The advent of recombinant allergens and allergen cloning

Wayne R. Thomas, PhD *Subiaco, Australia*

When the allergen nomenclature system was adopted in 1986, allergens were identified by their behavior on electrophoresis and chromatography and by reactivity to shared antisera. Not only was this unsatisfactory for standardization, but the processes of allergic sensitization and immunotherapy could not be studied in the framework of antigen processing and B- and T-cell epitopes. Recombinant technologies developed in the 1980s for cloning cDNA from low-abundance mRNA permitted the cloning of allergens, beginning with the major house dust mite allergen Der p 1 and hornet allergen Dol m 5. After this, a wave of cloning with IgE immunoscreening resulted in the cloning of Der p 2, Der p 5, Bet v 1, Bet v 2, and Dac g 2 along with Fel d 1 cloned after amino acid sequencing. Recombinant allergens have now been used to define the important allergens for a wide range of allergies and to develop new types of immunotherapy, some of which have shown efficacy in human trials. The clonally pure allergens have been used to solve the tertiary structures of allergens and from this how allergens might activate innate immunity. Proprietary recombinant allergens are now being used in improved diagnostic tests. (J Allergy Clin Immunol 2011;127:855-9.)

Key words: *Allergen, cloning, recombinant, history,* Dermatophagoides *species, allergen structure*

The immunochemical characterization of allergens was begun at a time when it was inconceivable that practical quantities of most common allergens could be purified to homogeneity. Indeed, it was a concern that the low doses of allergen required to elicit hypersensitivity reactions were less than those that could be monitored for purity. However, it was established that more than 1 allergen from a particular source was responsible for hypersensitivity and that each source had major and minor allergens. The importance of proteins was identified, which was an important question at the time, and it was established that major allergens, such as Amb a 1 from ragweed, could be a heterogeneous family of proteins.[1]

Amb a 1 was the vanguard of the immunochemical characterization of inhalant allergens, especially from the studies of Te Piao King.[1] Its biochemical identity and amino acid sequence, however, remained unknown because its blocked N-terminal precluded amino acid sequencing. The characterization of the codfish food allergen conducted at the same time was more informative. Its amino acid sequence was determined, and it was identified as the muscle calcium regulator parvalbumin. Parvalbumins remain the most important allergens known for fish.[2] Some other major allergens were biochemically characterized before their allergenicity was appreciated. Honey bee phospholipase A had been sequenced in toxicology studies, and the amino acid sequences of ovalbumin and ovomucoid had been determined before they were identified as the major egg allergens, as had the sequences of casein and β-lactoglobulin from milk. Although enormous advances were made in the 1960s and 1970s in the techniques of protein purification, the characterization of proteins required large amounts of purified material, and a period of stagnation in protein research occurred.[3] Amino acid sequences of the minor Amb a 3 and Amb a 5 allergens were solved,[1] but few other accomplishments were made before cDNA cloning was introduced.

IMMUNOLOGIC PUSH FOR MOLECULAR CLONING

Solid-phase synthesis of peptides made it routine to synthesize polypeptides of up to 50 amino acids, and synthetic peptides representing sequences of antigens were being used to immunize animals to induce antibodies that reacted with the native antigen. Similarly, even before the principles of antigen processing and presentation were elucidated, it was known that 10- to 12-mer synthetic peptides taken from a linear sequence of antigens could act as T-cell epitopes and immunize animals[4] or, as shown in tissue culture, inhibit the activation of human T cells.[5] Clearly this pointed to the possibility of new types of immunotherapy, as did the demonstration that variations in the sequences of T-cell epitopes in evolutionarily related antigens markedly altered immune responsiveness. The investigations of Ronald Schwartz and his colleagues showing that T-cell responses of mice to cytochrome c could be modified with sequence variants taken from different species were particularly influential for the initiation of allergen cloning by the author.[6] The use of urea-denatured allergens as a method of targeting T-cell responses had already been investigated in murine experiments with ragweed,[1] and site-directed mutagenesis, which is required to produce genetically engineered antigens, had been established in 1978.[7]

The quest to understand the link between the extraordinary polymorphism of the genes of the MHC and the immunologic responsiveness to allergens was also a major impetus. Experiments with mice had shown that the amino acid sequences of the epitopes determined which antigens could be presented by the MHC molecules of different alleles, and therefore it was thought that similar phenomena would be uncovered for allergy. The propensity of people with certain MHC alleles to produce

From the Centre for Child Health Research, University of Western Australia, Telethon Institute for Child Health Research.

Supported by the National Health and Medical Research Council of Australia.

Disclosure of potential conflict of interest: W. R. Thomas has received royalties from the license of house dust mite allergen patents owned by the Telethon Institute for Child Health Research.

Received for publication October 14, 2010; revised November 29, 2010; accepted for publication December 10, 2010.

Available online January 20, 2011.

Reprint requests: Wayne R. Thomas, PhD, Telethon Institute for Child Health Research, 100 Roberts Rd, Subiaco, Western Australia 6008, Australia. E-mail: wayne@ichr.uwa.edu.au.

0091-6749/$36.00

doi:10.1016/j.jaci.2010.12.1084

antibodies to the minor ragweed allergens had been described by Marsh et al,[8] especially the association of antibodies to Amb a 5 with the class II allele HLA-Dw2. However, Amb a 5 is a 45-amino-acid peptide and not a strong allergen, so studies to discover MHC restrictions of responses to major allergens and how they varied with allergen sequences, such as isoforms of grass allergens, were high on the research agenda. It is now known that immune responses to major allergens typically do not show consistent associations with particular MHC alleles, but these investigations were a major stimulus for sequencing allergens.

