[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

3 Allergen Nomenclature

Martin D. Chapman
Indoor Biotechnologies, Inc., Charlottesville, Virginia, U.S.A.

HISTORICAL INTRODUCTION

As with most biochemical disciplines, the history of allergen nomenclature dates back to the time when allergens were fractionated using a variety of "classical" biochemical separation techniques, and the active (most allergenic) fraction was usually named according to the whim of the investigator. Early attempts were made to purify pollen and house-dust allergens, using phenol extraction, salt precipitation, and electrophoretic techniques in the 1940–1950s. In the 1960s, ion exchange and gel filtration media were introduced and ragweed "antigen E" was the first allergen to be purified. This allergen was named by King and Norman because it was one of five precipitin lines (labeled A–E) that reacted rabbit polyclonal antibodies to ragweed in Ouchterlony immunodiffusion tests. Following purification, precipitin line E, or antigen E, was shown to be a potent allergen (1). Later, Marsh, working in Cambridge, England, isolated an important allergen from rye grass pollen (*Lolium perenne*) and used the name "Rye 1" to indicate that this was the first allergen purified from this species (2,3). In the 1970s, many allergens were purified from ragweed, rye grass, insect venoms, and other sources. The field was led by the laboratory of the late Dr. David Marsh, who had moved to the Johns Hopkins University, Baltimore, Maryland. At Hopkins, ragweed allergens, Ra3, Ra4, Ra5, and Ra6, and rye grass allergens, Rye 2 and Rye 3, were isolated and used for immunological and genetic studies of hay fever (4–6). At the same time, Ohman et al. identified the major cat allergen (Cat-I) (7) and Elsayed purified allergen M from codfish (8,9).

The state of the art in the early 1970s was reviewed in a seminal book chapter by Marsh in *The Antigens* (ed. Michael Sela), which described the molecular properties of allergens, the factors that influenced allergenicity, the immune response to allergens, and immunogenetic studies of IgE responses to purified pollen allergens (10). This chapter provided the first clear definition of a "major" allergen, which Marsh defined as a highly purified allergen that induced immediate skin test responses in >90% of allergic individuals in contrast to a "minor" allergen, to which <20% of patients gave skin test responses. Today, a major allergen is generally regarded as one to which >50% of allergic patients react (11).

With the introduction of crossed immunoelectrophoresis (CIE) and crossed radio-immunoelectrophoresis (CRIE) for allergen identification by Lowenstein and colleagues in Scandinavia, there was a tremendous proliferation of the number of antigenic proteins and CIE/CRIE peaks identified as allergens. Typically, 10 to 50 peaks could be detected in a given allergen based on reactivity with rabbit polyclonal antibodies or IgE antibodies (6,11–13). These peaks were given a plethora of names such as Dp5, Dp42, Ag 12, etc. Inevitably, the same allergens were referred to by different names in different laboratories, e.g., mite Antigen P_1 was also known as Dp42 or Ag12. It was clear that a unified nomenclature was urgently needed.

Three Men in a Boat

The origins of the systematic allergen nomenclature can be traced to meeting among Drs. David Marsh (at that time, Johns Hopkins University, Baltimore, U.S.), Henning Lowenstein (at that time, University of Copenhagen, Denmark), and Thomas Platts-Mills (at that time, Clinical Research Centre, Harrow, U.K.) on a boatride on Lake Boedensee, Konstanz, Germany, during the 13th Symposium of the Collegium Internationale Allergologicum in July 1980 (14). The idea was simply to develop a systematic nomenclature based on the Linnean system, with numerals to indicate different allergens. It was decided to adopt a system whereby the allergen was described based on the first three letters of the genus and the first

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

letter of the species (in italics) and then by a Roman numeral to indicate the allergen in the chronological order of purification. Thus, ragweed antigen E became *Ambrosia artemisiifolia* allergen I or *Amb a* I and Rye 1 became *L. perenne* allergen I or *Lol p I*.

An allergen nomenclature subcommittee was formed under the auspices of the World Health Organization and International Union of Immunological Societies (WHO/IUIS) and criteria for including allergens in the systematic nomenclature were established. These included strict criteria for biochemical purity as well as criteria for determining the allergenic activity of the purified protein. A committee chaired by Marsh and including Lowenstein, Platts-Mills, Drs. Te Piao King (Rockerfeller University, New York, U.S.), and Larry Goodfriend (McGill University, Canada) prepared a list of allergens that fulfilled the inclusion criteria and established a process for investigators to submit names of newly identified allergens. The original list, published in the Bulletin of the WHO in 1986, included 27 highly purified allergens from grass, weed and tree pollens, and house-dust mites (15).

The systematic allergen nomenclature was quickly adopted by allergy researchers and proved to be a great success. It was logical, easily understood, and readily assimilated by allergists and other clinicians who were not directly involved with the nitty-gritty of allergen immunochemistry. The nomenclature, *Der p* I, *Fel d* I, *Lol p* I, *Amb a* I, was used at scientific meetings and in the literature, and expanded rapidly to include newly isolated allergens.

THE REVISED ALLERGEN NOMENCLATURE

Allergens

The widespread use of molecular cloning techniques to identify allergens in the late 1980s and 1990s led to an exponential increase in the number of allergens described. Many allergen nucleotide sequences were generated from cDNA cloning or PCR-based sequencing, and it soon became apparent that the use of Roman numerals was unwieldy (e.g., *Lol p* I through *Lol p* XI) (16,17). The use of italics to denote a purified protein was inconsistent with nomenclature used in bacterial genetics and the HLA system, where italicized names denote a gene product and regular typeface indicates expressed proteins. In 1994, the allergen nomenclature was revised so that the allergen phenotype was shown in regular type and Arabic numerals were adopted. Thus, *Amb a* I, *Lol p* I, and *Der p* I in the original 1986 nomenclature are now referred to as Amb a 1, Lol p 1, and Der p 1 in the current nomenclature, which has been published in several scientific journals (18–20).

