

Molecular Biology of Indoor Allergens

***Alisa M. Smith,* Anna Pomes,
and Martin D. Chapman***

*Asthma and Allergic Diseases Center, University of Virginia
School of Medicine, Charlottesville, VA 22908*

Introduction

The first aeroallergens to be identified were pollen grains, described in 1873 by Charles Blackley, who showed these particles were able to induce symptoms in patients with hayfever (1). It was nearly 100 years later that the outdoor allergens from ragweed and rye were purified and biochemically characterized (2,3). Significant changes in the lifestyle of Western society took place during the course of the 20th century, and people are now spending up to 95% of their time indoors. Epidemiologic studies have demonstrated a strong relationship between development of IgE antibodies to indoor allergens and asthma. Allergens from cats, dogs, house dust mites, cockroaches, and fungi consistently show the strongest association with bronchial hyperreactivity using a variety of study designs and different population groups (4,5).

The characterization of indoor aeroallergens began in the 1970s. Ohman and colleagues (6,7) characterized allergens from mammalian sources including cat, rat, mouse, guinea pig and rabbit. In 1980, the house dust mite allergen Der p 1, was purified from dust mite extract using conventional protein chemistry techniques and subsequently shown to be carried within the mite fecal particle (8,9). Additional dust mite allergens as well as allergens from other natural sources have since been identified including cockroach allergens Bla g 1 and Bla g 2, and fungal allergens from *Alternaria*, *Aspergillus*, and *Trichophyton* (10–13). The purification of important allergens led to the development of sensitive immunoassays for the quantitation of allergen levels in

*Author to whom all correspondence and reprint requests should be addressed. E-mail: ams3k@virginia.edu

environmental samples as well as for measuring allergen specific IgE from allergic patients (14). The introduction of molecular biology techniques over the past 15 years has provided a new approach to the identification and characterization of allergens. The aims of this article are to outline basic techniques for cloning allergen genes, the information obtained from cDNA sequence homology profiles, and the relevance of recombinant allergens to the diagnosis and treatment of allergic disease.

Allergen Cloning

Cloning involves the isolation of the genetic material that encodes a particular protein. Either DNA or messenger RNA (mRNA) can be used as the template for a recombinant allergen. In most cases mRNA is isolated from the natural source (for example, cockroach bodies or cat salivary gland), copied in vitro into complementary DNA (cDNA), and introduced into bacteriophage genomic DNA. The result is a cDNA expression library, containing DNA encoding proteins from the source organism. Allergens produced from bacteria infected by phage from such a library can be identified by immunoassay using sera from allergic subjects, or using polyclonal or monoclonal antibodies (MAb) raised to the natural allergen.

Bacteriophage containing allergen genes are isolated, and the nucleotide sequence encoding for the allergen is determined and translated into the primary amino-acid sequence of the protein. The correct reading frame is confirmed by comparison with any partial protein sequence from the purified natural allergen. Occasionally, several clones are obtained that correspond to different isoforms of the same allergen. Following the WHO/IUIS allergen nomenclature recommendations, the allergens are classified based on their degree of homology as isoallergens (when they share $\geq 67\%$ amino-acid sequence identity) or as variants (if the homology is close to 100%) (15).

By screening a cDNA library with sera from cockroach allergic patients, Bla g 4 and Bla g 5 from the German cockroach (*Blattella germanica*) were cloned in our laboratory (16,17). Using murine MAb raised to natural Bla g 1, the DNA encoding this allergen was isolated from the same library and Bla g 2 was detected using a high-titer polyclonal murine antisera also raised to the natural allergen (18,19). Per a 1, from a cDNA library prepared from the American cockroach (*Periplaneta americana*), was identified based on its antigenic crossreactivity with Bla g 1, that is, using IgE antibody in sera from patients allergic to the German cockroach (20).

Other approaches for cloning allergen genes have also been successful. These include polymerase chain reaction (PCR)-based techniques

that use synthetic oligonucleotides designed using knowledge of partial protein sequences obtained from purified natural allergens. This approach was used for the successful cloning of the major cat allergen Fel d 1 and for Eur m 1 and Eur m 2 from the house dust mite *Euroglyphus maynei* (21,22).

Protein Families and Function

Cloning techniques allow the study of allergens at the level of their primary, secondary, and tertiary structure. Once the primary structure of the allergen is known, homology data obtained by comparison with other protein sequences in data bases (such as GenBank), may provide clues about structure or biological function. Table 1 shows a classification of indoor allergens by function (F) or by homology (H) with proteins of known function. They are grouped mainly in four groups:

1. Enzymes;
2. Binding proteins;
3. Proteins involved in muscle contraction; and
4. Other homologies.

The table also lists allergens of unknown function or homology.

Enzymes

Many allergens show homology with proteases (23). Proteases are grouped into families based on their substrate specificity and their names usually reflect amino acids that are involved in the enzyme active site. Proteolytic enzymes are common components of digestion and are produced in the gastrointestinal tract in large excess. The Group 1 mite allergens are cysteine proteases that are produced in the digestive tract and excreted within the feces. The Group 3, 6, and 9 allergens are serine proteases with distinct substrate specificities similar to trypsin, chymotrypsin, and collagenase, respectively. Another subfamily within the serine protease family are the subtilisins and it has been shown that a *Trichophyton rubrum* allergen, Tri r 2, has sequence homology to these proteases (24). The cockroach allergen Bla g 2 has homology to the aspartic protease family of enzymes (19).

Functional enzymatic activity has been demonstrated for several of the mite allergens, including Der p 1 and Der f 3. Stewart and others (25–28) have proposed that proteolytic activity contributes to the allergenicity of mite proteins. An in vitro study has shown that Der p 1 and Der p 9 induce cytokine release from cultured epithelium (29). Other investigators have demonstrated that Der p 1 is able to cleave

