that *B. germanica* produced multiple allergens. The strongest intensity of IgE antibody binding was observed using protein encoded by clone *bg*12A, and ~60% (47/73) of sera from CR-allergic patients gave positive IgE antibody plaques to this protein. Nucleotide sequencing showed that *bg*12A cDNA contained a 546-bp open reading frame, coding for a 182-amino acid protein with an estimated molecular mass of 20,904 daltons (Fig. 2). The allergen encoded by clone *bg*12 A was provisionally designated *B. germanica* allergen 4, Bla g 4, in keeping with the revised WHO/IUIS allergen nomenclature (16). Sequence data on the other CR allergen clones isolated from the *B. germanica* cDNA library will be reported elsewhere.

Demonstration of IgE Antibody Responses to rBla g 4—PCR-

```
  1: C  GCA GTT TTG GCA CTA TGT GCA ACA GAT ACA TTG GCG AAC GAA
        A   V   L   A   L   C   A   T   D   T   L   A   N   E

 44:    GAT TGT TTT AGA CAT GAA TCA TTG GTT CCA AAC CTT GAT TAT
        D   C   F   R   H   E   S   L   V   P   N   L   D   Y

 86:    GAA AGG TTC AGA GGT TCG TGG ATT ATT GCA GCC GGC ACT TCC
        E   R   F   R   G   S   W   I   I   A   A   G   T   S

128:    GAA GCG CTC ACC CAA TAC AAA TGC TGG ATC GAC AGG TTT TCA
        E   A   L   T   Q   Y   K   C   W   I   D   R   F   S

170:    TAT GAC GAT GCG TTG GTT TCT AAG TAT ACT GAT TCA CAA GGA
        Y   D   D   A   L   V   S   K   Y   T   D   S   Q   G
                •   •       •
212:    AAG AAT AGG ACT ACT ATC AGA GGA CGA ACT AAA TTT GAA GGC
        K   N   R   T   T   I   R   G   R   T   K   F   E   G

254:    AAC AAG TTT ACT ATC GAT TAT AAT GAT AAA GGG AAA GCA TTT
        N   K   F   T   I   D   Y   N   D   K   G   K   A   F

296:    TCT GCG CCA TAC TCT GTT CTA GCA ACT GAT TAC GAA AAT TAC
        S   A   P   Y   S   V   L   A   T   D   Y   E   N   Y

338:    GCA ATT GTG GAA GGC TGT CCC GCT GCA GCT AAT GGA CAT GTA
        A   I   V   E   G   C   P   A   A   A   N   G   H   V

380:    ATT TAT GTT CAA ATC CGA TTT TCG GTG AGG AGA TTT CAC CCC
        I   Y   V   Q   I   R   F   S   V   R   R   F   H   P

422:    AAG CTG GGT GAT AAA GAA ATG ATA CAG CAC TAC ACT TTG GAT
        K   L   G   D   K   E   M   I   Q   H   Y   T   L   D

464:    CAG GTG AAT CAA CAC AAG AAG GCT ATA GAA GAA GAC TTA AAG
        Q   V   N   Q   H   K   K   A   I   E   E   D   L   K

506:    CAC TTC AAT TTG AAG TAC GAG GAC TTA CAC TCC ACA TGT CAC
        H   F   N   L   K   Y   E   D   L   H   S   T   C   H

548:    TAA GTATGAAATGTTCATATTTATTGTAGGAAAATAAAACCTTCTAATGAATTA₂₉
        ***
```

FIG. 2. **Nucleotide and deduced amino acid sequence of *B. germanica* clone *bg*12A.** The initiation methionine is lacking from the deduced sequence. The N-terminal amino acid residues (1–12) contain the conserved features of a signal peptide, with a predicted cleavage site located between Ala[12] and Asn[13] (75–80% predictive accuracy) (25). The estimated molecular mass of the protein encoded by clone *bg*12A without the signal peptide sequence is 20,126 Da. One potential *N*-linked glycosylation site is indicated (●). The stop codon TAA is shown (*), and a polyadenylation signal sequence, AATAAA, in the 3′-non-coding region, is *underlined*.

