100 Years later: Celebrating the contributions of x-ray crystallography to allergy and clinical immunology

Anna Pomés, PhD,[a] Maksymilian Chruszcz, PhD,[b] Alla Gustchina, PhD,[c] Wladek Minor, PhD,[d] Geoffrey A. Mueller, PhD,[e] Lars C. Pedersen, PhD,[e] Alexander Wlodawer, PhD,[c] and Martin D. Chapman, PhD[a] *Charlottesville, Va, Columbia, SC, Frederick, Md, and Research Triangle Park, NC*

Current knowledge of molecules involved in immunology and allergic disease results from the significant contributions of x-ray crystallography, a discipline that just celebrated its 100th anniversary. The histories of allergens and x-ray crystallography are intimately intertwined. The first enzyme structure to be determined was lysozyme, also known as the chicken food allergen Gal d 4. Crystallography determines the exact 3-dimensional positions of atoms in molecules. Structures of molecular complexes in the disciplines of immunology and allergy have revealed the atoms involved in molecular interactions and mechanisms of disease. These complexes include peptides presented by MHC class II molecules, cytokines bound to their receptors, allergen-antibody complexes, and innate immune receptors with their ligands. The information derived from crystallographic studies provides insights into the function of molecules. Allergen function is one of the determinants of environmental exposure, which is essential for IgE sensitization. Proteolytic activity of allergens or their capacity to bind LPSs can also contribute to allergenicity. The atomic positions define the molecular surface that is accessible to antibodies. In turn, this surface determines antibody specificity and cross-reactivity, which are important factors for the selection of allergen panels used for molecular diagnosis and the interpretation of clinical symptoms. This review celebrates the contributions of x-ray crystallography to clinical immunology and allergy, focusing on new molecular perspectives that influence the diagnosis and treatment of allergic diseases. (J Allergy Clin Immunol 2015;136:29-37.)

Key words: *Allergens, allergy, function, structure, cross-reactivity, x-ray crystallography*

In 1912, the German physicist Max von Laue published a demonstration of x-ray diffraction from a crystal of copper sulfate pentahydrate.[1] This pioneering event led to a Nobel Prize 2 years later and marked the beginning of x-ray crystallography. The discipline celebrated its 100th anniversary in 2014.

Crystallography started in the hands of physicists with studies of nonbiological molecules and was extended to the fields of chemistry, biology, and medicine. Crystal structures define the spatial location of atoms and the folding of macromolecules involved in biological processes. The first x-ray diffraction data from a small protein, pepsin, were collected by Crowfoot and Bernal[2] in 1934 (see Fig E1, *A*, in this article's Online Repository at www.jacionline.org). Myoglobin and hemoglobin were the first protein structures obtained by Kendrew et al[3] in 1958 and Perutz et al[4] in 1960, respectively. In 1965, the first enzyme structure to be determined by Blake et al[5] was hen egg white lysozyme, which happens to be the chicken food allergen Gal d 4. With the advent of more powerful x-ray sources (eg, synchrotrons), new detectors, and new experimental protocols (eg, reduction of radiation damage by means of cryocooling) and the development of modern software, the number of macromolecular structures has exponentially increased (see Fig E2 in this article's Online Repository at www.jacionline.org).[6,7] The bulk of macromolecular diffraction data collection has moved from in-house facilities to high flux sources, mainly at synchrotrons. Nowadays, the Cambridge Structural Database contains more than 750,000 structures of organic molecules, mostly determined by using x-ray crystallography, and the Protein Data Bank contains more than 105,000 structures of biomacromolecules, almost 90% of which were determined by means of x-ray diffraction.[8]

This review is a tribute to the contributions of x-ray crystallography to clinical immunology and allergy. Notable achievements include the determination of the 3-dimensional structure of allergens, antibodies, receptors, and other molecules involved in immunologic processes and allergic disease. The review also highlights findings derived from

From [a]Basic Research, INDOOR Biotechnologies, Charlottesville; [b]the Department of Chemistry and Biochemistry, University of South Carolina, Columbia; [c]the Macromolecular Crystallography Laboratory, National Cancer Institute, Frederick; [d]the Department of Molecular Physiology and Biological Physic, University of Virginia, Charlottesville; and [e]the Genome Integrity and Structural Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park.

Supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH; award no. R01AI077653; A.P. contact PI), the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, and the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences Research (project no. Z01-ES050147 to G.A.M. and ZIA ES102645 to L.C.P.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure of potential conflict of interest: A. Pomés and M. Chruszcz have received research support from the National Institutes of Health (NIH). A. Gustchina and A. Wlodawer are employed by the National Cancer Institute. W. Minor has received research support and fees for participation in review activities from the NIH; is a board member for and has stock/stock options in HKL Research, Inc; is employed by the University of Virginia; has patents through the University of Virginia and receives royalties from the University of Virginia for these patents; and was involved in developing some of the crystallography techniques presented in this paper which were later commercialized and became widely used. M. Chapman has received research support from the National Institute of Allergy and Infectious Disease and the National Institute of Environmental Health Sciences, is employed by and is co-owner of Indoor Biotechnologies, and has patents through the University of Virginia. The rest of the authors declare that they have no relevant conflicts of interest.

Received for publication March 9, 2015; revised April 21, 2015; accepted for publication May 14, 2015.

Corresponding author: Anna Pomés, PhD, INDOOR Biotechnologies, 1216 Harris St, Charlottesville, VA 22903. E-mail: apomes@inbio.com.

0091-6749/$36.00

© 2015 American Academy of Allergy, Asthma & Immunology

http://dx.doi.org/10.1016/j.jaci.2015.05.016

Abbreviations used
GFP: Green fluorescent protein
IL-4R: IL-4 receptor
IUIS: International Union of Immunological Societies
MBP: Maltose-binding protein
TLR: Toll-like receptor
WHO: World Health Organization

structural studies, which revealed mechanisms that contribute to allergic sensitization.

THE BASICS OF 3-DIMENSIONAL STRUCTURE DETERMINATION

The determination of an x-ray crystal structure starts with the purification of a sufficient amount of highly concentrated and homogeneous protein (Fig 1). Several factors that contribute to molecular variability need to be taken into consideration, including molecular variants, glycosylation, proteolysis, fragmentation, aggregation, precipitation, and/or molecular flexibility (see the text in this article's Online Repository at www.jacionline.org). The expression of recombinant allergens, either alone or as fusion proteins, has often proved successful for obtaining homogeneous protein preparations.[9-12] Molecules naturally embedded in membranes, such as the histamine receptor, might require strategies involving protein solubilization.[13] Procuring soluble, properly folded, and pure protein is the most promising way to obtain diffraction-quality crystals and the major bottleneck in the structure determination process. The difficulty in obtaining diffraction-quality crystals promotes the development of complementary technologies, such as cryo-electron microscopy.[14]

Screening of optimal conditions for crystal growth is performed either manually or by means of robotics.[15] Hundreds of conditions are tested by mixing the highly concentrated protein with different precipitants. A commonly used technique is based on "hanging" or "sitting" drops in multiwell plates, where crystals form by slowly concentrating the protein and precipitant through vapor diffusion. Once crystals are obtained, the x-rays to probe them are produced either by rotating anode generators found in most crystallography laboratories or at synchrotron stations. There are more than 120 stations dedicated to macromolecular crystallography experiments worldwide.[16]

The diffraction pattern produced by exposing protein crystals to x-rays is analyzed to obtain an electron density map that is used for determination of a molecular model.[17-19] The overall quality of structural models, as measured by various parameters used for structure validation, has significantly improved over time.[20] Structure quality depends not only on the resolution of the collected diffraction data but also on the processing of these data. Resolution is the smallest distance in angstroms between 2 atoms that can be shown to be separated (see Fig E3 in this article's Online Repository at www.jacionline.org). The myoglobin structure determined in 1958 had a resolution of 6 Å. Nowadays, resolutions as high as 0.48 Å can be obtained.[21]

