

Short sequence-paper

Sequence analysis of the first complete cDNA clone encoding an American cockroach Per a 1 allergen

Chiou-Ying Yang ^{a,*}, Jen-Dau Wu ^a, Chii-Huei Wu ^b^a Institute of Molecular Biology, National Chung Hsing University, Taichung 402, Taiwan^b Department of Medical Research, Taichung Veterans General Hospital, Taichung 407, Taiwan

Received 13 July 2000; received in revised form 7 September 2000; accepted 7 September 2000

Abstract

C3, designated as Per a 1.0103 according to WHO/IUIS nomenclature, encoding a novel American cockroach allergen of 395 aa (44.6 kDa), is the first complete cDNA clone with a translatable immunoreactive protein among the reported group 1 cockroach allergens. Its deduced amino acid sequence possesses a signal sequence, phosphorylation sites, mitochondrial energy transfer protein signatures, and four repeats each sharing striking similarity with the corresponding repeats of the other cockroach allergens. The latter similarity suggests that the group 1 cockroach allergens might have evolved from a primordial mitochondrial sequence by exon duplication. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Cockroach; Allergen

Cockroach is one of the major indoor allergens that induce IgE-mediated allergic respiratory illnesses, such as allergic rhinitis and asthma [1–5]. Cockroach allergens with molecular masses ranging from 6 to 120 kDa have been identified by various immunochemical methods [6–8]. In American cockroach (*Periplaneta americana*), two major allergenic fractions, Per a 3 (Cr-PI) and Per a 1 (Cr-P II), each consisting of a family of isoallergens, have been purified from the crude extract [6] and specific antibodies have been generated [9,10]. Per a 3 and Per a 1 elicit skin reactivity in 73 and 93%, respectively, of the cockroach-sensitive patients [6]. Recently, another *P. americana* major allergen, Per a 7 (tropomyosin) cross-reacting with mite and shrimp allergens, has been identified [11,12]. In German cockroach (*Blattella germanica*), several allergens including Bla g 1, Bla g 2, Bla g 4 and Bla g 5 have been identified (13–18), among which Bla g 1 is antigenically cross-reactive with Per a 1 [13,14,19] whereas Bla g 2, Bla g 4 and Bla g 5 are German cockroach specific [15–17]. The nucleotide sequence has been determined for several cDNA clones of the cockroach allergens, including (i) four Per a 1 cDNAs: Per a 1.0101, Per a 1.0102 (C6), Per

a 1.0104 (C17) and Per a 1.02 (C42), encoding polypeptides of 231, 228, 274 and 446 aa, respectively [20–22], (ii) four Per a 3 cDNAs: Per a 3.01 (C12), Per a 3.0201 (C20), Per a 3.0202 (C13) and Per a 3.0203 (C28), encoding polypeptides of 685, 631, 470 and 393 aa, respectively [23,24], (iii) two Per a 7 cDNAs encoding polypeptides of 284 aa with three different amino acids [11,12], (iv) three Bla g 1 cDNAs: Bla g 1.01, Bla g 1.02 and Bla g Bd90K encoding polypeptides of 412, 492 and undetermined aa, respectively [25], and (v) one cDNA clone each from Bla g 2, Bla g 4 and Bla g 5 encoding polypeptides of 352, 182 and 200 aa, respectively [15–17]. Among these clones, only Per a 3.01 [24], Per a 7 [11,12] and Bla g 2 [16] contain complete coding sequence. Here, we report sequencing and expression of the first complete cDNA sequence of Per a 1.0103 (C3) coding for a novel Per a 1 allergen.