amplified DNA encoding Bla g 4 was ligated into pGEX-4T1 and expressed as a glutathione *S*-transferase fusion protein in *E. coli*. Recombinant Bla g 4 was obtained from bacterial lysates by glutathione affinity chromatography and thrombin cleavage, and the pure protein migrated as a single band of 18 kDa on an 8–24% gradient SDS-PAGE (Fig. 3). Serum IgE ab to rBla g 4 was compared in 73 sera from CR-allergic patients by antigen binding radioimmunoassay. The prevalence of IgE ab was 41% among patients with a CR radioallergosorbent test value of >200 units/ml and 31% among patients with a CR radioallergosorbent test value of 40–200 units/ml (Fig. 4). This prevalence of reactivity was lower than that observed by plaque immunoassay (~60%) and may possibly be explained by increased sensitivity of the plaque assay, as compared to radioimmunoassay. The ¹²⁵I-rBla g 4 showed strong reactivity with IgE ab (up to 45,000 cpm bound, as compared to controls of <400 cpm), suggesting that the recombinant protein expressed the majority of B cell epitopes (Table I).

The biologic activity of rBla g 4 was assessed by quantitative intradermal skin testing of seven selected CR-allergic patients and three non-allergic controls, as reported in other studies. The results show that positive skin tests were obtained using 10^{-3}–10^{-5} μg/ml rBla g 4 and that skin test reactivity broadly correlated with serum IgE ab responses. In contrast, neither non-allergic controls nor CR-allergic patients with no detectable serum IgE ab to rBla g 4 gave positive skin tests using up to 1 μg/ml rBla g 4 (Table I). These results showed that rBla g 4 was capable of inducing specific immediate hypersensitivity responses in CR-allergic patients.

Homology between Bla g 4 and Calycins—Sequence similarity searches showed that Bla g 4 was a member of the calycin family of proteins (Fig. 5). Calycins are a diverse family of ~30 proteins, which include lipocalins and fatty acid binding pro-
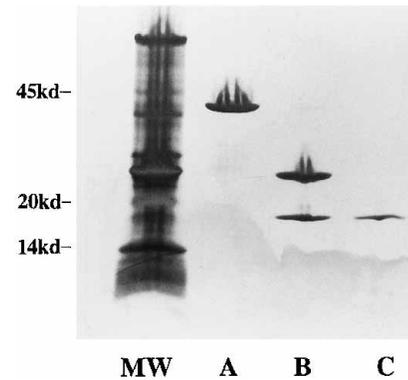
FIG. 3. **SDS-PAGE analysis of recombinant Bla g 4, expressed in *E. coli* as a glutathione *S*-transferase fusion protein.** Glutathione *S*-transferase-Bla g 4 was purified over glutathione-agarose and eluted with 10 mM reduced glutathione (*A*), followed by thrombin cleavage (*B*). Purified rBla g 4 was recovered in the flow-through after further purification over glutathione-agarose (*C*). Proteins were analyzed in a silver-stained 8–25% gradient SDS-PAGE gel using a PhastSystem. Molecular weight markers (*MW*) are indicated.
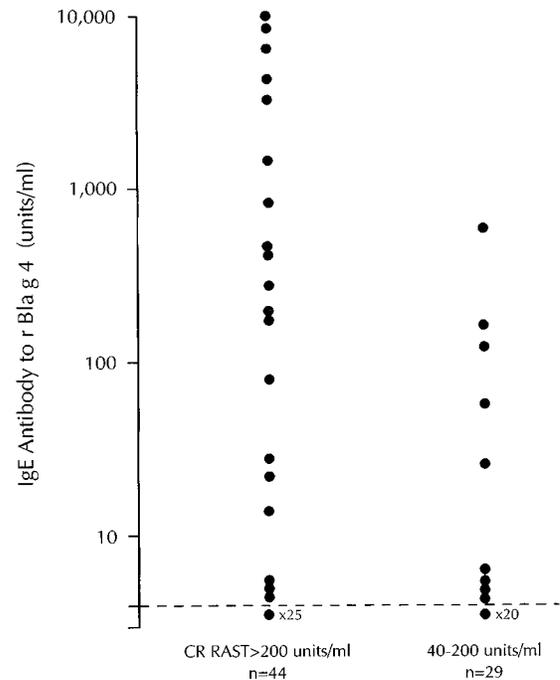
FIG. 4. **Serum IgE antibodies to rBla g 4 measured in sera from 73 CR-allergic asthmatic patients by antigen binding radioimmunoassay.** Results are expressed in units of IgE binding activity relative to a control serum (SW), which was arbitrarily assigned a value of 10,000 units/ml. The lower level of assay sensitivity is indicated (*dashed line*).