Crystallography has evolved from manual determination of molecular models to the use of hardware and software that greatly enhances the efficiency and quality of data collection, data reduction, model building and refinement. These transformative developments allow rapid data processing, model construction, and refinement (optimization).[22,23] Sophisticated software, such as HKL-3000, has enabled the determination of structures with "speed and finesse."[22] Indeed, a process that used to take years or months is now performed in days or even in hours in optimal cases (see Fig E2). The productivity of synchrotron stations, as measured based on structures determined, varies significantly and mostly depends on experimental protocols.[24] There is no correlation between beamline productivity and any aspect of physical setup of data collection hardware, beamline flux, or number of crystals used for diffraction. Frequent reports show that only 1 to 5 minutes of data collection time are needed to generate an entire data set sufficient for a good structure determination.[25] The most productive synchrotron stations still need roughly 20 hours of data collection time to produce 1

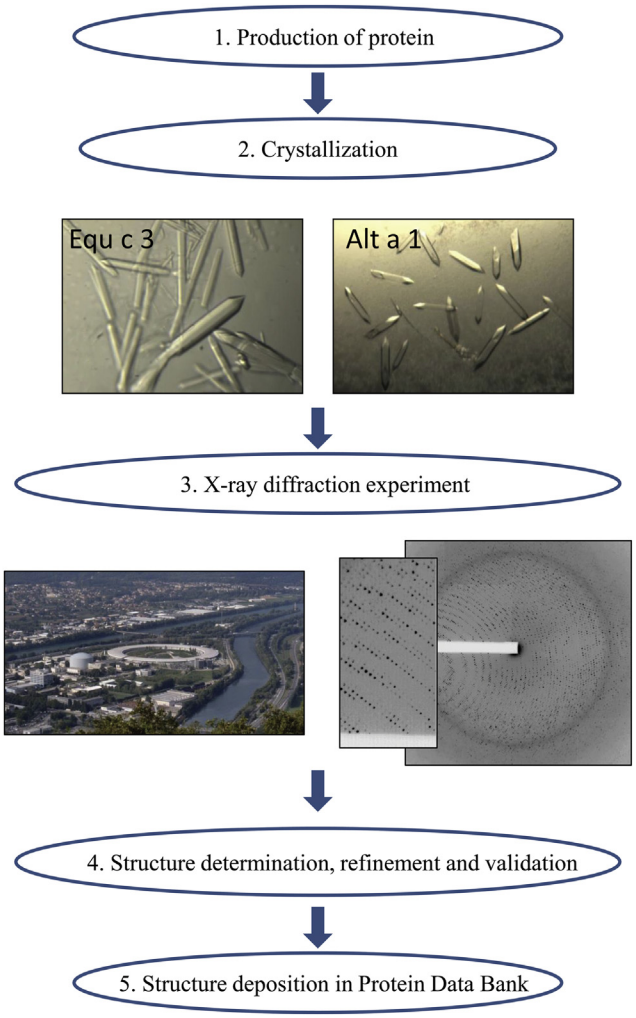

FIG 1. General outline of the process of structure determination by using x-ray crystallography. Crystals of horse Equ c 3 and Alt a 1 from *Alternaria alternata* grown by using the vapor diffusion technique and used for x-ray diffraction experiments are shown. The European Synchrotron Radiation Facility and Institute Laue-Langevin in Grenoble (France; *3, left*; photographer: Christian Hendrich) is shown next to diffraction images obtained from a protein crystal used for structure determination (3N99; *3, right*).