Inclusion Criteria

A key part of the systematic WHO/IUIS nomenclature is that the allergen should satisfy biochemical criteria, which define the molecular structure of the protein, and immunological criteria, which define its importance as an allergen. Originally, the biochemical criteria were based on establishing protein purity (e.g., by SDS-PAGE, IEF, or HPLC and physicochemical properties including MW, pI, and N-terminal amino acid sequence) (20). Nowadays, the full nucleotide or amino acid sequence is generally required. An outline of the inclusion criteria is shown in Table 1. An important aspect of these criteria is that they should provide a "handle" whereby other investigators can identify the same allergen and make comparative studies. Originally, this was achieved by purifying the protein, developing monospecific or monoclonal antibodies to it, and providing either the allergen or antibodies to other researchers for verification. Nucleotide and amino acid sequencing unambiguously identifies the allergen and enables sequence variation between cDNA clones of the same allergen to be defined (21–24). Allergen preparations, sequences, and antibodies submitted for inclusion in the systematic nomenclature are expected to be made available to other investigators for research studies.

A second set of inclusion criteria involves demonstrating that the purified allergen has allergenic activity, both in vitro and in vivo. Researchers use a variety of techniques for measuring IgE antibodies in vitro, including radioallergosorbent (RAST)-based techniques, immunoblotting, radioimmunoassays using labeled allergens, enzyme immunoassay (ELISA), and fluorescent enzymeimmunoassay (FEIA). It is important to screen a large number of sera from an unselected allergic population to establish the prevalence of IgE reactivity. Ideally,

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

Table 1 Allergens: Criteria for Inclusion in the WHO/IUIS Nomenclature

The molecular and structural properties should be clearly and unambiguously defined, including
- Purification of the allergen protein to homogeneity.
- Determination of molecular weight, pI, and carbohydrate composition.
- Determination of nucleotide and/or amino acid sequence.
- Production of monospecific or monoclonal antibodies to the allergen.

The importance of the allergen in causing IgE responses should be defined by
- Comparing the prevalence of serum IgE antibodies in large population(s) of allergic patients. Ideally, at least 50 or more patients should be tested.
- Demonstrating biological activity, e.g., by skin testing or histamine release assay.
- Investigating whether depletion of the allergen from an allergic extract (e.g., by immunoabsorption) reduces IgE-binding activity.
- Demonstrating, where possible, that recombinant allergens have comparable IgE antibody-binding activity to the natural allergen.

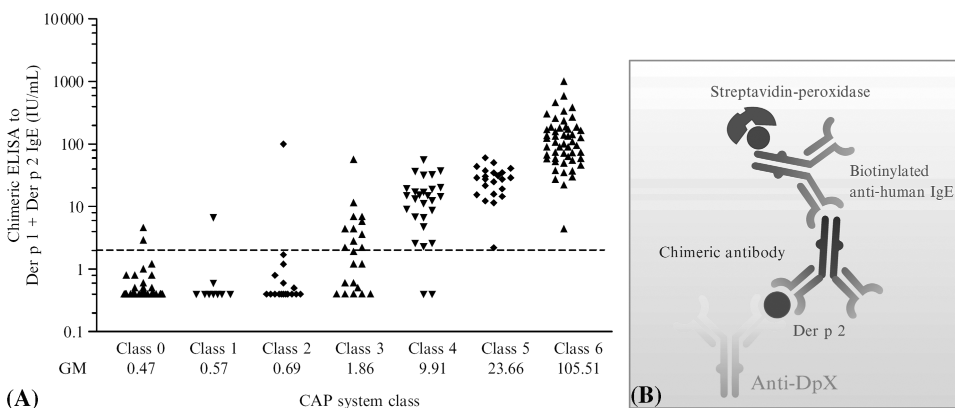

Figure 1 Chimeric ELISA for measuring allergen-specific IgE. (**A**) Schematic graphic of the ELISA. Microtiter plates are coated with monoclonal antibody followed by the relevant allergen and incubated with patient's serum. IgE antibodies that bind to the allergen complex are detected using biotinylated anti-IgE and streptavidin peroxidase. A chimeric IgE anti-Der p 2 is used to generate a control curve, and IgE values for patient's serum are interpolated from this curve. (**B**) Correlation between the chimeric ELISA for IgE antibody to Der p 1 and Der p 2 and FEIA (ImmunoCAP) for measuring IgE to house dust mite. There was an excellent quantitative correlation between the results for 212 sera from patients with asthma, wheezing, and/or rhinitis ($r = 0.86$, $p < 0.001$). *Source*: From Ref. 25.

50 or more sera should be screened, although allergens can be included in the nomenclature if the prevalence of IgE reactivity is >5% and if they elicit IgE responses in as few as five patients (Table 1) (20).

Several new methods for measuring IgE ab to specific allergens recently have been developed. "Chimeric" ELISA systems allow large numbers of sera to be screened for allergen-specific IgE ab by using a capture monoclonal antibody (mAb) to bind allergen. Serum IgE antibodies bind to the allergen complex and are detected with biotinylated anti-IgE (Fig. 1). The assay is quantified using a chimeric mouse anti-Der p 2 and human IgE epsilon antibody and provides results in ng/mL of allergen-specific IgE. Chimeric ELISA results for IgE ab to Der p 1, Der p 2, and Fel d 1 correlate with IgE measurements obtained by FEIA (25). A streptavidin-CAP assay using biotinylated allergens enables IgE antibodies to specific allergens to be routinely measured by FEIA (26). As with other diagnostic tests, chimeric ELISA and FEIA use separate tests to measure each IgE response, and these procedures use relatively large amounts of serum. Static or suspension microarray systems also have been developed that enable IgE antibodies to multiple allergens to be measured simultaneously. Microarrays provide a profile of IgE responses to specific allergens. One commercial test uses a static allergen array on allergen-coated glass slides to measure IgE antibodies in four sera to ~75 purified allergens at the same time. Results obtained with the microarray correlate with

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

FEIA using allergen extracts and the microarray uses only 30 μL serum (27–29). The sensitivity of the microarray is comparable to FEIA. Similarly, fluorescent multiplex suspension array technology has been developed in which allergens are covalently coupled to polystyrene microspheres containing different ratios of fluorescent dyes. Each microsphere bead can be distinguished by laser flow cytometry and forms a solid phase to which IgE antibodies bind and can be detected using biotinylated anti-IgE and streptavidin phycoerythrin. The fluorescent microarray currently measures total IgE and specific IgE to 10 purified allergens simultaneously using 20 μL serum (30). Array technologies are especially suited to large population surveys or birth cohorts for monitoring IgE responses to multiple allergens and for pediatric studies where serum is often in short supply.