teins, whose function is to bind or transport small hydrophobic molecules. Examples include human retinol binding protein, butterfly bilin binding protein (BBP) and tobacco hornworm insecticyanin (pigment binding protein) (26–28). Calycins were not previously known to cause IgE responses, but sequence analyses unexpectedly revealed that this protein family also contained three major allergens: β-lactoglobulin from cow milk and rodent urinary proteins (mouse urinary protein, MUP, and rat $α_{2u}$-globulin). The overall homology between Bla g 4 and calycins was 18.9–23.9%, consistent with the low degree of sequence homology between other members of the family. Assignment of Bla g 4 to this family was based on the presence of two concensus motifs, originally described by Sawyer (26) and Li and Riddiford (27) (Fig. 5). Subsequent comparisons showed

TABLE I

Skin tests and serum IgE antibodies to recombinant B. germanica allergen Bla g 4

Patient	Skin test to *B. germanica*	rBla g 4		
		CR radioallergosorbent test	Skin test	Serum IgE ab[a]
		units/ml	*μg/ml*	*cpm bound*
JS	10^{-5}	1,395	10^{-3}	45,043
SW	10^{-5}	1,290	10^{-5}	42,628
TB	10^{-5}	450	10^{-5}	38,579
RM	10^{-5}	1,065	10^{-3}	24,488
MB	10^{-5}	1,465	>1	374
ST	10^{-5}	1,080	>1	248
CT	Prick	125	>1	257
Controls ($n = 3$)[b]	ND	<10	>1	300 ± 33

[a] Measured by antigen-binding radioimmunoassay.
[b] Mean ± 2 S.D. of three non-allergic subjects.

```
              17- - - - - -32    102- - - - - - - 119   129 - - - 139
              oo*    o * *        oo *** *o            o    *
Bla g 4       NLDYERFRGGSWIIAAG   PYSVLATDYENYAIVEGC     VQIRFSVRRFH
Insecticyanin DFDLSAFAGAWHEIAK    VPWVLATDYKNYAINYNC     SIHAWILSKSK
Bilin-BP      NFDWSNYHGKWWEVAK    VFNVLSTDNKNYIIGYYC     QDFVWVLSRSK
Apolipoprotein D NFDVNKYLGRWYEIEK PYWILATDYENYALVYSC     VDFAWILARNP
Retinol-BP    NFDKARFSGTWYAMAK    DHWIVDTDYDTYAVQYSC     DSYSFVFSRDP
α₂ᵤ-globulin  NLDVAKLNGDWFSIVV    TFTILKTDYDRYVMFHLI     FQLMVLYGRTK
MUP           NFNVEKINGEWHTIIL    TFTIPKTDYDNFLMAHLI     FQLMGLYGREP
ß-Lactoglobulin GLDIQKVAGTWYSLAM  KVLVLDTDYKKYLLFCME     SLACQCLVRTP

              ----SCR 1------     -------SCR 2------    ---SCR 3---
```

FIG. 5. **Amino acid sequence homology between *B. germanica* Bla g 4 and selected calycins.** Conserved residues in the Sawyer (26) and Li and Riddiford (27) motifs are located at positions 17–32 and 102–118 of insecticyanin and are indicated in *bold type* (*). Conservative substitutions of acidic or hydrophobic residues, which also form part of the motifs, are indicated (○). Also shown are the three structurally conserved regions (SCR1–3) defined using iterative comparisons of Bilin-BP, Insecticyanin, Retinol-BP, and MUP (29). The two sequence motifs broadly correspond to SCR1 and SCR2, respectively. SCR3 has a conserved arginine or lysine residue (at position 137 in insecticyanin), which is an important contact residue.

that the Bla g 4 sequence contained each of the three structurally conserved regions (SCR) of calycins proposed by Flower *et al.* (29).