Table 1
Structural and Functional Properties of Indoor Allergens

Source	Allergen	Molecular weight (kDa)	Sequence	F/H	Other functional characteristics
Enzymes					
Cysteine proteases					
<i>Dermatophagoides spp.</i>	Group 1	25	cDNA	F	Papain/actinidin
<i>Euroglyphus maynei</i>	Eur m1	25	PCR	H	
Serine proteases					
<i>Dermatophagoides spp.</i>	Group 3	~30	cDNA	F	Trypsin
<i>Dermatophagoides spp.</i>	Group 6	25	cDNA	F	Chymotrypsin
<i>Dermatophagoides spp.</i>	Group 9	~30	Protein	F	Collagenase
<i>Trichophyton rubrum</i>	Tri r 2	29	cDNA	H	Subtilisin
<i>Trichophyton spp.</i>	Group 4	83	Protein/cDNA	H	
Amylase					
<i>Dermatophagoides spp.</i>	Der p 4	~60	Protein	F	
<i>Aspergillus oryzae</i>	Asp o 2	52	Protein	H	
Aspartic protease					
<i>Blattella germanica</i>	Bla g 2	36	cDNA	H	
Glutathione transferase					
<i>Dermatophagoides spp.</i>	Der p 8	26	cDNA	H	
<i>Blattella germanica</i>	Bla g 5	36	cDNA	F	
Ribonuclease					
<i>Aspergillus spp.</i>	Asp f 1	18	PCR	F	Cytotoxin
Binding proteins					
Lipocalins					
<i>Canis familiaris</i>	Can f 1	21–25	cDNA	H	Protease inhibitor

<i>Canis familiaris</i>	Can f 2	19-27	cDNA	H	
<i>Mus musculus</i>	Mus m 1	19-21	cDNA	F	Pheromone binding
<i>Rattus norvegicus</i>	Rat n 1	19-21		F	Pheromone binding
<i>Blattella germanica</i>	Bla g 4	21	cDNA	H	
Muscle regulatory or structural proteins					
Tropomyosin					
<i>Dermatophagoides spp.</i>	Group 10	33-37	cDNA	H	
<i>Periplaneta americana</i>	Per a 7	33	cDNA	H	
Troponin C					
<i>Blattella germanica</i>	Bla g 6	18	cDNA	H	
<i>Blonia tropicalis</i>	bt8	18	cDNA	H	
Paramyosin					
<i>D. farinae</i>	Der f 11	98	cDNA	H	
Other homologies					
<i>Felis domesticus</i>	Fel d 1	36	PCR	H	Uteroglobin
<i>Dermatophagoides spp.</i>	Group 2	14	cDNA	H	Epididymal protein/moth molting protein
<i>Lepidoglyphus destructor</i>	Lep d 2	14	cDNA	H	
<i>Euroglyphus maynei</i>	Eur m 2	14	PCR	H	
<i>Tyrophagus putrescentiae</i>	Tyr p 2	14	PCR	H	
<i>Periplaneta americana</i>	Per a 3	72-78	cDNA	H	Hemocyanin
Unknown function/homology					
<i>D. pteronyssinus</i>	Der p 5	14	cDNA		
<i>D. pteronyssinus</i>	Der p 7	22-28	cDNA		
<i>Blonia tropicalis</i>	Blo t 5	14	cDNA		
<i>Blattella germanica</i>	Bla g 1	Variable	cDNA		
<i>Periplaneta americana</i>	Per a 1	Variable	cDNA		
<i>Alternaria alternata</i>	Alt a 1	29	cDNA		
<i>Cladosporium herbarium</i>	Cla h 1	13	cDNA		

H, predicted function based on sequence homology; F, confirmed function.

the low-affinity IgE receptor (CD23) from the surface of B-cells as well as the alpha subunit of the IL-2 receptor (CD25) (27,28).

The glutathione-S-transferases (GST) are a group of enzymes that catalyze the oxidation of glutathione; glutathione is an important antioxidant and GST is upregulated during the cell stress responses. Bla g 5 and Der p 8 are members of the GST family (17,30).

Amylase, an enzyme that hydrolyses starch, is also a common constituent of the digestive system of most organisms, and allergens belonging to this family of enzymes have been isolated from mite (Der p 4) and fungi (Asp o 2) (31,32). Another allergen from *Aspergillus* (Asp f 1) is homologous to the cytotoxic ribonuclease U₂ family of proteins that inhibit protein synthesis (12).

Binding Proteins

When the cockroach allergen Bla g 4 was cloned, a new family of allergens was described that includes mouse and rat urinary allergens, dog allergens Can f 1 and Can f 2, and others (16,33,34). These allergens are members of the lipocalin family of ligand-binding proteins. The lipocalins are extracellular proteins that bind hydrophobic molecules with high affinity and selectivity. They have a broad range of biological functions, from transport of serum proteins to modulation of the immune system. The proteins within this group have diverse sequences but retain three structurally conserved regions (SCR), with characteristic sequence motifs. These regions are also found in another group of binding proteins, the intracellular fatty-acid binding proteins (FABP). Flower and colleagues (35) have proposed the term calycin to describe a structural superfamily that includes the lipocalins and FABP.

Figure 1 shows the residues in common (bold) in the three SCR among seven different ligand-binding proteins. The lipocalin indoor allergens (Can f 1, Can f 2, and mouse urinary proteins [Mus m 1]) display this characteristic in common with other lipocalin allergens: the bovine allergen Bos d 2, the horse allergen Equ c 1 (36,37). A food allergen, the milk protein β -lactoglobulin, is also a member of the lipocalin family (38).

The ligands carried by the mouse and rat urinary proteins are pheromones that influence rodent behavior (39). The dog lipocalin Can f 1 is 50% identical to a human salivary protein, the von Ebner's gland protein (VEGh) (40). Can f 1 and VEGh are cysteine protease inhibitors, also known as cystatins. Recently, Ichikawa and colleagues (41) have reported the cDNA cloning of a cat allergen which shows homology to the cystatins. The cystatins are considered a subfamily of the lipocalins.

	17- - - - -32	102- - - - -119	129 - - - 139
	oo* o * *	oo *** *o	o *
Bla g 4	NLDYERFRGSWIIAAG	PYSVLATD Y ENYAIVEGC	VQIRFSVRRFH
Rat n 1	NLDVAKLNGDWFSIVV	TFTILK TDY DRYVMFHLI	FQLMVLYGR T K
Mus m 1	NFNVEKINGEWHTIIL	TFTIPK TDY DNFLMAHLI	FQLMG L YGR E P
Dog, Can f 1	GKDTVAVSGKWYLKAM	PSPVRDH-YILY-----C	-QIRMA-----
Dog, Can f 2	QGGLEELSGRWHSVAL	PKS----YLILYMINQYN	LMVR-DLSR Q Q
Bovine, Bos d 2	EIDPSKIPGEWRIIYA	SENMLVTYVENYDGERIT	-----TSFT
Equine, Equ c 1	NFDISKISGEWYSIFL	EDGVYSLNYDGYNVFRIS	FQLFEFYAREP
	-----SCR 1-----	-----SCR 2-----	---SCR 3---

Fig. 1. Structurally conserved regions (SCR) of the ligand-binding protein or lipocalin family of allergens. The SCR1 and SCR2 contain conserved residues, indicated in bold type (*), and conservative substitutions (o), that are common to the lipocalins. In addition, SCR3 contains a conserved contact residue (lysine or arginine) at position 137.