Demonstrating that the allergen has biological activity in vivo is important, especially since many allergens are now produced as recombinant molecules before the natural allergen is purified (if ever). Several mite, cockroach, and fungal allergens (e.g., *Aspergillus, Alternaria, Cladosporium*) have been defined solely using recombinant proteins, and it is unlikely that much effort will be directed to isolating their natural counterparts. Ideally, the biological activity of recombinant proteins should be confirmed in vivo by quantitative skin testing or in vitro by histamine release assays. Skin testing studies were carried out using a number of recombinant allergens, including Bet v 1, Asp f 1, Bla g 4, Bla g 5, Der p 2, Der p 5, and Blo t 5 (22,31,32). These allergens showed potent biological activity and gave positive skin tests at the picogram level.

Resolving Ambiguities in Nomenclature

Early on it was recognized that because the system had Linnaean roots, some unrelated allergens would have the same name: *Candida* allergens could be confused with dog allergen (*Canis domesticus*); there are multiple related species of *Vespula* (Vespid) allergens; and *Periplaneta americana* (American cockroach) allergen needs to be distinguished from *Persea americana* (avocado)! These ambiguities have been overcome by adding a further letter to either the genus or species name. Examples thus become Cand a 1 (*C. albicans* allergen 1); Ves v 1 or Ves vi 1, to indicate *V. vulgaris* or *V. vidua* allergens, respectively; Per a 1 and Pers a 1 for the cockroach or avocado allergens. Dog allergen is referred to as Can f 1, from *Canis familiaris*. Many allergens have biochemical names that describe their biological function and which precede the allergen nomenclature. Examples include egg allergens (ovomucoid and ovalbumin); insect allergens (phospholipases and hyaluronidases); and tropomyosins from shrimp, mite, and cockroach. Sequence homology searches have assigned allergens to particular protein families and have provided important clues to their biological function. To some extent, allergens segregate among protein families that are according to whether they are indoor allergens, outdoor allergens, plant and animal food allergens, or injected allergens:

1. Indoor allergens (mite, animal allergens, cockroach, and molds): proteolytic enzymes (serine and cysteine proteases), lipocalins (ligand-binding proteins), tropomyosins, albumins, calcium-binding proteins, protease inhibitors (22,33)
2. Outdoor allergens (grass, tree and weed pollens, and mold spores): plant pathogenesis-related (PR-10) proteins, pectate lyases, β-expansins, calcium-binding proteins (polcalcins), defensin-like proteins, trypsin inhibitors (21,23,34,35)
3. Plant and animal food allergens (fruits, vegetables, nuts, milk, eggs, shellfish, and fish): lipid-transfer proteins, profilins, seed storage proteins, lactoglobulins, caseins, tropomyosins, parvalbumins (36–38)
4. Injected allergens (insect venoms and some therapeutic proteins): phospholipases, hyaluronidases, pathogenesis-related proteins, asparaginase (39,40)

Allergens belonging to these protein families are likely to have biological functions that are important to the host. Proteolytic enzymes are involved in digestion, tropomyosins and parvalbumins in muscle contraction, and profilins in actin polymerization in plants. The mouse lipocalin allergen, Mus m 1, is produced in the liver of male mice, secreted in large amounts in the urine and serves to mark the territories of male mice (41). The cockroach lipocalin allergen, Bla g 4, is produced in accessory glands of the male reproductive system and has an as yet unknown reproductive function (42,43) (Fig. 2). Crystallographic studies

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

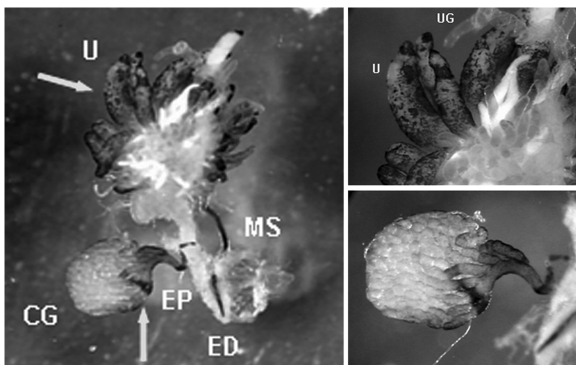

Figure 2 (*See color insert.*) Localization of German cockroach allergen Bla g 4 to the male reproductive tissues [conglobate glands (CG) and utricles (U)] by in situ hybridization (*left panel*). Right panel shows higher magnification. Bla g 4 is only found in male accessory reproductive glands and is transferred to the female during copulation. *Source*: From Ref. 42.

showed that Bet v 1, a plant pathogenesis–related (PR-10) protein, contained a hydrophobic pocket that could bind brassinosteroids and functions as a plant steroid carrier. The PR-10 proteins are important in plant defense, growth, and development (44).

In the allergy literature, it is preferable to use the systematic allergen nomenclature. However, in other contexts, such as comparisons of biochemical activities or protein structure, it may be appropriate or more useful to use the biochemical names. A selected list of the allergen nomenclature and biochemical names of inhalant, food, and venom allergens is shown in Table 2. There are now over 50 three-dimensional allergen structures in the Protein Database (PDB) and allergens are found in ~150 protein families in the Pfam protein family database (www.sanger.ac.uk/Software/Pfam). Breiteneder has argued that this is a relatively small number, given that almost 9000 protein families reside in Pfam (23,45). However, the 150 allergen protein families that have been identified still represent a huge degree of diversity at both the structural and biological level. Such diversity precludes any common structural feature, e.g., amino acid sequence motif or protein structure, which makes an allergen an allergen (21,24).