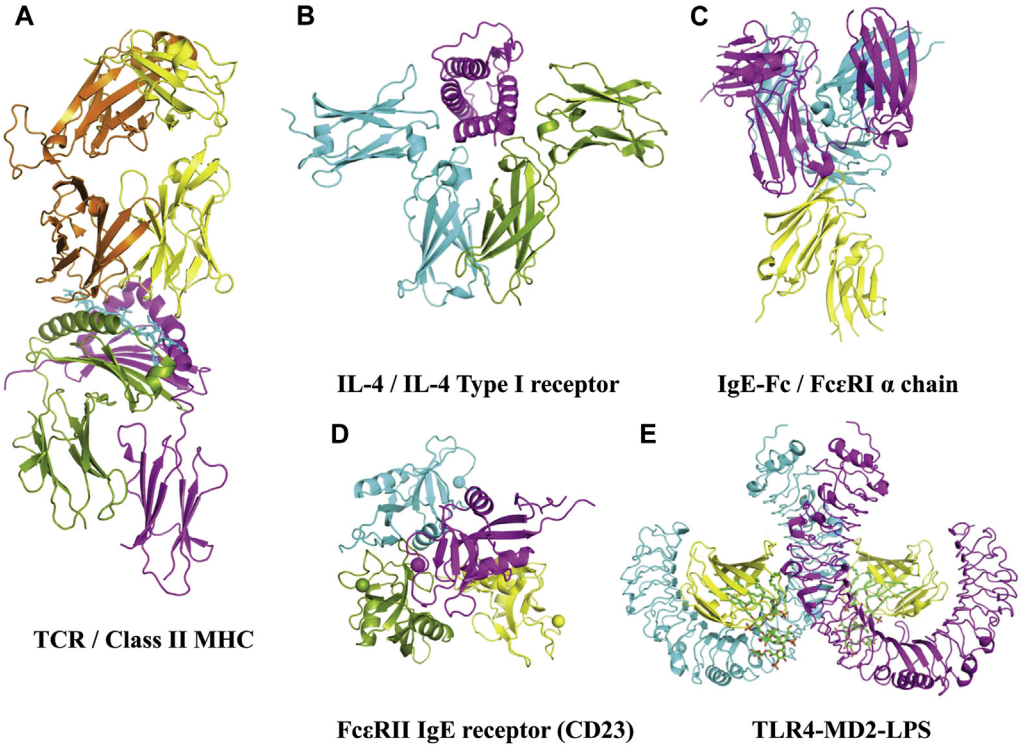

FIG 2. X-ray crystal structures of human molecules involved in immunology and allergic disease. **A,** Complex of a human α/β T-cell receptor (*TCR*; *orange* and *yellow*), influenza HA antigen peptide *(blue)*, and MHC class II molecule HLA-DR1 (*green* and *fuchsia*; 1FYT). **B,** Cytokine receptor IL-4 type I (*blue* and *green*) bound to cytokine IL-4 (*fuchsia*; 3BPL). **C,** IgE-Fc (Cε2, Cε3, and Cε4 domains; *blue* and *fuchsia*) bound to the extracellular domains of the high-affinity IgE receptor (FcεRI) α chain (*yellow*; 2Y7Q). **D,** Low-affinity IgE receptors (CD23, FcεRII; 4G9A). **E,** TLR4 (*blue* and *fuchsia*)–MD-2 *(yellow)*–LPS *(green)* complex (3FXI).

structure. There is still room for improvement of experimental protocols and software in the future.

THREE-DIMENSIONAL STRUCTURES OF MOLECULES INVOLVED IN IMMUNOLOGY AND ALLERGY

Immunologic processes result from the interaction of specialized molecules, the structure of which defines their function. Selected examples involved in adaptive and innate immunity are shown in Fig 2. The structure of MHC class II molecules expressed *in vivo* on the surfaces of antigen-presenting cells clearly revealed a groove that binds the peptides generated from antigen processing and presents them to T-cell receptors (Fig 2, *A*). This action, together with costimulatory signals, leads to activation of T cells to release cytokines that interact with cytokine receptors in B cells (Fig 2, *B*). Crystal structures of ovalbumin (Gal d 2) peptides bound to mouse MHC class II and a peptide from Cry j 1, a major allergen from the Japanese red cedar *(Cryptomeria japonica)*, bound to HLA-DP5 also aid in our understanding of molecular recognition of allergens.[26,27] The structural basis of the interaction of cytokines with their receptors has been investigated to better understand what induces the production of antibodies by B cells. Cytokine receptors of the IL-4/IL-13 system result from the assembly of different subunits. Interestingly, a single subunit can be shared by different cytokine receptors (ie, IL-4 receptor [IL-4R] α shared by IL-4/IL-4Rα/γc and IL-13/IL-4Rα/IL-13 receptor α1; Fig 2, *B*). These molecular interactions determine mechanisms of cytokine action and influence disease phenotype and response to treatment.[28,29]

IgE antibodies eventually bind to high-affinity IgE receptors (FcεRI) on mast cells and basophils (Fig 2, *C*). X-ray crystal structures showed the flexibility of IgE-Fc (consisting of the Cε2, Cε3, and Cε4 domains). IgE-Fc adopts different conformations, ranging from an acutely bent structure of IgE-Fc when bound to the extracellular domains of the FcεRI α chain (Fig 2, *C*) to a diversity of conformations in solution, including a fully extended symmetric one.[30,31] The unique slow dissociation rate of IgE from FcεRI was attributed in part to these conformational changes. This observation provides a strategy for the design of asthma therapeutics using peptides that would disrupt the interaction between IgE and its high-affinity IgE receptor.[31]