Proteins Involved in the Regulation of Contraction

Several allergens are proteins that are associated with the regulation of actin. The Group 10 allergens from *Dermatophagoides spp.* and a recently cloned allergen from the American cockroach *Periplaneta americana* are tropomyosins (42,43). The cockroach protein shares homology to other allergenic proteins described in arthropods, crustacea (shrimp), and arachnids (44). Allergens from *Blomia tropicalis* and *Blattella germanica* are troponins (45,46). Bla g 6 is homologous to troponin C. Troponin C is the calcium-binding subunit of the troponin complex, which also includes troponin I (actin-binding subunit) and troponin T (tropomyosin-binding subunit). Tropomyosin and the troponin complex inhibit the interaction of actin with myosin until the intracellular Ca²⁺ levels are boosted by a nerve impulse. Together, these protein constitute about one-third of the mass of a contractile thin filament.

A high molecular weight allergen from *Dermatophagoides farinae* has been identified as paramyosin, a structural protein of invertebrate muscle (47).

Other Homologies

The Group 2 mite allergens show 29 and 25% sequence identity to epididymal secretory proteins from human (HE1) and chimpanzee (EPI-1), respectively, and 25% sequence identity to the moth protein esr1 (48,49). Although the function of esr1 is not known, mRNA levels increase prior to molting of the pupal and larval stages, suggesting that this protein is involved in molting or ecdysis. Mueller and

colleagues (50) speculate that the Group 2 mite allergens may be involved in shedding of the exoskeleton. Based on the homology with the epididymal proteins, first described by Chua and Thomas (51), these allergens may have a role in mite reproduction.

The major cat allergen, Fel d 1, is composed of two polypeptide chains linked by disulfide bonds. Chain 1 is homologous to rabbit uteroglobulin; to date, no homologies have been identified for Chain 2 (21). Uteroglobulin is thought to have immunomodulatory activity based on its ability to bind steroids as well as inhibit phospholipase A2 activity. However, uteroglobulin is a homodimer and the Fel d 1 chains are distinctly different from each other, thus the significance of the homology to the function of Fel d 1 is unclear.

The crossreactive allergens Bla g 1 and Per a 1, show 30% homology to a mosquito (*Anopheles gambiae*) protein precursor ANG12, which is secreted only in the female insect after a blood meal (18). This suggests that these allergens may have a digestive function. Homology searches for Bla g 1 and Per a 1 in the GenBank revealed the existence of a 4058 nucleotide sequence from *B. germanica* described as Bla g Bd90K by Helm et al. (52), and four sequences from *P. americana* described by Wu et al. (53). These proteins, denominated the Group 1 cockroach allergens, share a singular primary structure: each protein is composed of tandem repeats of ~100 amino acids. Figure 2 illustrates a computer dot-plot matrix analysis that compares one sequence with another. When both are the same, the program plots a diagonal to indicate the identity of the sequence with itself. A parallel line to the diagonal shows that part of the sequence is repeated in another fragment of the molecule. Sequence analysis also identified a mitochondrial energy transfer protein domain in the primary structure of the Group 1 cockroach allergens, which suggests that they may have originated in this organelle. In addition, sequences of Group 1 allergens contain myristoylation sites suggesting possible post-translational modifications of the molecules. Immunohistochemical studies have identified Bla g 1 and Per a 1 in the digestive tract and fecal particles of cockroaches (18,20).

Allergens of Unknown Function or Homology

Although homology searches of the protein and nucleic acid data bases has yielded a wealth of information about the biochemical characteristics of allergens, there are still several cloned allergens that do not show homology to other proteins, including the Group 5 and 7 mite allergens, and fungal allergens from *Alternaria* and *Cladosporium* (54–58). In these cases, more studies are required to uncover the biologic function.

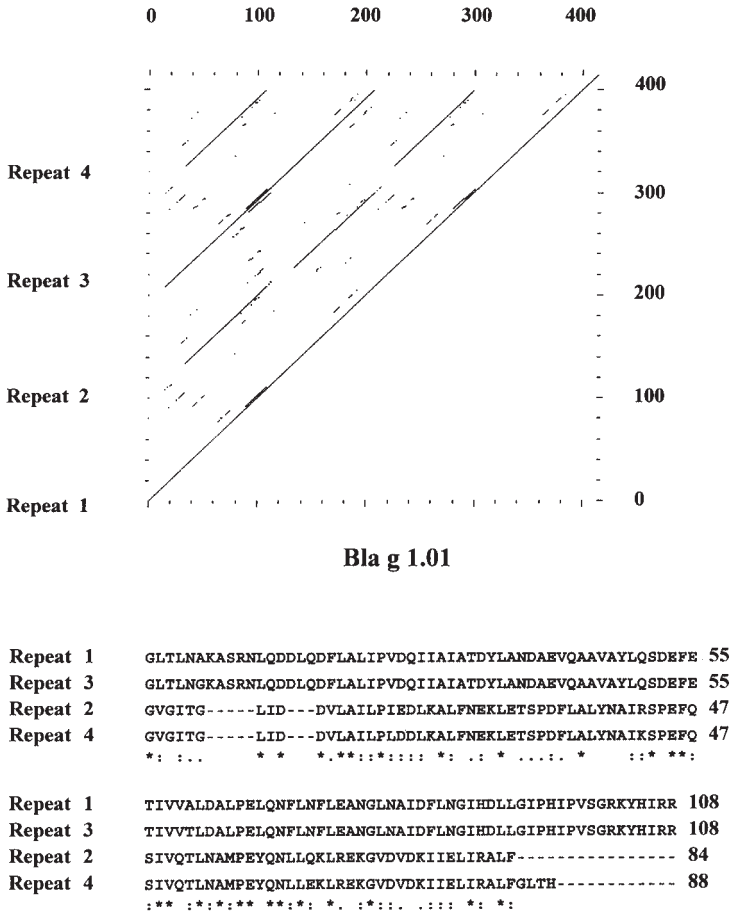


Fig. 2. Sequence analysis of the Group 1 cockroach allergens reveals a repeated domain of 100 amino acids. The top panel illustrates the dot-plot matrix analysis of Bla g 1.01 showing four amino-acid repeats as parallel lines. The bottom panel shows the amino acid sequence alignment of the four Bla g 1.01 repeats.