Isoallergens, Isoforms, and Variants

Originally, isoallergens were broadly defined as multiple molecular forms of the same allergen, sharing extensive antigenic (IgE) cross-reactivity. The revised nomenclature defines an isoallergen as an allergen from a single species, sharing similar molecular size, identical biological function, and $\geq$67% amino acid sequence identity (8). Some allergens, which were previously "grand fathered" into the nomenclature as separate entities, share extensive sequence homology and some antigenic cross-reactivity, but are named independently and are not considered to be isoallergens. Examples include Lol p 2 and Lol p 3 (65% homology) and Amb a 1 and Amb a 2 (65% homology). The word "Group" is now being used more often to describe structurally related allergens from different species within the same genus, or from closely related genera. In these cases, the levels of amino acid sequence identity can range from as little as 40% to ~90%. Similarities in tertiary structure and biological function are also taken into account when describing allergen Groups. Examples include the Group 2 mite allergens (Der p 2, Der f 2 and Lep d 2, Gly d 2 and Tyr p 2) showing 40% to 88% homology, and the Group 5 ragweed allergens (Amb a 5, Amb t 5, and Amb p 5) showing ~45% homology. The *Dermatophagoides* Group 2 allergen structures have been determined by X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR). The structures of the Group 2 allergens from other species were modeled on the *Dermatophagoides* structures (Fig. 3). This enabled the structural basis for antigenic relationships between members of the Group to be defined (20–22).

The term "variant" or "isoform" is used to indicate allergen sequences that show a limited number of amino acid substitutions (i.e., polymorphic variants of the same allergen). Typically, variants may be identified by sequencing several cDNA clones of a given allergen. Variants have been reported for Der p 1, Der p 2, Amb a 1, Cry j 1 and, for the most prolific Bet v 1, for which 42 sequences have been deposited in the GenBank database. Isoallergens and variants are denoted by the addition of four numeral suffixes to the allergen name. The first two numerals distinguish between isoallergens and the last two between variants. Thus, for

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

Table 2 Molecular Properties of Common Allergens

Source	Allergen	MW (kDa)	Homology/function
Inhalants			
Indoor			
House-dust mite (*Demlatophagoides pteronyssinus*)	Der p 1	25	Cysteine protease[b]
	Der p 2	14	Lipid binding protein
	Der p 3	30	Serine protease
	Der p 5	14	Unknown
Cat (*Felis domesticus*)	Fel d 1	36	Secretoglobin[b]
Dog (*Canis familiariss*)	Can f 1	25	Cysteine protease inhibitor?[b]
Mouse (*Mus muscularis*)	Mus m 1	21	Lipocalin (territory marking protein)
Rat (*Rattus norvegicus*)	Rat n 1	21	Pheromone-binding lipocalin[b]
Cockroach (*Blattella germanica*)	Bla g 2	36	Inactive aspartic protease
Outdoor			
Pollens—grassses			
Rye (*Lolium perenne*)	Lol p 1	28	Unknown
Timothy (*Phleum pratense*)	Phl p 5	32	Unknown
Bermuda (*Cynodon dactylon*)	Cyn d 1	32	Unknown
Weeds			
Ragweed (*Artemisia artemisiifolia*)	Amb a 1	38	Pectate lyase[b]
	Aruba 5	5	Neurophysins[b]
Trees			
Birch (*Betula verrucosa*)	Bet v 1	17	Pathogenesis-related protein[b]
Foods			
Milk	β-Lactolobulin	36	Retinol-binding protein[a,b]
Egg	Ovomucoid	29	Trypsin inhibitor
Codfish (*Gadus callarias*)	Gad c 1	12	Ca-binding protein (muscle parvalbumin)
Peanut (*Arachis hypogea*)	Ara h l	63	Vicilin (seed-storage protein)[b]
Venoms			
Bee (*Apis mellifera*)	Api m 1	19.5	Phospholipase A_2[b]
Wasp (*Polistes annularis*)	Pol a 5	23	Mammalian testis proteins
Homet (*Vespa crabro*)	Ves c 5	23	Mammalian testis proteins
Fire ant (*Solenopsis invicta*)	Sol i 2	13	Unknown
Fungi			
Aspergillus fumigatus	Asp f 1	18	Cytotoxin (mitogillin)
Alternaria altenata	Alt a 1	29	Unknown
Latex			
Hevea brasiliensis	Hev b 1	58	Elongation factor
	Hev b 5	16	Unknown—homologous to kiwi fruit protein of unknown function

[a]Most allergens have a single polypeptide chain; dimers are indicated.
[b]Allergens of known three-dimensional structure are also indicated.

ragweed Amb a 1, which occurs as four isoallergens, showing 12% to 24% differences in amino acid sequence, the nomenclature is as follows:

Allergen: Amb a 1
Isoallergens: Amb a 1.01; Amb a 1.02; Amb a 1.03; Amb a 1.04

Three variants of each isoallergen occur, showing >97% sequence homology

Isoforms: Amb a 1.0101, Amb a 1.0102, Amb a 1.0103
Amb a 1.0201, Amb a 1.0202, Amb a 1.0203, etc.

The Group 1 allergens from tree pollen have an unusually high number of isoallergens and variants. The 42 Bet v 1 sequences are derived from 31 isoallergens, which show from 73% to 98% sequence homology and are named from Bet v 1.0101 through Bet v 1.3101. The Group 1 allergen from hornbeam (*Carpinus betulus*), Car b 1, has three isoallergens that show 74% to

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

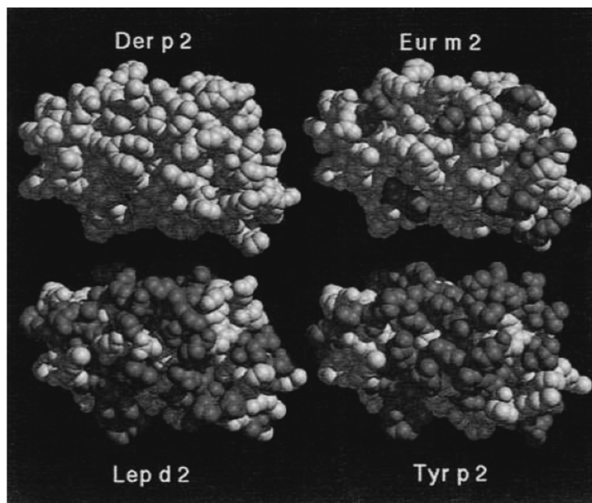

Figure 3 (*See color insert.*) Space-filling models of Group 2 allergens from house-dust mite. Amino acid substitutions are shown in gray scale. The space-filling model of Der p 2 (*Dermatophagoides pteronyssinus*) was generated from nuclear magnetic resonance spectroscopy studies and has subsequently been confirmed by X-ray crystallography (22). Eur m 2 (*Euroglyphus maynei*) shows 85% sequence identity with Der p 2 and seven of the substituted amino acids are shown in gray on the surface structure. There is extensive cross-reactivity between Der p 2 and Eur m 2. In contrast, Lep d 2 and Tyr p 2 show only 40% amino acid identity with the other Group 2 allergens. They show many substitutions on the antigenic surface of the molecules and show limited antigenic cross-reactivity for mAb and human IgE. *Source*: From Refs. 51, 52.