A cDNA clone recognized by rabbit anti-Per a 1 (Cr-P II) and reactive with IgE antibodies of atopic serum pool has previously been isolated from a λ gt22A cDNA expression library derived from *P. americana* mRNA and designated as C3 [21]. In this study, the 1.5-kb insert was released by digestion of C3 with *SalI* plus *NotI* and ligated into the compatible sites of pBluescript II SK(+), followed by subcloning the subfragments into pBluescript vectors. Both strands of the inserts were sequenced by the dideoxynucleotide chain termination method [26] using T3 and T7 primers. A total of 1433 bp was revealed (GenBank

* Corresponding author. Fax: +886-4-287-4879;
E-mail: cyyang@dragon.nchu.edu.tw

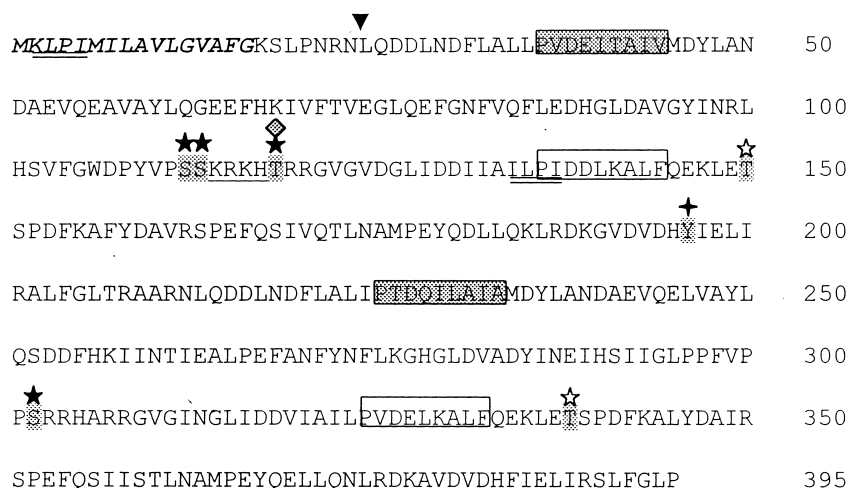


Fig. 1. The deduced amino acid sequence of Per a 1.0103 (C3) allergen. Single-letter amino acid code is used. The putative signal peptide is italic and bold; the predicted cleavage site for mitochondrial presequence is indicated by an arrowhead. Mitochondrial energy transfer protein signatures are boxed, and the degenerate sequences are shaded. Also indicated are phosphorylation sites for cAMP- and cGMP-dependent protein kinase (◆), protein kinase C (★), casein kinase II (☆), and tyrosine kinase (+), a nuclear localization signal (underlined), and two possible vacuolar targeting motifs (≡).

accession number U69957). Beginning with a G, this cDNA contained a 22-bp 5'-untranslated region, an open reading frame (ORF) starting with ATG at nt 23 and terminating with TGA at nt 1207, and a 223-bp-long 3'-untranslated region including an AATAAA polyadenylation signal and a 53-bp poly A tail. The presence of these features suggests C3 to be a complete cDNA molecule. Interestingly, a Shine–Dalgarno (S/D) sequence (AGAGGAA) of *Escherichia coli* was present at nt 5–11 of the C3 cDNA. This putative S/D sequence appeared to be functional in our expression experiments with *E. coli* (see below). Translation with the first ATG (nt 23) as the initiation codon would produce a polypeptide of 395 amino acids (Fig. 1) with a calculated molecular mass of 44.62 kDa and a predicted isoelectric point value of 4.4. No cysteine or any potential *N*-glycosylation site was found

in the protein product deduced from the C3 nucleotide sequence.

To ensure that the predicted ORF was correct, the inserted 1.5-kb *SalI–NotI* fragment containing the full-length cDNA was cloned into pET30a, b, and c vectors generating pET30a-C3, pET30b-C3, and pET30c-C3, respectively. Total proteins from *E. coli* BL21(DE3) carrying the respective plasmids were analyzed by SDS-PAGE and Western blotting using anti-Cr-PII monoclonal antibodies [9] as the primary antibody (Fig. 2). A strong immunoreactive band above the 45-kDa marker was detected in the pET30a-C3 clone after IPTG induction, confirming the notion that the predicted ORF indeed encodes a product of Per a 1 family. A band at the position below the 45-kDa marker was also observed in the IPTG-induced cells containing pET30b-C3 and pET30c-C3, suggesting that