Allergen Structure

Determining the tertiary structure of allergens can provide clues about function as well as important information concerning the potential antigenicity of these proteins. For those allergens with homology to proteins of known structure, a three-dimensional model of the allergen can be generated. For example, Der p 1 has been modeled based on its homology with papain and actinidin (59). The X-ray crystal structures of these cysteine proteases have been determined; the proteins consist of two globular domains. One domain is composed of three alpha helices and the other domain is a beta sheet structure with two helices at opposite ends. The model of Der p 1 shows this

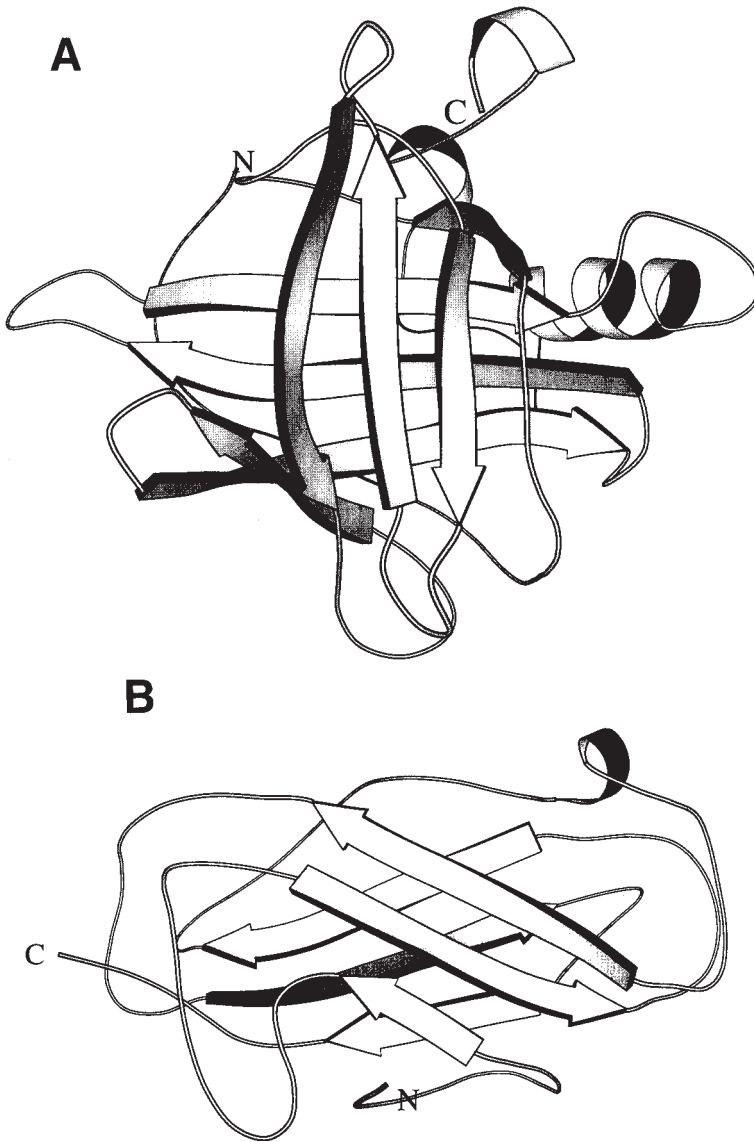


Fig. 3. Tertiary structure of Bla g 4 and Der p 2. The tertiary structure of Bla g 4 (A) was modeled on the structures of the bilin binding protein from the butterfly and the rat urinary protein, using SYBYL (Tripos Inc., St. Louis, MO) and CONGEN (Congenemix, Princeton, NJ) modeling programs. The structure of Der p 2 (panel B) was solved using nuclear magnetic resonance (NMR) spectroscopy.

arrangement of the domains as well as the active site cleft and the juxtaposition of the residues involved in proteolysis.

Bla g 4 has been modeled on the lipocalins bilin binding protein and rat urinary protein (Fig. 3A). The lipocalins have a characteristic cup shape formed from an eight strand antiparallel beta barrel (16).

Table 2
Expressed Recombinant Indoor Allergens

Allergen	Expression system	Vector	Yield (mg/L)	MW (kDa)
Cockroach				
Per a 1	<i>P. pastoris</i>	pPIC9	10	20
Bla g 2	<i>P. pastoris</i>	pPIC9	14	36
Bla g 4	<i>P. pastoris</i>	pPIC9	50	23
Bla g 5	<i>E. coli</i>	pET21a	3–4	23
Mite				
Der f 1	<i>Baculovirus</i>	pBlue Bac III	50	28
Der p 2	<i>E. coli</i>	pET	20	14
Der p 2	<i>S. cerevisiae</i>	pSAY 2, 3-p 2	8	14
Der p 5	<i>E. coli</i>	pGEX-2T	3	14
Blo t 5	<i>E. coli</i>	pGEX-4T1	3	14
Mammalian				
MUP	<i>P. pastoris</i>	pHIL-D2	270	18.7
Fel d 1*	<i>Baculovirus</i>	pAcSAG-LIC	1–2	50

*Chain 1 and chain 2 of Fel d 1 are joined by a 15aa linker and fused to the C-terminus of the H22 sFv (62).

For allergens of unknown function, protein structure can often provide clues by revealing structural motifs or domains. Recombinant allergens can be produced in large quantities (*see* Table 2) and purified to homogeneity. This makes it possible to perform X-ray crystallographic and spectroscopic studies to determine the tertiary structure. The structure of Der p 2 has been determined using nuclear magnetic resonance spectroscopy (NMR) (60) and X-ray crystallography (D. Benjamin and U. Derewenda, personal communication). These studies reveal that the protein is composed of two beta sheets arranged in an immunoglobulin-like fold (Fig. 3B). The relevance of this structure to the function of the Group 2 allergens is still unknown, but Ichikawa and colleagues (61) have suggested that Der f 2 may be involved in host defense. They speculate that Der f 2 may have antibacterial properties based on their study showing binding of the allergen to *Escherichia coli*.

Expression of Recombinant Allergens In Vitro

Cloning of allergen genes has been accompanied by the development of high-level expression systems for the production of recombinant allergens. In vitro systems for protein expression take advantage of the molecular machinery of a cell to produce large quantities of a certain protein encoded by a gene that is introduced into that cell.

Depending on the expression system, the host is bacteria (*E. coli*), yeast (*Pichia pastoris* or *Saccharomyces cerevisiae*), or insect cells. The choice of expression system may depend on characteristics of the allergen as well as the end use of the allergen (e.g., skin testing, which requires microgram amounts of allergen, or X-ray crystallography, which requires tens of milligrams). Some examples of expression systems and yields of important recombinant allergens are shown in Table 2.