The molecular structures of BBP, insecticyanin, MUP, and rat $α_{2u}$-globulin had previously been determined at high resolution by x-ray crystallography. The characteristic calycin structure is a cup-shaped, eight-strand, anti-parallel β-barrel, with a +1 topology (26–32). Structural models of Bla g 4 were constructed from the x-ray coordinates of butterfly BBP (which was selected as the base molecule because the positions of cysteine residues involved in disulfide bonds were conserved in the two proteins). Although the overall sequence identity between BBP and Bla g 4 is but 20%, it is comparable to that between BBP and the rodent urinary proteins, which have very similar conformations (29). The amino acid structures of Bla g 4 were substituted for those of BBP at positions where the two proteins differed. The positions of the main chain and Cβ atoms were retained for the substituted amino acids and the side chains rebuilt. Loop regions, whose conformations could not be modeled directly from the BBP structure, were generated by conformational search calculations or loop searches. The model was minimized and equilibrated, followed by simulated annealing with a slow cool and a final minimization (22, 23). Two models of Bla g 4 were generated. Both models fit the eight-strand, anti-parallel β-barrel calycin structure but differ primarily in the conformation of the large loop between the α-helix and the COOH-terminal β strand (Fig. 6). The eight conserved residues, Asp, Gly, Trp (SCR 1), Thr-Asp-Tyr, Tyr (SCR 2), and Arg (SCR 3) (Fig. 5) are all located in similar positions (at the base of the "cup"), suggesting that this region could form a

binding site for the respective ligands (Fig. 6).

Expression of Bla g 4 in P. americana—Northern analysis was used to compare expression of mRNA encoding Bla g 4 in *B. germanica* and *P. americana*. An 0.75-kilobase *B. germanica* mRNA transcript hybridized with ^{32}P-labeled Bla g 4 cDNA probe but failed to hybridize to *P. americana* mRNA. Control experiments using a labeled *N. crassa* ribosomal DNA probe showed equivalent loading of mRNA from both CR species (Fig. 7A). To investigate the presence of genomic DNA encoding Bla g 4, PCR reactions on *B. germanica* and *P. americana* genomic DNA were carried out using Bla g 4-specific primers. A predicted 523-bp DNA was amplified from the genomic DNA of both CR species, and, in addition, a larger DNA (653 bp) was also detected (Fig. 7B). The PCR products were analyzed using ^{32}P-labeled Bla g 4 cDNA to hybridize to the two DNA bands on Southern blots. The two DNAs from both *B. germanica* and *P. americana* hybridized to the Bla g 4 DNA probe. Taken together, these results suggest that *B. germanica* and *P. americana* have genomic DNA encoding Bla g 4, that the DNA is only transcribed into mRNA in *B. germanica*, and, consequently, that the Bla g 4 protein is only expressed in *B. germanica*.

DISCUSSION

Using molecular cloning techniques, we have identified and sequenced an important *B. germanica* allergen, Bla g 4, which binds IgE antibodies in 40–60% of CR-allergic patients' sera. The recombinant allergen gave wheal and flare skin test responses at concentrations as low as 10^{-5} μg/ml. These results show unequivocally that Bla g 4 can elicit classical immediate hypersensitivity responses, and the allergen would, therefore, be expected to contribute toward the symptoms of CR-allergic patients. Bla g 4 is one of several CR allergens that we have cloned and sequenced, including Bla g 2 (an aspartic protease), Bla g 5 (a glutathione transferase), and Bla g 6 (troponin) (Fig. 1) (33, 34). Current evidence suggests that *B. germanica* produces at least five allergens that elicit IgE responses in 30–70% of CR-allergic patients (11–13, 33, 34). A high molecular mass allergen (72 kDa) has recently been cloned from *P. americana*, designated Per a 3; however, the homologous allergen does not appear to occur in *B. germanica* (35). Cloning of high molecular mass *B. germanica* allergens has also been reported (36).

Although previous studies have shown allergenic cross-reactivity between *B. germanica* and *P. americana* extracts, only one of the allergens purified to date (Bla g 1 and its homologue, Per a 1) have been shown to cross-react (11–13). The nucleic acid hybridization studies reported here show that genomic DNA encoding Bla g 4 is present in *P. americana* (as well as in *B. germanica*) but that the *P. americana* DNA does not appear to be transcribed into Bla g 4 mRNA. On the basis of this evidence, it appears that Bla g 4 is produced by *B. germanica* and not by *P. americana* and may be *Blattella* spp. specific, although its expression in a larger number of CR species remains to be tested. In the United States, most patients are primarily exposed and sensitized to *B. germanica*, whereas in the Far East (*e.g.* Taiwan and Japan), *P. americana* and *P. fuliginosa* appear to be more important causes of sensitization (9, 37, 38). Our skin test and serum IgE antibody results with rBla g 4 suggest that the recombinant allergen will be useful for diagnosis of allergic reactions to *B. germanica* and raise the possibility that, by using a mixture of allergens (including Bla g 2 and Bla g 5), it will be possible to use 3–4 recombinant proteins for both diagnostic (and therapeutic) purposes.