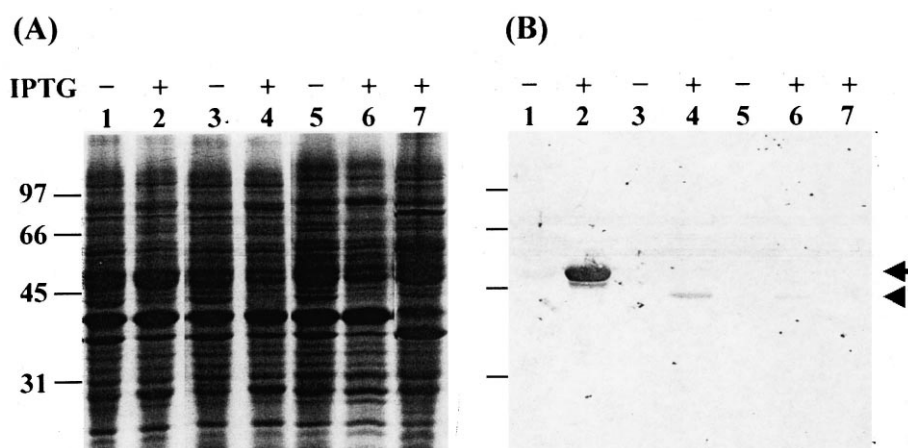


Fig. 2. SDS-PAGE (A) and immunoblot (B) analyses of total proteins from *E. coli* BL21(DE3) containing pET30a-C3 (lanes 1 and 2), pET30b-C3 (lanes 3 and 4), pET30c-C3 (lanes 5 and 6), and the control pET30a (lane 7). The cultures were grown in LB for 6 h at 30°C with (+) or without (–) IPTG (1 mM). The SDS-polyacrylamide gel was stained with Coomassie blue (A) or probed with murine anti-Cr-PII monoclonal antibodies (B). The arrow and the arrowhead point to the positions corresponding to the fusion and non-fusion C3 proteins, respectively. Numbers at left indicate sizes of protein markers in kDa.

C3N	22	RNLQDDLND	LALLPVDEIT	AIVMDYLAND	AEVQEAV-AY	LQGEFHKIV	70
C3C	211	*****T*	Q*L	**A*****	*****L*-*	**SDD****I	259
C6	44	*****T*	Q*L	**A*****	*****L*-*	**SDD****I	92
C17	90	*****T*	Q*L	**A*****	*****L*-*	**SDD****I	138
C42	113	*****HS*	.EP*S	PRHRHG*PRQ	RRRSAR*S**	*HADD****I	160
Bla g 1.01	226	*****I*	QII	**AT*****	*****A*-*	**SD**ET**	274
Bla g 1.02	308	*****I*	Q*L	**AA*****	*****A*-E*	*KSD**ET**	356
Cr4H	1	*****K*	V**Q*I	**AT**I**	*****A*-*	**SD**ET**	49
C3N	71	FTVEGLQEF	GF NQVQFLEDHG	LDAVGYNRL	HSVFGWDPYV	PSSKRK-HTR	119
C3C	260	T*I***P**A	**YN**KE**	**VADFL*EI	**II*LP*F*	*P*-R-*A*	307
C6	93	T*I***P**A	**YN**KE**	**VADFL*EI	**II*LP*F*	*P*-R-*A*	140
C17	139	N*I*A*P**A	**YN**KG**	**VAN**EI	**II*LP*F*	*P*-R-*A*	186
C42	161	T*I*A*P**A	**YN**KE**	**V*D***EI	**II*LP*F*	*P*-R-*A*	208
Bla g 1.01	275	V*LDA*P*LQ	**LN***AN*	*N*IDFL*GI	*DLL*I-*HI	*V*G**Y*I*	323
Bla g 1.02	357	V**DS*P**K	**LN**QTN*	*N*IEF**NI	*DLL*I-*HI	*ATG**-*V*	404
Cr4H	50	V**DA*P*LQ	**LN***TN*	*N*IDFL*GI	*DLL*I-*HI	*V*G**Y*I*	98
C3N	120	RGVGVGLID	DIILAILPIDD	LKALFQEKLE	TSPDFKAFYD	AVRSPEFQSI	169
C3C	308	****IN****	*V*****V*E	*****L**	*I*****		357
C6	141	****IN****	*V*****V*E	*****D****	*****L**	*I*****	190
C17	187	****IN****	*V*****V*E	**T*****	*****L**	*I*****	236
C42	209	****IN****	*V*****V*E	*****L**	*I*****		258
Bla g 1.01	324	****IT****	*VL*****L**	*****N****	*****L*L*N	*IK*****	373
Bla g 1.02	405	****IN****	*V*****V*E	*Y*****	S**E***L**	*I*****	454
Cr4H	99	****IT****	*VL*****	*****N****	*****L*L*N	*IK*****	148
C3N	170	VQTLNAMPEY	QDLLQKLRDK	GVDVDHYIEL	IRALFGLT		207
C3C	358	IS*****E*	*N****	A*****F***	**S****P		395
C6	191	IS*****E*	*N****	*****F***	**S****P		228
C17	237	IS*****E*	*N****	*****F***	**SW****P		274
C42	259	IS*****E*	*N****	*****F*RV	DQGTLRITL		296
Bla g 1.01	374	*****N*	E***E**E*	*****KI***	*****S		401
Bla g 1.02	455	*****K****	*I*R*K**	*****F***	*KK*****S		492
Cr4H	149	*****D****	*N*****E*	*****KI***	*****S		186