E. coli is often the first choice as an expression host system based on the minimal nutrient requirements and rapid generation time of this organism. Indeed, most recombinant allergens have been expressed in bacteria. However, prokaryotic organisms lack the posttranslational modification machinery of higher organisms and are unable to glycosylate proteins or process proenzymes into their mature form. Yeast and insect cell systems can provide these processing and modification steps.

Most commercially available expression systems include features designed to aid in the purification of the recombinant protein. These may include plasmids that incorporate a protein or peptide tag in frame with the cloned allergen. These tags permit affinity chromatography, independent of the recombinant allergen, to be used for purification.

To provide useful reagents, recombinant allergens should show comparable immunoreactivity to their natural counterparts, when available. This is determined by in vitro immunoassays comparing the reactivity of the natural allergen with the recombinant, by histamine release assays, and in vivo, by quantitative skin testing. The rDer p 2 is recognized by MAb specific for the natural allergen, as well as with IgE from the majority of mite allergic subjects (50,63–65). The major allergen rDer p 1 has been the exception. Initially produced in *E. coli*, rDer p 1 showed poor reactivity with IgE antibody, suggesting that the recombinant allergen was incorrectly folded (66,67). IgE reactivity was improved when the allergen was expressed in *Saccharomyces*, but the protein yield was low (68). Shoji and colleagues (69,70), working with the Group 1 homolog Der f 1, have established an expression system in insect cells and have produced Der f 1 at 25 mg/L. This rDer f 1 gives overlapping IgE binding curves with the natural Der f 1 and also shows comparable protease activity. The key to this successful system appears to be in the correct processing of the proenzyme sequence to produce the mature protein.

Increasingly, recombinant allergens are identified before the natural allergen is known. For example, there are no purified allergens from *Blomia tropicalis*, a mite important in tropical and subtropical regions of the world. Arruda and colleagues (46) cloned a 14kDa allergen, designated Blo t 5, from a *Blomia* cDNA library. In this case the immunoreactivity of the recombinant allergen was established by

radioallergosorbent testing (RAST) using sera from patients with IgE antibody to *Blomia* extract. Skin testing was also used to demonstrate that rBlo t 5 was an important allergen. *Blomia* extracts are not commercially available and the production of recombinant allergens will make it possible to investigate the role of *Blomia* sensitization in asthma.

Clinical Relevance

Allergen cloning has become the method of choice for identifying and defining important allergens. The use of recombinant allergens has had an impact on several aspects of allergic disease as follows.

Standardization of Allergen Extracts

MAB-based immunoassays are now used for the standardization of dust-mite and cat extracts. The availability of recombinant allergens will aid in the development of immunoassays for standardizing other preparations such as cockroach extracts and fungal extracts (71). The goal is to provide a measurement, in mass units, of the major allergens in an extract. This will alleviate the variability of the biological units. These assays can also be applied to environmental samples for assessing allergen exposure and for assessing the effectiveness of allergen avoidance measures. In addition, recombinant allergens themselves provide reproducible standards and could potentially replace natural allergens for both diagnosis and treatment.

Allergen Crossreactivity

Allergen cloning and sequencing has helped to address the issue of allergen crossreactivity. The relationship between allergens can now be described in terms of primary amino acid sequence homology and families of allergens can be defined. For example, tropomyosins from mite and cockroach share homology with tropomyosins from shrimp and snails. van Ree and colleagues (72) have documented the development of oral allergy syndrome to shellfish in three patients receiving immunotherapy for dust mite hypersensitivity. In two of the three patients, tropomyosin was identified as the crossreacting allergen (72). Knowledge of the relationship between major allergens may help improve the diagnosis of specific allergen hypersensitivity.

Diagnosis

Most recombinant allergens show comparable immunoreactivity to their native counterparts. In vitro IgE assays and skin testing of allergic individuals with recombinant mite and cockroach allergens

has shown promising results. Arruda et al. (46) measured IgE antibody to recombinant Blo t 5 in 69% of sera from *Blomia* allergic children in Brazil. The recombinant allergen also showed good skin test reactivity in patients that were skin test positive to the natural *Blomia* extract. It may be possible to replace allergen extracts with one or two well defined recombinant allergens for diagnosis (71,73). This would be especially useful for allergens such as *Blomia*, where the availability of commercial extracts is limited.

Treatment

New treatment strategies have emerged as a result of cloning studies. The availability of the primary amino-acid sequence of a given allergen allows investigators to use synthetic peptides to determine T-cell epitopes. Peptide immunotherapy has gone through clinical trials for the treatment of cat allergic subjects and also for ragweed (74,75). Another strategy derived from recombinant allergens is the use of engineered hypoallergenic variants for immunotherapy. These variants have reduced IgE reactivity but retain T-cell epitopes; it is thought that these proteins may be more efficacious than conventional immunotherapy owing to decreased risk of side effects (76). Plasmid DNA vaccines and allergen-DNA conjugates are also being developed for the treatment of allergic disease. In a rat model, Hsu and colleagues (77) showed that a plasmid harboring the gene for the Der p 5 mite allergen could effectively prevent the development of IgE Ab and airway hyperresponsiveness when given intramuscularly to naïve animals. In addition, treatment of previously sensitized animals with the plasmid vaccine could reduce airway hyperresponsiveness upon subsequent exposure to the allergen (77). Raz and colleagues found similar results in a mouse model using the antigen β -galactosidase (78). This group characterized the DNA sequences, termed ISS for immunostimulatory sequences, outside the allergen gene regions that were responsible for an adjuvant effect that elicited IFN- γ and IL-12 (79). Subsequently, it has been shown that ISS can be administered with a protein antigen to achieve an antigen-specific Th1-type response (80,81). It is unclear how bacterial DNA sequences interact with cells of the immune system, and this approach to treatment of allergic disease is under active investigation.

In summary, the application of techniques of gene cloning and sequencing has led to the identification and characterization of a large number of indoor allergens. Research is now directed to using this information to address the immune mechanisms underlying allergic disease, to develop sensitive tests for diagnosis, to evaluate environmental exposure, and to develop new forms of treatment.

Acknowledgments

The studies in our laboratory are supported by National Institutes of Health Grants AI 34607 and AI 32557 and a grant from the American Lung Association (A.M.S.). We thank Dr. David Benjamin for preparing the graphic representations of allergen structures (Fig. 3) and Ms. Wanda Harvey for assistance with the preparation of the manuscript.