FROM IMMUNOCHEMISTRY TO MOLECULAR CLONING

Molecular cloning resulted from the realization that DNA containing a gene of one organism can be transferred and made to function in a genetically different organism. The most frequent incarnation of this is to transfer a gene encoding a protein of interest to the bacterium *Escherichia coli*, where each *E coli* produces 10 to 100 copies of the gene. The cloning results from the ability to transfer 1 gene to each *E coli* and to propagate it to produce limitless progeny. The cloning of the DNA allows it to be isolated for analysis and genetic engineering, and the bacterial hosts of the DNA can be made to produce the product of the gene. In practice, the DNA used for cloning the "genes" that code for proteins from eukaryotes is made as a DNA copy of the mRNA that codes for the protein (cDNA). The cloning of DNA transferring antibiotic resistance between bacteria, accomplished in 1973, was one of the most clear-cut conceptual and practical advances made in biology. Even so, many milestones needed to be passed before the ability to confer antibiotic resistance to bacteria became a technology for the characterization and production of recombinant proteins.[7] Because of the abundance of its mRNA in the hen oviduct, the ovalbumin allergen Gal d 2 was a trailblazer in molecular cloning along with rabbit globin. Its nucleotide sequence was solved during the development of sequencing methodology. The expression of recombinant ovalbumin in *E coli* reported in 1978 was made at the same time as the more exalted expression of growth hormone from cDNA constructed from the mRNA of a pituitary tumor.[7]

The use of cDNA cloning for proteins produced from low-abundance mRNA required substantial technical advances in library construction and screening methodologies. The production of large cDNA libraries suitable for high-density screening was achieved by cloning with the λ gt10 and gt11 bacteriophages. Their high-density lawns of plaques could be screened by using DNA hybridization, and importantly for IgE antibody based screening, high-throughput immunoscreening methods for bacterial colonies and plaques were devised.[9]

Another important development was the microsequencing of proteins immobilized on polyvinylidene difluoride membranes.[3] This was not only able to provide 20 to 30 N-terminal residues from submicrogram amounts of protein, but sequencing could be directly performed from bands electroblotting onto the membranes after SDS-PAGE.

CLONING OF THE FIRST ALLERGENS

The feasibility of cDNA cloning of house dust mite allergens was demonstrated by the detection of Der p 1 and other IgE-binding polypeptides among the *in vitro* translation products of mRNA made from the bodies of *Dermatophagoides pteronyssinus* mites. Libraries were then constructed with λ gt10 and λ gt11 vectors and screened by using a plaque immunoassay with anti–Der p 1 antisera and by using DNA hybridization with oligonucleotides synthesized from the sequences of the N-terminal and peptide fragments from trypsin digests.[10] Clones that encoded Der p 1 were identified by their translated amino acid sequences, which also had high identity to the N-terminal sequence published for Der f 1. Publication was delayed so that IgE binding could be demonstrated but was eventually made with the knowledge that IgE binding to Der p 1 was very sensitive to denaturation and might need considerable work to achieve. The reports of Der p 1 cloning only just preceded the report of the cloning of Dol m 5 from white-faced hornet venom. This was accomplished by using an identical strategy, except that the mRNA was extracted from a specific organ, the acid venom gland.[11] IgE binding was also not demonstrated.

The amino acid amino sequence of Der p 1 was immediately of considerable interest.[12] Der p 1 was revealed to be a cysteine protease similar to papain, and therefore the potent allergenicity might be linked with an adjuvant effect of protease activity. The overall tertiary structure was also immediately revealed because the structures of papain and the related actinidin enzymes had been solved. Although not perfectly folded, high IgE-binding recombinant allergens were soon produced, and the publication in 1992 of rDer p 1 made in *Saccharomyces cerevisiae* with near-natural and high-frequency IgE binding formally demonstrated the cloning of a major allergen.[13] Because, like Der p 1, IgE binding of the Dol m 5 had been well established, the main reason for cloning was to complete the characterization of the venom allergens and to explore the molecular basis for allergenic cross-reactivity between vespid species. The sequence showed identity to plant pathogenesis-related proteins, now designated as the V5/Tpx-1–related family. The family members are remarkably conserved across fungi, plants, animals, and parasites, but their precise functions are unclear.

CLONING BY MEANS OF IGE SCREENING

The first reports of IgE-binding recombinant allergens resulted from the use of IgE immunoassays used to screen cDNA expression libraries. Der p 2[14,15] and Der p 5[16] from the house dust mite, Bet v 1 from birch pollen,[17] and Dac g 2 from orchard grass[18] were cloned this way. The cloning of Bet v 1 was a watershed. It unveiled the allergenicity of PR10 pathogenesis-related proteins, which are now known to induce strong IgE responses in a diverse range of pollen and food hypersensitivities and to be a cause of the oral allergy syndrome. It was also an ideal allergen for pioneering studies of recombinant allergens for immunotherapy.[19] Most birch pollen–sensitive patients direct 80% of their anti-birch IgE to this allergen, and immunotherapy to treat birch pollen–induced allergic rhinitis is commonly practiced. Bet v 1 was also the first recombinant allergen used to determine NMR and X-ray crystallographic structures.[2] Recombinant Der p 2 and Der f 2 have also been frequently used for structural studies[2] and were the first members of ML (MD2-like) domain proteins, which is an inappropriate name because MD2 was modeled on Der p 2.

MAJOR ALLERGENS FROM MAJOR SOURCES

The cloning of major allergens from other important sources quickly followed with the cloning of the group 1 and group 5 grass