88% homology (Car b 1.01, 1.02, and 1.03) and the nomenclature committee's most recent records show 15 sequences of Car b 1. Ten variants of hazel pollen allergen, Cor a 1, have also been recorded. The reasons why the Group 1 tree pollen allergens have so many variants are unclear. Latex provides another example of distinctions in nomenclature. Hevein is an important latex allergen, designated Hev b 6, which occurs as a 20-kDa precursor with two fragments all derived from the same transcript. These moieties are all variants of Hev b 6 and are distinguished as Hev b 6.01 (prohevein, 20-kDa precursor), Hev b 6.02 (5-kDa hevein), and Hev b 6.03 (a 14-kDa C-terminal fragment). Studies have also uncovered a prodigious number of isoforms among mite Group 1 and Group 2 allergens. High fidelity PCR sequencing of environmental isolates of dust mites revealed 23 isoforms of Der p 1 and 13 isoforms of Der p 2) (24). Because isoforms differ in only a few amino acid substitutions, analysis of immunoreactivity to isoforms can be useful in defining antibody-binding sites and T-cell epitopes on allergens (46).

NOMENCLATURE FOR ALLERGEN GENES AND RECOMBINANT OR SYNTHETIC PEPTIDES

In the revised nomenclature, italicized letters are reserved to designate allergens genes. Two genomic allergen sequences have been determined from animal dander allergens: cat allergens, Fel d 1; and mouse urinary allergen, Mus m 1. Fel d 1 has two separate genes encoding chain 1 and chain 2 of the molecule, which are designed *Fel d 1A* and *Fel d 1B*, respectively (24). Genomic sequences of Bet v 1, Cor a 1, and apple "allergen," Mal d 1, have also been determined (47,48). Mal d 1 is an example of an incomplete or nonsensitizing allergen, i.e., an allergen that can interact with IgE antibodies but is unable to induce the production of IgE. Thus, symptoms of oral allergy syndrome in birch pollen allergic patients who eat apple are due to IgE cross-reactivity between Bet v 1 (the primary sensitizer) and Mal d 1 (with which the IgE anti-Bet v 1 interacts).

When recombinant allergens were first introduced, researchers often used the term "native allergen" to distinguish the natural protein from the recombinant allergen. However, because "native" has implications for protein structure (i.e., native conformation), it was decided that the term "natural allergen" should be used to indicate any allergen purified from natural source material. Natural allergens may be denoted by the prefix (n) to distinguish them from recombinant allergens, which are indicated by the prefix (r) before the allergen name (e.g., nBet v 1 and rBet v 1). There is no distinction between recombinant allergens produced in bacterial, yeast, or mammalian expression systems. Synthetic peptides are indicated by the prefix (s), with the particular peptide residues indicated in parentheses after the allergen name.

Thus, a synthetic peptide encompassing residues 100 to 120 of Bet v 1.0101 would be indicated: sBet v 1.0101 (100–120). At this point, the nomenclature, while technically sound, begins to become cumbersome and rather long-winded for most purposes. There are also additional refinements to the nomenclature that cover substitutions of different amino acid residues within synthetic peptides. This aspect of the nomenclature (which is based on that used for synthetic peptides of immunoglobulin sequences) is detailed in the revised nomenclature document to which aficionados are referred for full details (20).

THE IUIS SUBCOMMITTEE ON ALLERGEN NOMENCLATURE

Allergens to be considered for inclusion in the nomenclature are reviewed by the IUIS Subcommittee, which is currently chaired by Dr. Heimo Breiteneder (Medical University of Vienna, Austria) and has 19 members from all over the world (Table 3). The committee meets annually at an international allergy/immunology meeting and discusses new proposals it has received during the year, together with any proposed changes or additions to the nomenclature. There is also a committee-at-large, which is open to any scientist with an interest in allergens, to whom decisions made by the subcommittee are circulated. The procedure for submitting candidate names for allergens to the subcommittee is straightforward. Having purified the allergen and shown that it reacts with IgE ab, investigators should download the "new allergen name" form from the nomenclature committee web site (www.allergen.org) and send the completed form to the subcommittee prior to publishing articles describing the allergen. The subcommittee will provisionally accept the author's suggested allergen name, or assign the allergen a name, provided that the inclusion criteria are satisfied. The name will later be confirmed at a full meeting of the subcommittee. Occasionally, the subcommittee has to resolve differences between investigators who may be using different names for the same allergen, or disputes concerning the chronological order of allergen identification. These issues can normally be resolved by objective evaluation of each case.