References

1. Blackley CH. (1873), Experimental Researches on the Causes and Nature of Cattharus Aestivus (Hay-Fever or Hay-Asthma). Bailliere, Tindal and Cox, London; (reprint: Dawson's of Pall Mall, 1959.)
2. Johnson P, Marsh DG. (1965), Isoallergens from ryegrass pollen. *Nature* 206: 935-937.
3. King TP, Norman PS, Lichtenstein LM. (1967), Isolation and characterization of allergens from ragweed pollen IV. *Biochemistry* 6:1992-2000.
4. Pope AM, Patterson R, Burge H (eds.). Indoor allergens: assessing and controlling adverse health effects. Washington DC: National Academy Press, 1993.
5. Custovic A, Chapman MD. Indoor allergens as a risk factor for asthma. In: Barnes PJ, Grunstein MM, Leff A, Woolcock AJ (eds.). *Asthma*. Philadelphia: Lippincott-Raven, 1997; pp. 1-21.
6. Ohman JL, Lowel FC, Bloch KJ. (1974), Allergens of mammalian origin III. Properties of a major feline allergen. *J Immunol* 113:1668-1677.
7. Ohman JL, Lowel FC, Bloch KJ. (1975), Allergens of mammalian origin II. Characterization of allergens extracted from rat, mouse, guinea pig, and rabbit pelts. *J Allergy Clin Immunol* 55:16-24.
8. Chapman MD, Platts-Mills TAE. (1985), Purification and characterization of the major allergen from *Dermatophagoides pteronyssinus*. *J Aller Clin Immunol* 76:753-761.
9. Tovey ER, Chapman MD, Platts-Mills TAE. (1981), Mite fecaes are a major source of house dust mite allergens. *Nature* 289:592, 593.
10. Pollart SM, Mullins DE, Vailes LD, Sutherland WM, Chapman MD. (1991), Identification, quantitation and purification of cockroach allergens using monoclonal antibodies. *J Aller Clin Immunol* 87:511-521.
11. Yunginger JW, Jones RT, Nesheim ME, Geller M. (1980), Studies on *Alternaria* allergens. III. Isolation of a major allergenic fraction (ALT-I). *J Allergy Clin Immunol* 66:138-147.
12. Arruda LK, Platts-Mills TAE, Fox JW, Chapman MD. (1990), *Aspergillus fumigatus* allergen I, a major IgE binding protein, is a member of the mitogillin family of cytotoxins. *J Exp Med* 172:1529-1534.
13. Deuell BL, Arruda LK, Hayden ML, Chapman MD, Platts-Mills TAE. (1991), *Trichophyton tonsurans* allergen I: characterization of a protein that causes immediate but not delayed hypersensitivity. *J Immunol* 147:96-101.
14. Chapman MD, Smith AM, Slunt JB, Vailes LD, Arruda LK. (1995), Immunochemical methods for defining and measuring indoor allergens: in dust and in the air. *Paed Allergy Clin Immunol* 6 (suppl. 7):8-12.
15. King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TAE, Thomas WR. (1994), Allergen nomenclature. *Intl Arch Allergy Immunol* 105:224-233.
16. Arruda LK, Vailes LD, Hayden ML, Benjamin DC, Chapman MD. (1995), Molecular cloning of cockroach allergen, Bla g 4, identifies ligand binding

- proteins (or calycons) as a cause of IgE antibody responses. *J Biol Chem* 270:31,196–31,201.
17. Arruda LK, Vailes LD, Platts-Mills TAE, Hayden ML, Chapman MD. (1997), Induction of IgE antibody responses by glutathione-S-transferase from the German cockroach (*Blattella germanica*). *J Biol Chem* 272:20,907–20,912.
 18. Pomes A, Melen E, Vailes LD, Retief JD, Arruda LK, Chapman MD. (1998), Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. *J Biol Chem* 273:30,801–30,807.
 19. Arruda LK, Vailes LD, Mann BJ, Shannon J, Fox JW, Vedvick TS, et al. (1995), Molecular cloning of a major cockroach (*Blattella germanica*) allergen, Bla g 2: sequence homology to the aspartic proteases. *J Biol Chem* 270:19,563–19,568.
 20. Melen E, Pomes A, Vailes LD, Arruda LK, Chapman MD. (1999), Molecular cloning of Per a 1 and definition of the cross-reactive group 1 cockroach allergens. *J Allergy Clin Immunol* 103:859–864.
 21. Morgenstern JP, Griffith IJ, Brauer AW, Rogers BL, Bond JF, Chapman MD, and Kuo MC. (1991), Determination of the amino acid sequence of Fel d 1, the major allergen of the domestic cat. protein sequence analysis and cDNA cloning. *Proc Natl Acad Sci USA* 88:9690–9694.
 22. Smith W, Mills KL, Hazell LA, Hart BJ, Thomas WR. (1999), Molecular analysis of the groups 1 and 2 allergens from the house dust mite, *Euroglyphus maynei*. *Int Arch Allergy Immunol* 118:15–22.
 23. Stewart GA, Thompson PJ. (1996), The biochemistry of common aeroallergens. *Clin Exp Allergy* 26:1020–1044.
 24. Woodfolk JA, Wheatley LM, Piyasena RV, Benjamin DC, Platts-Mills TAE. (1998), *Trichophyton* antigens associated with IgE antibodies and delayed type hypersensitivity. *J Biol Chem* 273:29,489–29,496.
 25. Stewart GA, Thompson PJ, McWilliam AS. (1993), Biochemical properties of aeroallergens: contributory factors in allergic sensitization. *Pediatr Allergy Immunol* 4:163–172.
 26. Shakib F, Schulz O, Sewell H. (1998), A mite subversive: cleavage of CD23 and CD25 by Der p 1 enhances allergenicity. *Immunol Today* 19:313–316.
 27. Schulz O, Laing P, Sewell HF, Shakib F. (1995), Der p 1, a major allergen of the house dust mite, proteolytically cleaves the low affinity receptor for human IgE (CD23). *Eur J Immunol* 25:3191–3194.
 28. Hewitt CRA, Brown AP, Hart BJ, Pritchard DI. (1995), A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: innate protection by antiproteases. *J Exp Med* 182:1537–1544.
 29. King C, Brennan S, Thompson PJ, Stewart GA. (1998), Dust mite proteolytic allergens induce cytokine release from cultured airway epithelium. *J Immunol* 161: 3645–3651.
 30. O'Neill G, Donovan GR, Baldo BA. (1994), Identification of a major allergen of the house dust mite *Dermatophagoides pteronyssinus*, homologous with glutathione-S-transferase. *Biochim Biophys Acts* 1219:521–524.
 31. Lake FR, Ward LD, Simpson RJ, Thompson PJ, Stewart GA. (1991), House dust mite derived amylase: allergenicity and physicochemical characterization. *J Aller Clin Immunol* 81:1035–1042.
 32. Baur X, Fruhman G, Haug B, Rasche B, Reiher W, Weiss W. (1986), Role of *Aspergillus* amylase in baker's asthma. *Lancet* 1:43.
 33. Konecny A, Morgenstern JP, Bizinkauskas CB, Lilley CH, Brauer AW, Bond JF, et al. (1997), The major dog allergens, Can f 1 and Can f 2, are salivary lipocalin proteins: cloning and immunological characterization of the recombinant forms. *Immunol* 92:577–586.
 34. Longbottom JL. Purification and characterization of allergens from the urines of mice and rats. In: Oehling A, Glazer I, Mathov E, Arbesman E (eds.) *Advances in Allergology and Applied Immunology*. Oxford: Pergamon, 1980, pp. 483–490.