Fig. 3. Sequence alignments of the two tandem repeats, C3N and C3C, of Per a 1.0103 with Per a 1.0102 (C6) (U78970), Per a 1.0104 (C17) (U69261), and Per a 1.02 (C42) (U69260) of Per a 1 allergen, Bla g 1.01 (AF072219), Bla g 1.02 (AF072220), and the 'repeat 1' sequence of Bla g Bd90K (L47595) from *B. germanica*. Single-letter amino acid code is used, and gaps are introduced to optimize alignment. Asterisk represents a residue identical to that of C3N.

the potential S/D sequence might be functional in producing non-fusion C3 protein molecules. To test this possibility, the *HindIII*–*NotI* fragment spanning nt 26–1433 of C3 cDNA, with the potential S/D sequence being truncated, was cloned into pET30b and pET30c that were then separately subjected to IPTG induction. The results showed that no specific protein product was produced (data not shown), indicating that the smaller (<45 kDa) immunoreactive band was indeed a non-fusion C3 protein.

Sequence comparison using BLAST revealed that the amino acid sequence deduced from C3 showed very high degrees of identity with the Per a 1 allergens, Per a 1.0101 (95%), Per a 1.0102 (95%), Per a 1.0104 (97.8%), and Per a 1.02 (74%) [20–22], indicating that they are isoallergens. According to the WHO/IUIS nomenclature for allergens, the allergen encoded by C3 clone was designated as Per a 1.0103 [20]. Like the other members of the Per a 1 family, Per a 1.0103 allergen is also highly homologous to the *B. germanica* allergens [18,25]; 67.8% identity and 81.8% similarity were shared with the 412-aa Bla g 1.01 clone, 65.3% identity and 79.9% similarity were shared with the 492-aa

Bla g 1.02 clone, and 70.9% identity and 80.8% similarity were shared with the translatable 192-aa repeats of Bla g Bd90K cDNA clone. Moderate degree of homology, 32% identity and 58% similarity, was found between aa 211–393 of Per a 1.0103 and aa 24–209 of ANG12 antigen (Z22925), the precursor protein (211 aa) of African mosquito (*Anopheles gambiae*) that is found exclusively in the midgut of female mosquito and is induced following a blood meal. Significant homology was also found between Per a 1.0103 and some other proteins. Among these are six myosin genes from human, rat, and mouse, each sharing 33% identity and 54% similarity aligning with aa 326–379 of Per a 1.0103, one hypothetical protein from *Caenorhabditis elegans* similar to serine/threonine protein kinase aligning with aa 269–376 of Per a 1.0103 (23% identity and 42% similarity), one adenylate cyclase homolog from *Spirulina platensis* aligning with aa 326–374 of Per a 1.0103 (37% identity and 55% similarity), one putative calcium P-type ATPase from *Neurospora crassa* aligning with aa 194–275 of Per a 1.0103 (31% identity and 43% similarity), one polycystic kidney disease-related gene