Table 3 The World Health Organization and International Union of Immunological Societies Sub-Committee on Allergen Nomenclature, 2006–2008[a]

Name	Institution	Country
Heimo Breiteneder, PhD, Chairman	Medical University of Vienna	Vienna, Austria
Stefan Vieths, PhD, Secretary	Paul Ehrlich Institute	Langen, Germany
Wayne R Thomas, PhD, Past Chair	Western Australia Institute for Child Health	Perth, Australia
Naveen Arora, PhD	Institute of Genomics and Integrative Biology	Delhi, India
L. Karla Arruda, MD, PhD	School of Medicine of Ribeirão Preto, University of São Paulo	Ribeirão Preto, Brazil
Martin D Chapman, PhD	Indoor Biotechnologies, Inc.	Charlottesville, VA, USA
Fatima Ferreira, PhD	University of Salzburg	Salzburg, Austria
Richard Goodman, PhD	University of Nebraska	Lincoln, NE, USA
Donald Hoffman, PhD	East Carolina University	Greenville, NC, USA
Viswanath P Kurup, PhD	Medical College of Wisconsin	Milwaukee, WI, USA
Jørgen N Larsen, PhD	ALK-Abelló	Hørsholm, Denmark
Jonas Lidholm, PhD	Phadia AB	Uppsala, Sweden
Kåre Meno, PhD	ALK-Abelló	Hørsholm, Denmark
Andreas Nandy, PhD	Allergopharma Joachim Ganzer KG	Reinbek, Germany
Thomas AE Platts-Mills, MD, PhD	University of Virginia	Charlottesville, VA, USA
Christian Radauer, PhD	Medical University of Vienna	Vienna, Austria
Marianne van Hage MD, PhD	Karolinska Institute	Stockholm, Sweden
Ronald van Ree, PhD	Academic Medical Centre	Amsterdam, The Netherlands

[a]Past chairs of the committee: David G Marsh, PhD (1980–1989); Te Piao King, PhD (1990–1994); Henning Løwenstein, PhD (1994–1997); Wayne R. Thomas, PhD (1997–2006). Henning Løwenstein, PhD, DMSc (Hørsholm, Denmark) is an Emeritus member of the committee.

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

Table 4 Online Databases for Allergen Nomenclature and Structural Biology

Database	Locator
WHO/IUIS Allergen Nomenclature Sub-Committee	**www.allergen.org**[a]
Structural Database of Allergenic Proteins (SDAP)	http://fermi.utmb.edu/SDAP
Food Allergy Research and Resource Program (Farrp)	www.allergenonline.com
Protall	www.ifr.bbsrc.ac.uk/protall
ALLERbase	www.dadamo.com/allerbase
Allergome	www.allergome.org
Central Science Laboratory (York, UK)	http://www.csl.gov.uk/allergen/
AllFam	http://www.meduniwien.ac.at/allergens/allfam/

[a]Official website of the WHO/IUIS Sub-committee on allergen nomenclature.

Allergen Databases

The official web site for the WHO/IUIS Subcommittee on Allergen Nomenclature is www.allergen.org. This site lists all allergens and isoforms that are recognized by the Subcommittee and is updated on a regular basis. Over the past five years, several other allergen databases have been generated by academic institutions, research organizations, and industry sponsored groups (Table 4). These sites differ in their focus and emphasis, but are useful sources of information about allergens. The Structural Database of Allergenic Proteins (SDAP) was developed at the Sealy Center for Structural Biology, University of Texas Medical Branch, and provides detailed structural data on allergens in the WHO/IUIS nomenclature, including sequence information, PDB-files and programs to analyze IgE epitopes. Amino acid and nucleotide sequence information is also compiled in the SWISS-PROT and NCBI databases.

The Farrp and Protall databases have a focus on food allergens and provide sequence similarity searches (Farrp) and clinical data (skin tests and provocation tests) on food allergens (Protall). The Allergome database provides regular updates on allergens from publications in the scientific literature. Recently, a new database, AllFam, was developed that merges the Allergome allergens database with data on protein families from the Pfam database. Allfam contains all allergens with known sequences that can be assigned to at least one Pfam family. The database is maintained by Drs. Breiteneder and Radauer at the University of Vienna and can be accessed at: http://www.meduniwien.ac.at/allergens/allfam/.

CONCLUDING REMARKS

The three men in a boat did a remarkably good job! The use of the systematic allergen nomenclature has been extremely successful, has significantly enhanced research in the area, and continues to be revised and updated. The use of the generic terms "major" and "minor" allergen continues to evoke discussion. Relatively few allergens fulfill the criteria originally used by Marsh to define a major allergen (i.e., an allergen that causes IgE responses in $\geq$90% of allergic patients, such as Bet v 1, Fel d 1, Der p 2, Lol p 1) (10). However, there are a large number of allergens that cause sensitization in $>$50% of patients and Lowenstein et al. used this figure of 50% to define major allergens in the early 1980s (6). Scientists like to describe their allergen as "major" because it is effective in promoting their research and carries some weight in securing research funding. The question continues to be "What defines a major allergen"? Demonstrating a high prevalence of IgE-mediated sensitization and that the protein has allergenic activity is a minimal requirement, given the increasing sensitivity of assays to detect IgE antibodies. The contribution of an allergen to the total potency of the vaccine should be considered (e.g., by absorption studies), as well as the amount of IgE antibody directed against the allergen, compared with other allergens purified or cloned from the same source. Other criteria include whether the allergen induces strong T-cell responses and, for indoor allergens, whether it is a suitable marker of exposure in house-dust and air samples. The author, together with Dr. Rob Aalberse (University of Amsterdam), has developed eight criteria for "Allergens that make a difference."

It is clear from many studies that some allergens play a pre-eminent role in causing immune responses in atopic individuals; are better marker proteins for immunological, clinical, and epidemiological studies; and are usually considered to be high-profile targets for

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

Table 5 Eight Criteria for Defining Allergens That Make a Difference

1. A sensitization rate of >80% (>2 ng allergen specific IgE/mL) in a large panel of allergic patients
2. A significant proportion of total IgE (>10%) can be allergen specific
3. Absorption of the allergen from the source material significantly reduces the potency of the extract
4. Absorption of serum with purified allergen significantly reduces specific IgE to the allergen extract
5. The allergen accounts for a significant proportion of the extractable protein in the source material
6. The allergen can be used as a marker for environmental exposure assessment
7. Both antibody and cellular responses to the allergen can be measured in a high proportion of allergic patients
8. The allergen has been shown to be effective as part of an allergy vaccine

allergy diagnostics and therapeutics. Table 5 lists the eight criteria for defining the properties of these "allergens that make a difference." Examples of allergens that we consider to fulfill most of these criteria are as follows:

Mite	Group 1 and Group 2 (*Dermatophagoides* sp) allergens
Animal	Fel d 1, Mus m 1, Rat n 1
Tree pollen	Bet v 1 (and structurally homologous allergens); Ole e 1
Grass pollen	Phl p 1, Phl p 5
Weed pollen	Amb a 1
Peanut	Ara h 1, Ara h 2
Shellfish	Pen a 1 and other tropomyosins from shellfish
Insect allergens	Api m 1 (and homologous insect venom allergens)

Some of these recombinant allergens have already been shown to be effective as vaccines in clinical trials (Phl p 1 and Phl p 5), and most of the other allergens listed are being used as targets for vaccine development (4,22,49,50).