The low-affinity IgE receptor (CD23, FcεRII) on B cells is a calcium-dependent molecule with a C-type lectin domain that contributes to the regulation of IgE levels. Calcium induces structural changes in CD23 that lead to additional interactions with IgE and a 30-fold increase in affinity for IgE (Fig 2, *D*).[32] The crystal structure of IgE bound to CD23 revealed a mechanism of reciprocal allosteric inhibition with the high-affinity IgE receptor.[33]

Structures of molecules involved in innate immunity have also been determined, including the extracellular domain of Toll-like receptors (TLRs). The structures of TLRs differ from the "immunoglobulin-like" extracellular domains of the B- and

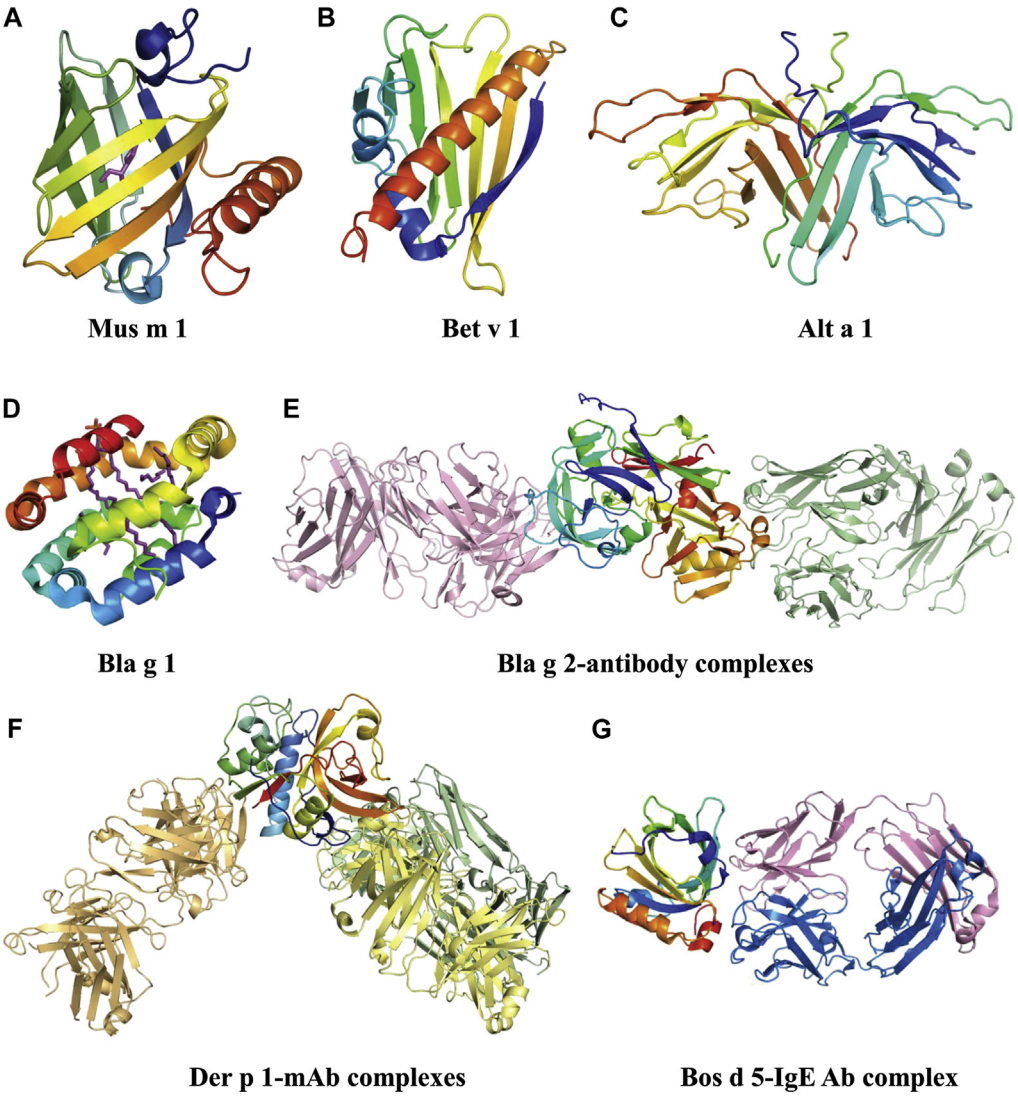

FIG 3. X-ray crystal structures of selected allergens. **A,** Mus m 1 (1MUP). **B,** Bet v 1 (1BV1). **C,** Alt a 1 (3V0R). **D,** Bla g 1 (4JRB). **E,** Bla g 2 in complex with Fab of mAb 7C11 (2NR6; *pink*) superimposed with the mAb-Fab from the structure of Bla g 2 in complex with Fab of mAb 4C3 (3LIZ; *green*). **F,** Der p 1 in complex with 4C1 (*pale yellow*, 3RVW) superimposed with the mAb-Fab from the 2 structures of the allergen in complex with 10B9 (*pale green*; 4PP2) or 5H8 (*pale orange*, 4PP1). **G,** Bos d 5 in complex with the Fab of a recombinant IgE antibody. Allergens are shown in rainbow colors from the N-terminus *(blue)* to C-terminus *(red)* and ligands in magenta (Fig 3, *A* and *D*).

T-cell receptors mentioned above. For all these receptors, only the structures of the extracellular domains have been determined. Ten TLRs are known in human subjects, and they are formed by an N-terminal recognition domain, a single transmembrane helix, and a C-terminal signaling domain. The N-terminal domain recognizes pathogen-derived compounds or endogenous molecules released by the host in response to infection.