Recognition of Bla g 4 as a calycin (the first to be described in CR) provides insights into the biologic function of this protein. The homology with rodent urinary proteins, which are male pheromone transport proteins (32), raises the interesting possibility that Bla g 4 serves a similar function in CR. It is well
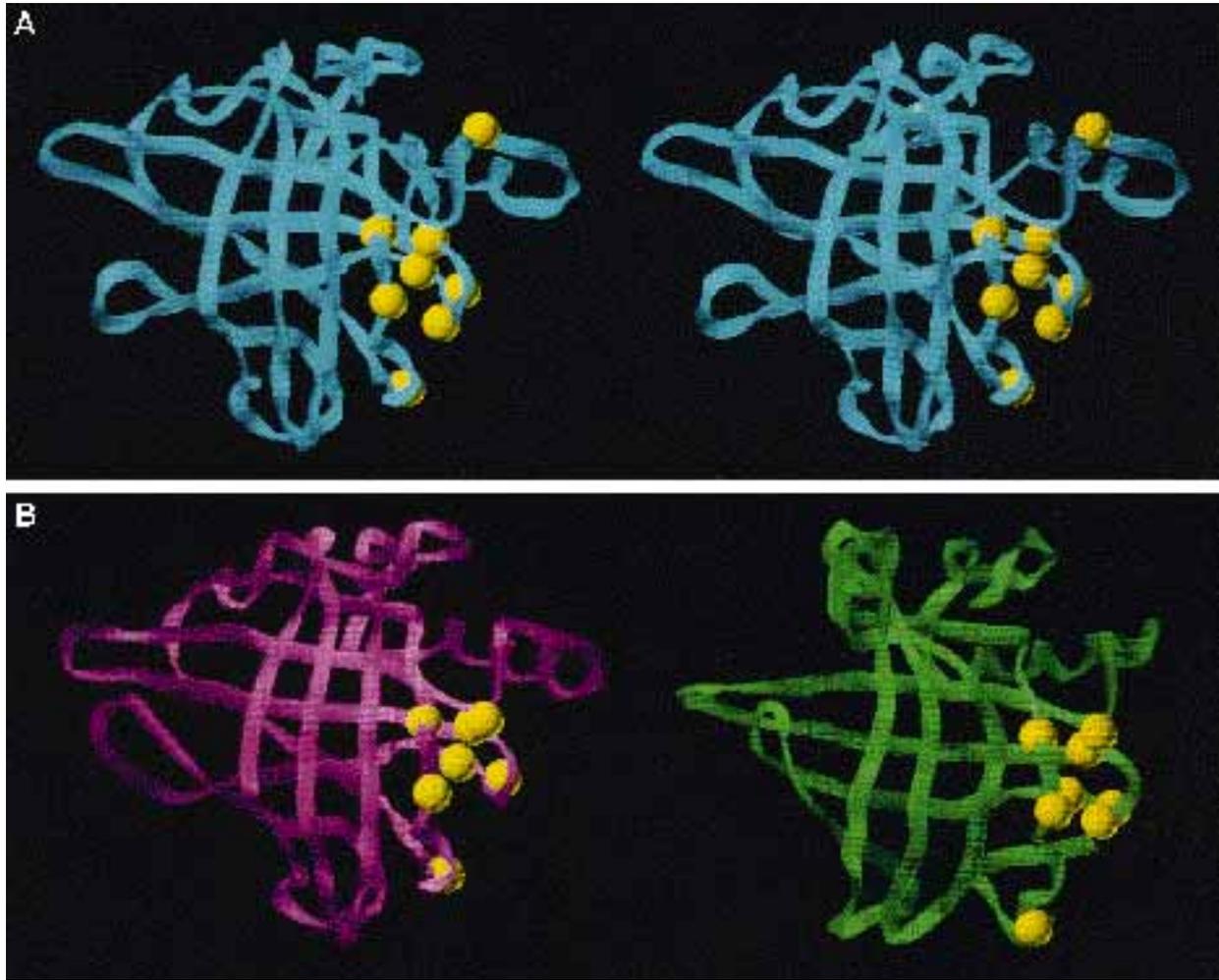
FIG. 6. **Molecular models of Bla g 4.** The Cα backbone structures for two models of Bla g 4 (*panel A*) are compared with those of butterfly BBP and rat urinary protein (*panel B*) (31, 32). Conserved residues within the calycin sequences are shown as *spheres*.
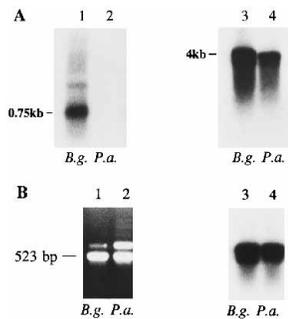
FIG. 7. **Analysis of mRNA and PCR products amplified from *B. germanica* (*B.g.*) or *P. americana* (*P.a.*) DNA.** *Panel A*, 2 μg of *B. germanica* or *P. americana* mRNA were hybridized with ^{32}P-labeled 650-bp Bla g 4 cDNA probe (*lanes 1* and *2*) or control ^{32}P-labeled *N. crassa* ribosomal DNA (*lanes 3* and *4*). *Panel B*, PCR primers were used to amplify Bla g 4 DNA from 100 ng of *B. germanica* or *P. americana* genomic DNA, and PCR products were analyzed on ethidium bromide gel (*lanes 1* and *2*) or by Southern blotting with ^{32}P-labeled Bla g 4 cDNA probe (*lanes 3* and *4*). A predicted DNA fragment of 523 bp, along with a 653-bp fragment, was amplified from both *B. germanica* and *P. americana*.