For most purposes, allergists need only be familiar with the nomenclature for allergens, rather than isoallergens, isoforms, peptides, etc. As measurements of allergens in diagnostics and vaccines, and in environmental exposure assessments become a routine part of the care of allergic patients, allergists will need to understand more about the structure and functions of allergens and how to distinguish them. Having a systematic nomenclature is an important part of this process. The systematic nomenclature is a proven success and is versatile enough to evolve with advances in molecular biology and proteomics that will occur over the next decade.

SALIENT POINTS

- A systematic nomenclature for all allergens that cause disease in humans has been formulated by a subcommittee of the WHO and the IUIS.
- Allergens are described using the first three letters of the genus, followed by a single letter for the species and an Arabic numeral to indicate the chronological order of allergen purification (e.g., *Dermatophagoides pteronyssinus* allergen 1 = Der p 1).
- To be included in the systematic nomenclature, allergens have to satisfy criteria of biochemical purity and criteria to establish their allergenic importance. It is important that the molecular structure of an allergen is defined without ambiguity and that allergenic activity is demonstrated in a large, unselected population of allergic patients.
- Modifications of the nomenclature are used to identify isoallergens, isoforms, allergen genes, recombinant allergens, and synthetic peptides. For example, Bet v 1.10 is an isoallergen of Bet v 1, and Bet v 1.0101 is an isoform or variant of the Bet v 1.10 isoallergen. Allergen genes are denoted by italics; e.g., *Fel d* 1A and *Fel d* 1B are the genes encoding chain 1 and chain 2 of Fel d 1, respectively.

This chapter has reviewed the systematic IUIS allergen nomenclature as revised in 1994. Other views expressed in the chapter are personal opinions and do not necessarily reflect the views of the WHO/IUIS Subcommittee on Allergen Nomenclature. The author is grateful to

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

Dr. Anna Pomés for assistance in preparing this chapter, and to Drs. Chad Gore and Coby Schal for permission to reproduce Figure 2.

REFERENCES

1. King TP, Norman PS. Isolation of allergens from ragweed pollen. Biochemistry 1962; 1:709–720.
2. Johnson P, Marsh DG. 'Isoallergens' from rye grass pollen. Nature 1965; 206(987):935–937.
3. Johnson P, Marsh DG. Allergens from common rye grass pollen (Lolium perenne). II. The allergenic determinants and carbohydrate moiety. Immunochemistry 1966; 3(2):101–110.
4. Marsh DG, Bias WB, Santilli J Jr., et al. Ragweed allergen Ra5: a new tool in understanding the genetics and immunochemistry of immune response in man. Immunochemistry 1975; 12(6–7):539–543.
5. Platts-Mills TA, Chapman MD, Marsh DG. Human immunoglobulin E and immunoglobulin G antibody responses to the "minor" ragweed allergen Ra3: correlation with skin tests and comparison with other allergens. J Allergy Clin Immunol 1981; 67(2):129–134.
6. Lowenstein H, King TP, Goodfriend L, et al. Antigens of Ambrosia elatior (short ragweed) pollen. II. Immunochemical identification of known antigens by quantitative immunoelectrophoresis. J Immunol 1981; 127(2):637–642.
7. Ohman JL Jr., Lowell FC, Bloch KJ. Allergens of mammalian origin. III. Properties of a major feline allergen. J Immunol 1974; 113(6):1668–1677.
8. Elsayed SM, Aas K. Characterization of a major allergen (cod.) chemical composition and immunological properties. Int Arch Allergy Appl Immunol 1970; 38(5):536–548.
9. Elsayed S, Bennich H. The primary structure of allergen M from cod. Scand J Immunol 1975; 4(2):203–208.
10. Marsh DG. Allergens and the genetics of allergy. In: Sela M III, ed. The Antigens. New York: Academic Press, 1975:271–350.
11. Lowenstein H. Quantitative immunoelectrophoretic methods as a tool for the analysis and isolation of allergens. Prog Allergy 1978; 25:1–62.
12. Lind P, Korsgaard J, Lowenstein H. Detection and quantitation of Dermatophagoides antigens in house dust by immunochemical techniques. Allergy 1979; 34(5):319–326.
13. Lowenstein H. Timothy pollen allergens. Allergy 1980; 35(3):188–191.
14. DeWeck A, Ring J. Collegium Internationale Allergologicum. History and Aims of a Special International Community Devoted to Allergy Research, 1954–1996. Munich, MMV Medezin Verlag 1996:1–120.
15. Marsh DG, Goodfriend L, King TP, et al. Allergen nomenclature. Bull World Health Organ 1986; 64 (5):767–774.
16. Scheiner O, Kraft D. Basic and practical aspects of recombinant allergens. Allergy 1995; 50(5):384–391.
17. Thomas WR. Mite allergens groups I-VII. A catalogue of enzymes. Clin Exp Allergy 1993; 23(5): 350–353.
18. King TP, Hoffman D, Lowenstein H, et al. Allergen nomenclature. WHO/IUIS Allergen Nomenclature Subcommittee. Int Arch Allergy Immunol 1994; 105(3):224–233.
19. King TP, Hoffman D, Lowenstein H, et al. Allergen nomenclature. Allergy 1995; 50(9):765–774.
20. King TP, Hoffman D, Lowenstein H, et al. Allergen Nomenclature. Bull World Health Org 1994; 72:797–800.
21. Aalberse RC. Structural biology of allergens. J Allergy Clin Immunol 2000; 106(2):228–238.
22. Chapman MD, Smith AM, Vailes LD, et al. Recombinant allergens for diagnosis and therapy of allergic disease. J Allergy Clin Immunol 2000; 106(3):409–418.
23. Radauer C, Breiteneder H. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. J Allergy Clin Immunol 2006; 117(1):141–147.
24. Chapman MD, Pomes A, Breiteneder H, et al. Nomenclature and structural biology of allergens. J Allergy Clin Immunol 2007; 119(2):414–420.
25. Trombone AP, Tobias KR, Ferriani VP, et al. Use of a chimeric ELISA to investigate immunoglobulin E antibody responses to Der p 1 and Der p 2 in mite-allergic patients with asthma, wheezing and/or rhinitis. Clin Exp Allergy 2002; 32(9):1323–1328.
26. Erwin EA, Custis NJ, Satinover SM, et al. Quantitative measurement of IgE antibodies to purified allergens using streptavidin linked to a high-capacity solid phase. J Allergy Clin Immunol 2005; 115 (5):1029–1035.
27. Harwanegg C, Laffer S, Hiller R, et al. Microarrayed recombinant allergens for diagnosis of allergy. Clin Exp Allergy 2003; 33(1):7–13.
28. Wohrl S, Vigl K, Zehetmayer S, et al. The performance of a component-based allergen-microarray in clinical practice. Allergy 2006; 61(5):633–639.
29. Hiller R, Laffer S, Harwanegg C, et al. Microarrayed allergen molecules: diagnostic gatekeepers for allergy treatment. FASEB J 2002; 16(3):414–416.
30. King EM, Vailes LD, Tsay A, et al. Simultaneous detection of total and allergen-specific IgE using purified allergens in a fluorescent multiplex array. J Allergy Clin Immunol 2007; 120(5):1126–1131.