[34] This domain adopts a typical horseshoe shape made of tandem copies of a motif known as the leucine-rich repeat. Its concave surface has ligand-binding capacity in most leucine-rich repeats. Two extracellular domains form a dimer upon ligand binding and activate signaling. The structure of TLR4, for example, has revealed the importance of its interaction with MD-2 and endotoxin for signaling activation (Fig 2, *E*).

ALLERGEN STRUCTURE AS A DETERMINANT OF ALLERGIC DISEASE

Immunologic processes that lead to the development of allergic disease are strongly associated with structural features of the allergen (Fig 3). Allergen exposure, which is critical to IgE sensitization, is determined in part by the function of the allergen and the degree of molecular stability required for the protein to become an allergen, both of which are defined by the allergen structure. Allergens must be either released to the environment (eg, pollens, spores, dander, or fecal particles) or made accessible through other paths (ingestion of foods, injection of venoms, or contact through skin or infection) to access the immune system. The function of the allergen can facilitate exposure in different ways. Some allergens are released to the environment because

of their reproductive function in pollen or spores. Others have a digestive function and are excreted in fecal particles. Additional factors determine exposure, such as the aerodynamic properties of particles that carry inhalant allergens.

One of the first functions gleaned from the x-ray crystal structure of allergens was obtained in 1992 for the major urinary proteins from mouse and rat. Both Mus m 1 and Rat n 1, which were not known as allergens at that time, are pheromone transporters and belong to the lipocalin family of proteins. Lipocalins consist of a β-barrel with a hydrophobic ligand-binding cavity (Fig 3, *A*) and are secreted in tears, urine, sweat, or saliva, which facilitates exposure.[35] Lipocalins are common mammalian inhalant allergens that are also found in cockroach (Bla g 4) and cow's milk (food allergen Bos d 5).