known that *B. germanica* produces a variety of pheromones, including aggregation pheromones, excreted in the frass, and volatile sex pheromones, produced in glands located on the posterior abdominal tergites (39, 40). There is also good evidence that CR allergens are secreted from CR bodies or excreted in the feces, and there are anecdotal reports that CR can

cause immediate skin reactions by crawling on the skin (41, 42). Thus, a plausible hypothesis is that Bla g 4 is a pheromone binding protein that is secreted along with male sex pheromones. The chemical structures of a number of CR pheromones have been defined, and it will now be possible to investigate this hypothesis using photoaffinity labeling techniques and NMR to analyze pheromone binding to rBla g 4 (43). Bla g 4 could bind other calycin ligands, such as odorants or pigments; however, insect odorant binding proteins belong to different protein families, and it is unlikely that odorant or pigment binding proteins would be secreted (44). If Bla g 4 is a pheromone binding protein, it could be a target for novel CR control strategies.

Our studies clearly show that calycins are a family of proteins that commonly cause IgE antibody responses. The calycin family contains several allergens associated with asthma (cockroach, rat, mouse, and dog) as well as an allergen associated with food hypersensitivity (cow milk β-lactoglobulin) (Fig. 5, Ref. 45). Equine allergen has also been reported to belong to the calycin family (46). Analysis of the published x-ray crystal structures of BBP and MUP calycins also allowed us to generate two molecular models of the tertiary structure of Bla g 4, which differ at a conformational loop region between the α-helix and COOH-terminal β strand. Using rBla g 4, either for crystallographic or NMR studies, it will be possible to determine which model is correct, and further refinement of the modeling programs may also be used to generate a "complete"

three-dimensional structure, as has recently been reported for mite Der p 1, by Topham *et al.* (47). Whether calycins themselves have intrinsic properties that stimulate IgE production remains to be established. IgG and IgE antibody responses and proliferative T cell responses to rodent urinary proteins have been measured, and recent data show IgE antibody binding to peptides from β-lactoglobulin (48–50). The availability of cloned Bla g 4 and other calycin sequences will enable these responses to be compared and the T cell epitopes involved in IgE responses to calycins to be defined. Since calycins have no enzymic activity, our results do not support the view that enzyme function *per se* is necessary to induce IgE responses, though enzymes could have adjuvant effects that enhance IgE production (51).