35. Flower DR, North ACT, Attwood TK. (1993), Structure and sequence relationships in the lipocalins and related proteins. *Protein Sci* 2:753–761.
36. Mantyjarvi R, Parkkinen S, Rytönen M, Penttinen J, Pelkonen J, Rautiainen J, et al. (1996), Complementary DNA cloning of the predominant allergen of bovine dander: a new member in the lipocalin family. *J Allergy Clin Immunol* 97:1297–1303.
37. Gregoire C, Rosinski-Chupin I, Robillon J, Alzari PM, David B, Dandeu J-P. (1996), cDNA cloning and sequencing reveal the major horse allergen Equ c 1 to be a glycoprotein member of the lipocalin superfamily. *J Biol Chem* 271:32,951–32,959.
38. Papiz MZ, Sawyer L, Eliopoulos EE, North ACT, Findlay JBC, Sivaprasadarao R, et al. (1986), The structure of beta-lactoglobulin and its similarity to plasma retinol-binding protein. *Nature* 324:383–385.
39. Bocskei Z, Groom CR, Flower DR, Wright CE, Phillips SEV, Cavaggioni A, et al. (1992), Pheromone binding to two rodent urinary proteins revealed by x-ray crystallography. *Nature* 360:186–188.
40. van't Hoff W, Blankenvoorde MFJ, Veerman ECI, Nieuw Amerongen AV. (1997), The salivary lipocalin Von Ebner's gland protein is a cysteine proteinase inhibitor. *J Biol Chem* 272:1837–1841.
41. Chapman MD, Ichikawa K, Vailes LD. (1999), Molecular cloning of a new cat allergen, cystatin, a cysteine protease inhibitor. *J Allergy Clin Immunol* 103(1) Pt. 2 (abstract #466), p. S122.
42. Aki T, Kodama T, Fujikawa A, Miura K, Shigeta S, Wada T, et al. (1995), Immunochemical characterization of recombinant and native tropomyosins as a new allergen from the house dust mite, *D. farinae*. *J Allergy Clin Immunol* 96:74–83.
43. Asturias JA, Arilla MC, Gomez-Bayon N, Martinez A, Martinez J, Palacios R. (1998), Sequencing and high level expression in *Escherichia coli* of the tropomyosin allergen (Der p 10) from *Dermatophagoides pteronyssinus*. *Biochim Biophys Acta* 1397:27–30.
44. Santos ABR, Tobias KR, Ferriani VPL, Rizzo MC, Nasipitz CK, Pomes A, et al. (1999), Identification of tropomyosin from *Periplaneta americana* as a major cockroach allergen. *J Allergy Clin Immunol* 104:329–337.
45. Witteman AM, Akkerdaas J, van Leeuwen J, van der Zee JS, Aalberse RC. (1994), Identification of a cross-reactive allergen (presumably tropomyosin) in shrimp, mite and insects. *Int Arch Allergy Immunol* 105:56–61.
46. Arruda LKA, Vailes LD, Platts-Mills TAE, Fernandez-Caldas E, Montealegre F, Lin KL, et al. (1997), Sensitization to *Blomia tropicalis* in patients with asthma and identification of allergen Blo t 5. *Am J Respir Crit Care Med* 155:343–350.
47. Arruda LK, Vailes LD, Benjamin DC, Chapman MD. (1995), Molecular cloning of German cockroach (*Blattella germanica*) allergens. *Int Arch Allergy Immunol* 107:295–297.
48. Tsai LC, Chao PL, Shen HD, Tang RB, Chang TC, Chang ZN, et al. (1998), Isolation and characterization of a novel 98kda *Dermatophagoides farinae* allergen. *J Allergy Clin Immunol* 102:295–303.
49. Thomas WR, Smith W. (1998), Allergy Review Series II. An update on allergens: house dust mite allergens. *Allergy* 53:821–832.
50. Mueller GA, Smith AM, Williams DC, Hakkaart GAJ, Aalberse RC, Chapman MD, et al. (1997), Expression and secondary structure determination by NMR methods of the major house dust mite allergen Der p 2. *J Biol Chem* 272:26,893–26,898.
51. Thomas WR, Chua KY. (1995), The major mite allergen Der p 2- a secretion of the male mite reproductive tract. *Clin Exp Allergy* 25:667–669.
52. Helm R, Cockrell G, Stanley JS, Brenner RJ, Burks W, Bannon GA. (1996), Isolation and characterization of a clone (Bla g Bd90K) encoding a major allergen involved in IgE mediated immediate hypersensitivity. *J Allergy Clin Immunol* 98:172–180.