The first 3-dimensional structure of a clinically important allergen was Bet v 1, a pathogenesis-associated protein from birch pollen (Fig 3, *B*).[36] Bet v 1 is the most extensively studied allergen from pollen. The IgE prevalence in temperate climate areas of the northern hemisphere (ie, Northern Europe) is high (>95%). Bet v 1 shows clinical cross-reactivity with homologous allergens from fruits and vegetables (ie, apples, celery, and carrot).[37] A large number of variants (n = 18) have been identified in natural Bet v 1, sharing high amino acid sequence identity (approximately 95%). Their nomenclature has recently been revised by the Allergen Nomenclature Sub-Committee from the World Health Organization (WHO) and International Union of Immunological Societies (IUIS; www.allergen.org).[38,39] Bet v 1.0101 is the major component of natural Bet v 1 (>50%) and the main sensitizer, whereas other isoforms (Bet v 1.0401 and Bet v 1.1001) induce only minimal IgE antibody responses.[40] Bet v 1.0101 and Bet v 1.0401 share the same fold, but differences in dimerization or aggregation could contribute to a decreased ability of the Bet v 1.0401 variant to elicit an allergic immune response. Interestingly, the fold of Bet v 1.0101 *per se* was demonstrated to be important for T_H2 polarization and induction of a strong IgE response by means of comparison with an engineered folding variant.[41]

Nowadays, the Protein Data Bank (www.rcsb.org) contains the 3-dimensional structures of more than 100 allergens representing approximately 50 protein families from approximately 800 allergens currently present in the WHO/IUIS official database of systematic Allergen Nomenclature (see Tables E1 and E2 in this article's Online Repository at www.jacionline.org). A database of allergen families (http://www.meduniwien.ac.at/allergens/allfam/) reports that allergens belong to a wide array of 186 protein families.[42,43] The most frequent biochemical functions of allergens are protein hydrolysis, carbohydrate metabolism, binding of metal ions and lipids, storage and functions associated with the cytoskeleton.[42] Measurement of biological activity with a specific functional assay is the best way to confirm function, but this option is not always available. Sequence homology to a protein of known function might also be insufficient.[44] Determination of tertiary structure can delineate allergen function by defining the overall shape of the molecule, revealing specific functional residues, or both.[45] The major dust mite allergens Der p 1 and Der f 1 are cysteine proteases. Their catalytic site has been identified in the 3-dimensional structures of the native allergens (Fig 3, *F*).[46,47] In contrast, the cockroach allergen Bla g 2 folds as a typical pepsin-like aspartic protease but is devoid of aspartic protease activity because of specific substitutions in the catalytic site that were revealed by the crystal structure (Fig 3, *E*, and see Fig E1, *A*).[48,49] In other cases molecular flexibility confers a regulatory function, as seen in the calcium-binding allergens commonly found in pollen and in troponin, which regulates muscle contraction (ie, Bla g 6).[50] The first 3-dimensional structure of a representative of the 2 EF-hand allergen family was reported for Phl p 7 bound to calcium.[51] Differences observed in IgE antibody binding depending on allergen conformation suggest that conditions and conformation with optimal IgE reactivity should be selected for molecular diagnosis.

The storage peanut proteins Ara h 1 and Ara h 2 are food allergens with very different structural complexities. Ara h 1 is a trimer of 60-kDa bicupin-fold monomers, whereas Ara h 2 is a small monomeric α-helix protein (17 kDa; see Fig E4 in this article's Online Repository at www.jacionline.org).[11,52] It has been suggested that the quaternary structure of Ara h 1 might play a role in its allergenicity by increasing molecular stability and preventing digestion of IgE antibody-binding epitopes.[53] However, there is no evidence that the differing tertiary structure of both allergens is responsible for differing allergenic potential. In fact, IgE antibody titers to Ara h 2, which has a simpler structure, have been reported as the best predictor of peanut allergy.[54-56] In general, 3-dimensional structural complexity of allergens does not seem to be related to their allergenicity.

Finally, determination of the x-ray crystal structure of allergens has revealed the existence of new structural groups of proteins. Alt a 1, the major allergen from *Alternaria alternata*, has a unique β-barrel structure that forms a "butterfly-like" dimer and is exclusively found in the Dothideomycetes and Sordariomycetes classes of fungi (Fig 3, *C*).[57] The cockroach allergen Bla g 1 has an α-helical structure thus far only found in insects (Fig 3, *D*).[12,58] Although the 3-dimensional structures of these and other allergens have been determined, their functions still are not well understood.[10,57,59] The cat allergen Fel d 1 has a uteroglobulin-like fold and consists of a dimer of heterodimers made exclusively of α-helices. Structures of recombinant Fel d 1 made by fusion of the monomers involved in each heterodimer have been determined.[60,61] However, the native assembly of this major cat allergen remains unknown.