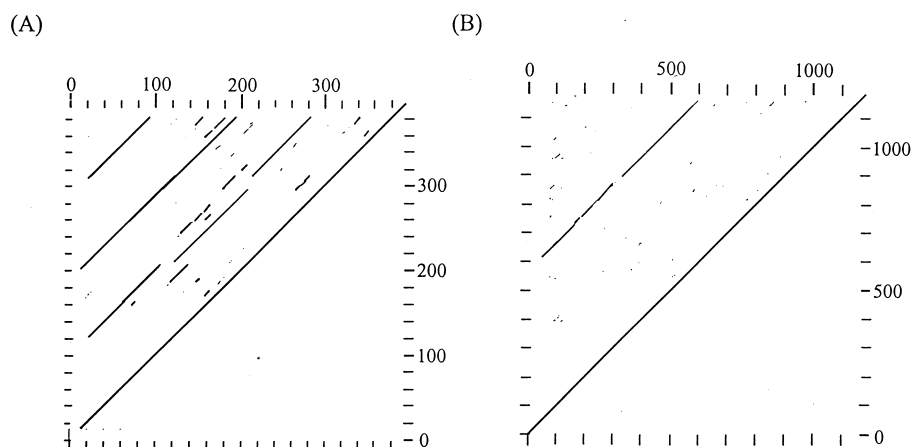


Fig. 4. Intrasequence comparison of Per a 1.0103 by dotplot matrix. (A) The primary translation product with a span of 30 residues and the Blosum62 matrix [30], and (B) the coding region of Per a 1.0103 cDNA with a window size of 21 nucleotides and a stringency of 66.7% (14/21 bp).

aligning with aa 48–99 of Per a 1.0103 (30% identity and 60% similarity), and 11 proteins with unknown function derived from *Drosophila melanogaster*, *Arabidopsis thaliana*, *Amycolatopsis orientalis*, and human aligning within aa 195–395 of Per a 1.0103 (with identities ranging 24–35%).

Similarity search using SCANPS revealed that Per a 1.0103 protein contained two tandem repeats, locating at positions 22–206 (C3N) and 211–395 (C3C), with 74.2% identity and 85.5% similarity (Fig. 3). Alignment of these repeats with Per a 1.0102 (C6), Per a 1.0104 (C17), Per a 1.02 (C42) and the *B. germanica* Bla g 1 and Bla g Bd90K (Cr4H) revealed a high level of homology among them (Fig. 3). The most conserved region was found in the C-termini of these sequences. Dotplot matrix analysis [27], in which repeated fragments appear as parallel lines above the diagonal, however, revealed four entire amino acid repeats of ~100 amino acids long and two nucleotide repeats of ~564 nucleotides in Per a 1.0103 (Fig. 4). This structural feature has also been described for Bla g 1 and Bla g Bd90K [25]. Because internal repeats in proteins often represent either structural modules or functional domains [28], the results of dotplot indicate that the Per

a 1.0103 protein might consist of four modules/domains. The four amino acid repeats were designated as repeat 1 (aa 1–99), repeat 2 (aa 100–198), repeat 3 (aa 199–297), and repeat 4 (aa 298–395). High degrees of similarity (69–97%) were found between repeats 1 and 3 (74%), repeats 2 and 4 (92%), and each of the repeats with the other members of Per a 1 and Bla g Bd90K allergens (Table 1).