Rodent urinary proteins are the most abundant proteins secreted in the urine and become airborne on ~7-micron particles in laboratory animal rooms (or houses containing rats) (48). Inhalation of these particles, which stay airborne for several hours, causes IgE antibody responses and can provoke acute asthma attacks. Similarly, when large CR populations develop in substandard housing, CR allergen accumulates at high levels in the dust and becomes airborne on >10-micron particles following natural disturbance (7, 8).[2] The ability of antigen-presenting cells and T cells to recognize this transient, low dose antigen exposure at mucosal surfaces to a large extent determines whether or not individuals will mount IgE antibody responses to environmental allergens. The present studies will make it possible to develop new immunotherapeutic strategies for CR allergy, including T cell-based vaccines, and will also facilitate further analysis of the molecular events that mediate chronic inflammatory responses in CR-allergic patients with asthma.

REFERENCES

1. Bernton, H. S., McMahon, T. F., and Brown, H. (1972) *Br. J. Dis. Chest* **66**, 61–66
2. Twarog, F. J., Picone, F. J., Strunk, R. S., So, J., and Colten, H. R. (1977) *J. Allergy Clin. Immunol.* **59**, 154–160
3. Kang, B. (1976) *J. Allergy Clin. Immunol.* **58**, 357–365
4. Steinberg, D. R., Bernstein, D. I., Gallagher, J. S., Arlian, L., and Bernstein, I. L. (1987) *J. Allergy Clin. Immunol.* **80**, 586–590
5. Kang, B., Vellody, D., Homburger, H., and Yunginger, J. W. (1979) *J. Allergy Clin. Immunol.* **63**, 80–86
6. Pope, A. M., Patterson, R., and Burge, H. (1993) *Indoor Allergens: Assessing and Controlling Adverse Health Effects*, National Academy Press, Washington, D. C.
7. Call, R. S., Smith, T. F., Morris, E., Chapman, M. D., and Platts-Mills, T. A. E. (1992) *J. Pediatr.* **121**, 862–866
8. Gelber, L. E., Seltzer, L., Bouzoukis, J. K., Pollart, S. M., Chapman, M. D., and Platts-Mills, T. A. E. (1993) *Am. Rev. Respir. Dis.* **147**, 573–578
9. Wu, C. H., and Lan, J. L. (1988) *J. Allergy Clin. Immunol.* **82**, 727–735
10. Stankus, R. P., Horner, W. E., and Lehrer, S. B. (1990) *J. Allergy Clin. Immunol.* **86**, 781–787
11. Schou, C., Lind, P., Fernandez-Caldas, E., Lockey, R. F., and Lowenstein, H. (1990) *J. Allergy Clin. Immunol.* **86**, 935–946
12. Pollart, S. M., Mullins, D. E., Vailes, L. D., Hayden, M. L., Platts-Mills, T. A. E., Sutherland, W. M., and Chapman, M. D. (1991) *J. Allergy Clin Immunol.* **87**, 511–521
13. Helm, R. M., Bandele, E. O., Swanson, M. C., Campbell, A. R., and Wynn, S. R. (1988) *Int. Arch. Allergy Appl. Immunol.* **87**, 230–238
14. Chirgwin, J. M., Pryzbyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299
15. Sanger, F., Nicklen, S., and Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
16. King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., Platts-Mills, T. A. E., and Thomas, W. (1994) *Int. Arch. Allergy Immunol.* **105**, 224–233
17. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
18. Smith, D. B., and Johnson, K. S. (1988) *Gene (Amst.)* **67**, 31–40
19. Chapman, M. D., and Platts-Mills, T. A. E. (1980) *J. Immunol.* **125**, 587–592
20. Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 2444–2448
21. Chothia, C., Lesk, A. M., Gherardi, E., and Thomlinson, I. M. (1992) *J. Mol. Biol.* **227**, 799–817
22. Bruccoleri, R. E., and Karplus, M. (1987) *Biopolymers* **26**, 137–168