[gajendra][7×10 Tight][D:/informa_Publishing/H1976_Lockey_112068/z_production/
z_3B2_3D_files/978-1-4200-6197-0_CH0003_O.3d] [10/7/08/11:46:7] [47–58]

31. Schmid-Grendelmeier P, Crameri R. Recombinant allergens for skin testing. Int Arch Allergy Immunol 2001; 125(2):96–111.
32. Pauli G, Purohit A, Oster JP, et al. Comparison of genetically engineered hypoallergenic rBet v 1 derivatives with rBet v 1 wild-type by skin prick and intradermal testing: results obtained in a French population. Clin Exp Allergy 2000; 30(8):1076–1084.
33. Arruda LK, Vailes LD, Ferriani VP, et al. Cockroach allergens and asthma. J Allergy Clin Immunol 2001; 107(3):419–428.
34. Radauer C, Willerroider M, Fuchs H, et al. Cross-reactive and species-specific immunoglobulin E epitopes of plant profilins: an experimental and structure-based analysis. Clin Exp Allergy 2006; 36(7): 920–929.
35. Gadermaier G, Dedic A, Obermeyer G, et al. Biology of weed pollen allergens. Curr Allergy Asthma Rep 2004; 4(5):391–400.
36. Breiteneder H, Radauer C. A classification of plant food allergens. J Allergy Clin Immunol 2004; 113(5): 821–830.
37. Reese G, Schicktanz S, Lauer I, et al. Structural, immunological and functional properties of natural recombinant Pen a 1, the major allergen of Brown Shrimp, Penaeus aztecus. Clin Exp Allergy 2006; 36(4):517–524.
38. Vieths S, Scheurer S, Ballmer-Weber B. Current understanding of cross-reactivity of food allergens and pollen. Ann N Y Acad Sci 2002; 964:47–68.
39. Henriksen A, King TP, Mirza O, et al. Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. Proteins 2001; 45(4):438–448.
40. King TP, Spangfort MD. Structure and biology of stinging insect venom allergens. Int Arch Allergy Immunol 2000; 123(2):99–106.
41. Hurst JL, Payne CE, Nevison CM, et al. Individual recognition in mice mediated by major urinary proteins. Nature 2001; 414(6864):631–634.
42. Fan Y, Gore JC, Redding KO, et al. Tissue localization and regulation by juvenile hormone of human allergen Bla g 4 from the German cockroach, *Blattella germanica* (L.). Insect Mol Biol 2005; 14(1):45–53.
43. Gore JC, Schal C. Cockroach allergen biology and mitigation in the indoor environment. Annu Rev Entomol 2007; 52:439–463.
44. Markovic-Housley Z, Degano M, Lamba D, et al. Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. J Mol Biol 2003; 325(1):123–133.
45. Jenkins JA, Griffiths-Jones S, Shewry PR, et al. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. J Allergy Clin Immunol 2005; 115(1): 163–170.
46. Piboonpocanun S, Malainual N, Jirapongsananuruk O, et al. Genetic polymorphisms of major house dust mite allergens. Clin Exp Allergy 2006; 36(4):510–516.
47. Schenk MF, Gilissen LJ, Esselink GD, et al. Seven different genes encode a diverse mixture of isoforms of Bet v 1, the major birch pollen allergen. BMC Genomics 2006; 7:168.
48. Gao ZS, van de Weg WE, Schaart JG, et al. Genomic cloning and linkage mapping of the Mal d 1 (PR-10) gene family in apple (*Malus domestica*). Theor Appl Genet 2005; 111(1):171–183.
49. Jutel M, Jaeger L, Suck R, et al. Allergen-specific immunotherapy with recombinant grass pollen allergens. J Allergy Clin Immunol 2005; 116(3):608–613.
50. Larche M, Akdis CA, Valenta R. Immunological mechanisms of allergen-specific immunotherapy. Nat Rev Immunol 2006; 6(10):761–771.
51. Smith AM, Benjamin DC, Hozic N, et al. The molecular basis of antigenic cross-reactivity between the group 2 mite allergens. J Allergy Clin Immunol 2001; 107(6):977–984.
52. Gafvelin G, Johansson E, Lundin A, et al. Cross-reactivity studies of a new group 2 allergen from the dust mite *Glycyphagus domesticus*, Gly d 2, and group 2 allergens from Dermatophagoides pteronyssinus, Lepidoglyphus destructor, and Tyrophagus putrescentiae with recombinant allergens. J Allergy Clin Immunol 2001; 107(3):511–518.