Pattern and profile searches indicated that the Per a 1.0103 protein contained phosphorylation sites (Fig. 1) for cAMP- and cGMP-dependent protein kinase, protein kinase C, casein kinase II, and tyrosine kinase. At aa positions 136–145 and 324–333 of Per a 1.0103 protein were a motif P-x-[DE]-x-[LIVAT][RK]-x-[LRH]-[LIVM-FY] characteristic of mitochondrial energy transfer proteins (Fig. 1). This motif is also present in the other members of Per a 1 and degenerate in Bla g 1.02 and the ANG12 precursor. Proteins belonging to this family are evolutionarily related. Structurally, they consist of three tandem repeats of approximately 100 amino acid residues, each of which containing two transmembrane regions [29]. However, the program PSORT predicted that Per a 1.0103 protein had a cleavable signal peptide (aa 1–16), a mitochondrial targeting sequence, a nuclear localization signal

Table 1

Pairwise comparison of the internal repeat sequences among Per a 1 members, Bla g 1 and Bla g Bd90K (Cr4H)^a

	Repeat 1 ^b	Repeat 2	Repeat 3	Repeat 4	Per a 1.0102	Per a 1.02	Per a 1.0104	Bla g 1.01	Bla g 1.02	Cr4H
Repeat 1 ^b		55	74	52	76	82	74	72	69	70
Repeat 2	28		51	92	78	89	89	86	90	88
Repeat 3	62	26		57	96	93	93	80	75	75
Repeat 4	27	80	30		97	97	97	85	91	85
Per a 1.0102	61	78	94	96						
Per a 1.02	72	77	92	96						
Per a 1.0104	62	77	97	96						
Bla g 1.01	56	74	66	76						
Bla g 1.02	52	73	58	83						
Cr4H	58	80	60	79						

^aThe numbers above and below the diagonal are the percentage of similarity and identity, respectively.

^bRepeat 1, aa 1–99; repeat 2, aa 100–198; repeat 3, aa 199–297; repeat 4, aa 298–395.

(KRKH, aa 114–117), and two possible vacuolar targeting motifs (KLPI, aa 2–5, and ILPI, aa 134–137) (Fig. 1) but no transmembrane region. Overall, these analyses have predicted that the Per a 1.0103 protein is most likely directed to the exterior of the cells following the secretory pathway through endoplasmic reticulum or into the mitochondrial matrix.

Based on the BLAST results showing that the sequence similarity shared between Per a 1.0103 and the other subject proteins falls within the C-terminus, the degrees of similarities shared between the amino acid repeats (51–92%) and that of the other members of Per a 1 and Bla g Bd90K (Table 1), and mitochondrial energy transfer protein signatures are present in repeats 2 and 4 and degenerated in repeats 1 and 3 (Fig. 1), we propose that repeat 4 may represent the primordial domain evolving from a mitochondrial sequence. In other words, Per a 1.0103 was most likely formed via the following steps: (i) duplication of repeat 4 followed by exon fusion together with the accompanied divergence resulted in the formation of the ‘duplex exon’ encoding the C3C, and (ii) duplication of the ‘duplex exon’ corresponding to the C3N to form the full-length Per a 1.0103. Because this unique structure is shared by all the members of Per a 1 and Bla g 1 allergens identified to date, it seems reasonable to propose that the ‘group 1’ allergens of cockroach were arisen by replicating the ‘duplex’ with varied times. The homology between Per a 1.0103 and ANG12 indicated that ANG12 precursor was degenerated from the original ‘duplex’ structure during evolution.

The presence of phosphorylation sites for a variety of protein kinases suggests that the biological function of Per a 1.0103 may be modulated by phosphorylation that is common for proteins involved in signal transduction and some metabolic pathway. Another possible biological role Per a 1.0103 might play is in muscle contraction since it shares homology with different isoforms of myosin and ATPase. Taken together the results described above, we propose that the fourth domain of Per a 1.0103 may be a copy of an ancient gene serving as a ‘generator’ to create new genes during evolution. Complex genes could then be formed by domain duplication. This proposal seems reasonable, taken into consideration that cockroach appeared in the very primitive time.