53. Wu CH, Lee MF, Liao SC, Luo SF. (1996), Sequencing analysis of cDNA clones encoding the American cockroach Cr-PI allergens. *J Biol Chem* 30:17,937–17,943.
54. Tovey ER, Johnson MC, Roche AL, Cobon GS, Baldo BA. (1989), Cloning and sequencing of a cDNA expressing a recombinant house dust mite protein that binds human IgE and corresponds to an important low molecular weight allergen. *J Exp Med* 170:1457–1462.
55. Lin KL, Chua KY, Thomas WR, Chiang BL, Hsieh KH. (1994), Characterization of Der p V allergen. cDNA analysis and IgE mediated reactivity to the recombinant protein. *J Allergy Clin Immunol* 94:989–996.
56. Shen H-D, Chua KY, Lin WL, Hsieh KH, Thomas WR. (1995), Molecular cloning and immunological characterization of the house dust mite allergen Der f 7. *Clin Exp Allergy* 25:1000–1006.
57. Shen HD, Chua KY, Lin NL, Hsieh KH, Thomas WR. (1995), Characterization of the house dust mite allergen Der p 7 by monoclonal antibodies. *Clin Exp Allergy* 25:416–422.
58. Achatz G, Oberkofler H, Lechenauer E, Simon B, Unger A, Kandler D, et al. (1995), Molecular cloning of major and minor allergens of *Alternaria alternata* and *Cladosporium herbarum*. *Mol Immunol* 32:213–227.
59. Topham CM, Srinivasan N, Thorpe CJ, Overington H, Kalsheker NA. (1994), Comparative modeling of major house dust mite allergen Der p I: structure validation using an extended environmental amino acid propensity table. *Protein Eng* 7:869–894.
60. Mueller GA, Benjamin DC, Rule GS. (1998), Tertiary structure of the major house dust mite allergen Der p 2: Sequential and structural homologies. *Biochemistry* 34:12,707–12,714.
61. Ichikawa S, Hatanaka H, Yuuki T, Iwamoto N, Kojima S, Nishiyama C, et al. (1998), Solution structure of Der f 2, the major allergen for atopic diseases. *J Biol Chem* 273:356–360.
62. Liu C, Goldstein J, Graziano RF, He J, O'Shea JK, Deo Y, et al. (1996), F(c) gamma RI-targeted fusion proteins result in efficient presentation by human monocytes of antigenic and antagonistic T-cell epitopes. *J Clin Invest* 98:2001–2007.
63. Chua KY, Dilworth RJ, Thomas WR. (1990), Expression of *Dermatophagoides pteronyssinus* allergen Der p II in *Escherichia coli* and binding studies with human IgE. *Int Arch Allergy Appl Immunol* 91:124–129.
64. Smith AM, Chapman MD. (1996), Reduction in IgE binding to allergen variants generated by site-directed mutagenesis: contribution of disulfide bonds to the antigenic structure of the major house dust mite allergen, Der p 2. *Mol Immunol* 33:399–405.
65. Hakkaart GAJ, Harmsen MM, Chua KY, Thomas WR, Aalberse RC, van Ree R. Expression of the house dust mite allergen Der p 2 in the baker's yeast *Saccharomyces cerevisiae*. *Clin Exp Allergy* 28:45–52.
66. Chua KY, Stewart GA, Thomas WR, Simpson RJ, Dilworth RJ, Plozza TM, Turner KJ. (1988), Sequence analysis of cDNA coding for a major house dust mite allergen, Der p 1. Homology with cysteine proteases. *J Exp Med* 167:175–182.
67. Greene WK, Chua KY, Stewart GA, Thomas WR. (1990), Antigenic analysis of group I house dust mite allergens using random fragments of Der p I expressed by recombinant DNA libraries. *Int Arch Allergy Immunol* 92:30–38.
68. Chua KY, Kehal PK, Thomas WR, Vaughan PR, Macreadie IG. (1992), High frequency of IgE binding to the Der p 1 allergen expressed in yeast. *J Allergy Clin Immunol* 89:95–102.
69. Shoji H, Hanawa M, Shibuya I, Hirae M, Yasuhara T, Okumura Y, Yamakawa H. (1996), Production of recombinant mite allergen Der f 1 in insect cells and characterization of products: removal of pro-sequence is essential to IgE binding activity. *Biosci Biotech Biochem* 60:621–625.

70. Shoji H, Shibuya I, Hirai M, Horiuchi H, Takagi M. (1997), Production of recombinant Der f 1 with the native IgE-binding activity using a baculovirus expression system. *Biosci Biotech Biochem* 61:1668–1673.
71. Chapman MD, Smith AM, Vailes LD, Arruda LK. (1997), Recombinant mite allergens: new technologies for the management of patient with asthma. *Allergy* 52:374–379.
72. van Ree R, Antonicelli L, Akkerdaas JH, Garritani MS, Aalberse RC, Bonifazi F. (1996), Possible induction of food allergy during mite immunotherapy. *Allergy* 51:108–1113.
73. Van Ree R, van Leeuwen WA, Aalberse RC. (1998), How far can we simplify in vitro diagnosis for grass pollen allergy? A study with 17 whole pollen extracts and purified natural and recombinant major allergens. *J Aller Clin Immunol* 102:184–190.
74. Norman PS, Ohman JL, Long AA, Creticos RS, Gefter MA, Shaked Z, et al. (1996), Treatment of cat allergy with T cell reactive peptides. *Am J Respir Crit Care Med* 154:1623–1628.
75. Creticos PS, Hebert J, Philip G, Allervax Ragweed Study Group. (1997), Efficacy of Allervax ragweed peptides in the treatment of ragweed induced allergy. *J Allergy Clin Immunol* 99(1 p 2): abstr 1631.
76. Smith AM, Chapman MD. Allergen specific immunotherapy: new strategies using recombinant allergens. In Bousquet J, Yssel H (eds.) *Immunotherapy for Asthma*. New York: Marcel Dekker 1999; pp. 99–118.
77. Hsu CH, Chua KY, Tao MH, Lai YL, Wu HD, Huang SK, Hsieh KH. (1996), Immunoprophylaxis of allergen induced immunoglobulin E synthesis and airway hyperresponsiveness in vivo by genetic immunization. *Nature Med* 2:540–544.
78. Raz E, Tighe H, Sato Y, Corr M, Dudler JA, Roman M, et al. (1996), Preferential induction of a Th1 immune response and inhibition of specific IgE antibody formation by plasmid DNA immunization. *Proc Natl Acad Sci USA* 93:5141–5145.
79. Sato Y, Roman M, Tighe H, Lee D, Corr M, Nguyen MD, et al. (1996), Immunostimulatory DNA sequences necessary for effective intradermal gene immunization. *Science* 273:352–354.
80. Roman M, Martin-Orozco E, Goodman JS, Nguyen M-D, Sato Y, Ronaghy A, et al. (1997), Immunostimulatory DNA sequences function as T helper-1-promoting adjuvants. *Nature Med* 3:849–854.
81. Broide D, Schwarze J, Tighe H, Gifford T, Nguyen MD, Malek S, et al. (1998), Immunostimulatory DNA sequences inhibit IL-5, eosinophilic inflammation, and airway hyperresponsiveness in mice. *J Immunol* 161:7054–7062.