23. Bruccoleri, R. E., Haber, H., and Novotny, J. (1988) *Nature* **335**, 564–568
24. Arruda, L. K., Mann, B. J., and Chapman, M. D. (1992) *J. Immunol.* **149**, 3354–3359
25. Von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690
26. Sawyer, L. (1987) *Nature* **327**, 659
27. Li, W., and Riddiford, L. M. (1992) *Eur. J. Biochem.* **205**, 491–499
28. Pervaiz, S., and Brew, K. (1985) *Science* **228**, 335–337
29. Flower, D. R., North, A. C. T., and Attwood, T. K. (1993) *Protein Sci.* **2**, 753–761
30. Huber, R., Schneider, M., Epp, O., Mayr, I., Messerschmidt, A., Pflugrath, J., and Kayser, H. (1987) *J. Mol. Biol.* **195**, 423–434
31. Holden, H. M., Rypniewski, W. R., Law, J. H., and Rayment, I. (1987) *EMBO J.* **6**, 1565–1570
32. Böcskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E. V., Cavaggioni, A., Findlay, J. B. C., and North, A. C. T. (1992) *Nature* **360**, 186–188
33. Arruda, L. K., Vailes, L. D., Mann, B. J., Shannon, J., Fox, J. W., Vedvick, T. S., Hayden, M. L., and Chapman, M. D. (1995) *J. Biol. Chem.* **270**, 19563–19568
34. Vailes, L. D., Chapman, M. D., Wright, M., and Arruda, L. K. (1995) *J. Allergy Clin. Immunol.* **95**, 347 (abstr.)
35. Wu, C.-H., Lee, M.-F., and Liao, S.-C. (1995) *J. Allergy Clin. Immunol.* **96**, 352–359
36. Helm, R. M., Cockrell, G., Stanley, J. S., Brenner, R., Burks, A. W., and Bannon, G. A. (1995) *J. Allergy Clin. Immunol.* **95**, 158 (abstr.)
37. Lan, J. L., Lee, D. T., Wu, C. H., Chang, C. P., and Yeh, C. L. (1988) *J. Allergy Clin. Immunol.* **82**, 736–740
38. Sakaguchi, M., Inouye, S., Miyazawa, H., Okabe, T., Yasueda, H., Muto, A., Tanaka, I., Akasaka, T., Wu, C. H., Chapman, M. D., and Schou, C. (1994) *Jpn. J. Allergol.* **43**, 1309–1315
39. Schal, C., Tobin, T. R., Surber, J. L., Vogel, G., Tourtellot, M. K., Leban, R. A., Sizemore, R., and Bell, W. J. (1983) *J. Insect Physiol.* **29**, 575–579
40. Prestwich, G. D. (1987) *Science* **237**, 999–1006
41. Vailes, L., Glime, T., Pollart, S., and Chapman, M. (1990) *J. Allergy Clin. Immunol.* **85**, 171 (abstr.)
42. Lehrer, S. B., Horner, W. E., Menon, P., and Stankus, R. P. (1991) *J. Allergy Clin. Immunol.* **87**, 574–580
43. Prestwich, G. D. (1993) *Arch. Insect Biochem. Physiol.* **22**, 75–86
44. Krieger, J., Gänßle, H., Raming, K., and Breer, H. (1993) *Insect Biochem. Mol. Biol.* **23**, 449–456
45. Koznieczny, A., Morgenstern, J. P., Bizinkauskas, C. B., Kasaian, M., Meyer, C., Aalberse, R. C., and Wallner, B. P. (1995) *J. Allergy Clin. Immunol.* **95**, 161 (abstr.)
46. Dandeu, J. P., Rabillon, J., Carmi-Leroy, A., Divanovic, A., Camoin, L., and David, B. (1995) *J. Allergy Clin. Immunol.* **95**, 348 (abstr.)
47. Topham, C. M., Srinivasan, N., Thorpe, C. J., Overington, J. P., and Kalsheker, N. A. (1994) *Protein Eng.* **7**, 869–894
48. Platts-Mills, T. A. E., Heymann, P. W., Longbottom, J. L., and Wilkins, S. R. (1986) *J. Allergy Clin. Immunol.* **77**, 850–857
49. Gurka, G., Ohman, J., and Rosenwasser, L. J. (1989) *J. Allergy Clin. Immunol.* **83**, 945–954
50. Ball, G., Shelton, M. J., Walsh, B. J., Hill, D. J., Hosking, C. S., and Howden, M. E. H. (1994) *Clin. Exp. Allergy* **24**, 758–764
51. Stewart, G. A., Ward, L. D., Simpson, R. J., and Thompson, P. J. (1992) *Immunology* **75**, 29–35

[2] F. DeBlay, J. Sanchez, M. Chapman, A. Perez-Infante, A. Verot, M. Ott, and G. Pauli, submitted for publication.