Although cDNA coding for different group I allergens of cockroach have been reported [20–22,25], the C3 clone (Per a 1.0103) described here is the first complete cDNA with a translatable product recognized by the Per a 1 specific antibodies. The results presented here support the proposal that the ‘duplex’ may represent the basic allergenic unit of cockroach ‘group 1’ allergen [25].

This work was supported by National Science Council of the Republic of China Grant NSC88-23111-B-005-038. The authors would like to thank Dr. Y.-H. Tseng for reading the manuscript and Mr. C.-H. Hung for assistance in preparation of the manuscript.

References

- [1] C. Cuesta, J.L. Placido, L. Delgado, M. Miranda, J.P. Moreira Silva, M.G. Castel-Branco, M. Vaz, Cockroach allergy: a study of its prevalence using skin tests with commercial extracts, *Allergol. Immunopathol. (Madr.)* 23 (1995) 295–300.
- [2] A. Custovic, A. Simpson, A. Woodcock, Importance of indoor allergens in the induction of allergy and elicitation of allergic disease, *Allergy* 53 (1998) 115–120.
- [3] B.C. Kang, Cockroach allergy, *Clin. Rev. Allergy* 8 (1990) 87–98.
- [4] S.B. Sarpong, J.P. Corey, Assessment of the indoor environment in respiratory allergy, *Ear Nose Throat J.* 77 (1998) 960–964.
- [5] N.S. Zhong, New insights into risk factors of asthma, *Respirology* 1 (1996) 159–166.
- [6] C.H. Wu, J.L. Lan, Cockroach hypersensitivity: isolation and partial characterization of major allergens, *J. Allergy Clin. Immunol.* 82 (1988) 727–735.
- [7] B.C. Kang, M. Wilson, K.H. Price, T. Kambara, Cockroach-allergen study: allergen patterns of three common cockroach species probed by allergic sera collected in two cities, *J. Allergy Clin. Immunol.* 87 (1991) 1073–1080.
- [8] R.P. Stankus, W.E. Horner, S.B. Lehrer, Identification and characterization of important cockroach allergens, *J. Allergy Clin. Immunol.* 86 (1990) 781–787.
- [9] C.H. Wu, M.J. Hsieh, J.H. Huang, S.F. Luo, Identification of low molecular weight allergens of American cockroach and production of monoclonal antibodies, *Ann. Allergy Asthma Immunol.* 76 (1996) 195–203.
- [10] C.H. Wu, B.T. Chiang, M.C. Fann, J.L. Lan, Production and characterization of monoclonal antibodies against major allergens of American cockroach, *Clin. Exp. Allergy* 20 (1990) 675–681.
- [11] J.A. Asturias, N. Gomez-Bayon, M.C. Arilla, A. Martinez, R. Palacios, F. Sanchez-Gascon, J. Martinez, Molecular characterization of American cockroach tropomyosin (*Periplaneta americana* allergen 7), a cross-reactive allergen, *J. Immunol.* 162 (1999) 4342–4348.
- [12] A.B. Santos, M.D. Chapman, R.C. Aalberse, L.D. Vailes, V.P. Ferriani, C. Oliver, M.C. Rizzo, C.K. Naspitz, L.K. Arruda, Cockroach allergens and asthma in Brazil: identification of tropomyosin as a major allergen with potential cross-reactivity with mite and shrimp allergens, *J. Allergy Clin. Immunol.* 104 (1999) 329–337.
- [13] S.M. Pollart, D.E. Mullins, L.D. Vailes, M.L. Hayden, T.A. Platts-Mills, W.M. Sutherland, M.D. Chapman, Identification, quantitation, and purification of cockroach allergens using monoclonal antibodies, *J. Allergy Clin. Immunol.* 87 (1991) 511–521.
- [14] C. Schou, P. Lind, E. Fernandez-Caldas, R.F. Lockey, H. Lowenstein, Identification and purification of an important cross-reactive allergen from American (*Periplaneta americana*) and German (*Blattella germanica*) cockroach, *J. Allergy Clin. Immunol.* 86 (1990) 935–946.
- [15] L.K. Arruda, L.D. Vailes, M.L. Hayden, D.C. Benjamin, M.D. Chapman, Cloning of cockroach allergen, Bla g 4, identifies ligand binding proteins (or calycons) as a cause of IgE antibody responses, *J. Biol. Chem.* 270 (1995) 31196–31201.
- [16] L.K. Arruda, L.D. Vailes, B.J. Mann, J. Shannon, J.W. Fox, T.S. Vedvick, M.L. Hayden, M.D. Chapman, Molecular cloning of a major cockroach (*Blattella germanica*) allergen, Bla g 2. Sequence homology to the aspartic proteases, *J. Biol. Chem.* 270 (1995) 19563–19568.
- [17] L.K. Arruda, L.D. Vailes, T.A. Platts-Mills, M.L. Hayden, M.D. Chapman, Induction of IgE antibody responses by glutathione S-transferase from the German cockroach (*Blattella germanica*), *J. Biol. Chem.* 272 (1997) 20907–20912.
- [18] R. Helm, G. Cockrell, J.S. Stanley, R.J. Brenner, W. Burks, G.A. Bannon, Isolation and characterization of a clone encoding a major allergen (Bla g Bd90K) involved in IgE-mediated cockroach hypersensitivity, *J. Allergy Clin. Immunol.* 98 (1996) 172–180.

- [19] C.H. Wu, S.F. Luo, D.W. Wong, Analysis of cross-reactive allergens from American and German cockroaches by human IgE, *Allergy* 52 (1997) 411–416.
- [20] E. Melen, A. Pomes, L.D. Vailes, L.K. Arruda, M.D. Chapman, Molecular cloning of Per a 1 and definition of the cross-reactive Group 1 cockroach allergens, *J. Allergy Clin. Immunol.* 103 (1999) 859–864.
- [21] C.H. Wu, N.M. Wang, M.F. Lee, C.Y.Y. Kao, S.F. Luo, Cloning of the American cockroach Cr-P11 allergens: evidence for the existence of cross-reactive allergens between species, *J. Allergy Clin. Immunol.* 101 (1998) 832–840.
- [22] N.M. Wang, M.F. Lee, C.H. Wu, Immunologic characterization of a recombinant American cockroach (*Periplaneta americana*) Per a 1 (Cr-P11) allergen, *Allergy* 54 (1999) 119–127.
- [23] C.H. Wu, M.F. Lee, N.M. Wang, S.F. Luo, Sequencing and immunochemical characterization of the American cockroach per a 3 (Cr-P1) isoallergenic variants, *Mol. Immunol.* 34 (1997) 1–8.
- [24] C.H. Wu, M.F. Lee, S.C. Liao, S.F. Luo, Sequencing analysis of cDNA clones encoding the American cockroach Cr-P1 allergens. Homology with insect hemolymph proteins, *J. Biol. Chem.* 271 (1996) 17937–17943.
- [25] A. Pomes, E. Melen, L.D. Vailes, J.D. Retief, L.K. Arruda, M.D. Chapman, Novel allergen structures with tandem amino acid repeats derived from German and American cockroach, *J. Biol. Chem.* 273 (1998) 30801–30807.
- [26] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, *Proc. Natl. Acad. Sci. USA* 74 (1977) 5463–5467.
- [27] J.V. Maizel Jr., R.P. Lenk, Enhanced graphic matrix analysis of nucleic acid and protein sequences, *Proc. Natl. Acad. Sci. USA* 78 (1981) 7665–7669.
- [28] W. Barker, L. Ketcham, M. Dayhoff, Duplications in protein sequences, in: M.O. Dayhoff (Ed.), *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, Natl. Biomed. Res. Found, Silver Spring, MD, 1978, pp. 359–362.
- [29] F. Palmieri, Mitochondrial carrier proteins, *FEBS Lett.* 346 (1994) 48–54.
- [30] S. Henikoff, J.G. Henikoff, Amino acid substitution matrices from protein blocks, *Proc. Natl. Acad. Sci. USA* 89 (1992) 10915–10919.