INFLUENCE OF ALLERGEN FUNCTION ON ALLERGENICITY

An interesting example of how an x-ray crystal structure revealed a function associated with allergenicity is illustrated by Der p 2. This dust mite allergen resembles MD-2, the LPS-binding component of the TLR4 signaling complex, which is involved in activation of the innate immune system. Both proteins have an immunoglobulin-like fold formed only of β-sheets (see Figs E1, *B*, and E4, *A*).[62,63] The surprisingly high structural similarity prompted an investigation as to whether Der p 2 would have a similar function. Der p 2 was not only able to mimic the function of MD-2 in mouse models of experimental asthma but was also able to reconstitute LPS-driven TLR4 signaling in the absence of MD-2, suggesting a possible role of Der p 2 in activation of innate immunity.[64] Results obtained with Der p 2 are supported by data demonstrating that low-level stimulation of TLRs drives T_H2 immune responses.[64-68] This discovery marked an important step in our understanding of possible innate immune pathways that lead to allergic sensitization. Until then, activation of the adaptive immune system was the main

recognized path for allergy development. Proteolytic function had been considered to contribute to allergenicity by cleaving molecules involved in the immune response.[69,70] Der p 1 can contribute to allergenicity through noncanonical pathways. Der p 1 was reported to directly promote IgE synthesis through cleavage of the low-affinity IgE receptor (CD23) on B cells and, indirectly, through cleavage of the α-subunit of the IL-2 receptor (CD25) in T cells.[71,72] Der p 1 can also contribute to inflammation by inducing cytokine production and disrupting gap junctions in the lung epithelium, which would increase membrane permeability and facilitate transepithelial allergen delivery and processing.[69,70,73,74]

Structural studies have shown that an increasing number of allergens belonging to different structural families bind hydrophobic ligands and that these are potent stimulators of the innate immune system.[65] The crystal structure of Bla g 1 provided the first clues that this protein might be a lipid-binding protein (Fig 3, *D*). Lipids in cockroach frass were identified as fatty acids that are known to activate TLR2 and TLR4.[75] Some allergens contain lipids in internal cavities that are formed by either β-sheets (Der p 2 and lipocalins) or α-helices (Fel d 1 and Bla g 1; Fig 3, *A* and *D*, and see Fig E1, *B*).[12,61,62,76]

Other allergens, lacking internal cavities, can also bind lipidic ligands in different ways. Der p 5 consists of 3-helix bundles that tend to form multimers (see Fig E4, *B*). A large hydrophobic cavity formed by each dimer has the potential to bind lipidic ligands.[10] Der p 7 has a similar structure to an LPS-binding protein involved in TLR4 activation and to the surfactant allergen Equ c 4.[9,77] The fold consists of two 4-strand antiparallel β-sheets that wrap around a long C-terminal helix (see Fig E4, *C*).[9] Although unable to bind LPS, Der p 7 bound the lipopeptide polymyxin B.[9] The specific contribution of allergen-associated lipids to allergenicity has been recently reviewed.[78] The interactions of these lipids with innate immunity are open to further investigation.

HOW CRYSTALLOGRAPHY CONTRIBUTES TO CLINICAL PRACTICE

The allergen structures determined by using x-ray crystallography provide the basis for improved diagnosis and therapy. An undesired side effect of immunotherapy is the induction of adverse reactions that can occur when administering increasing doses of the allergen during vaccination. To avoid these effects, hypoallergens with reduced IgE reactivity that preserve their capacity to induce T-cell responses have been designed as candidates for vaccination. One strategy is the disruption of the overall allergen fold. This has proved efficacious with hypoallergenic chemically modified extracts (allergoids) that are successful for immunotherapy in Europe.[79] Numerous variants of recombinant hypoallergens have also been designed, but only Bet v 1 variants have reached clinical trials, in which they have already shown promising results.[80] Recently, a dose-ranging immunotherapy study of a new recombinant hypoallergenic folding variant of Bet v 1 showed efficacy in an environmental exposure chamber.[81] Although knowledge of the 3-dimensional structure of the allergen is not always necessary for the production of unfolded variants, the disruption of the overall allergen fold can be effectively designed based on the modification of specific structural features. For example, mite group 2 hypoallergens were produced by mutating cysteines involved in the formation of disulfide bonds that preserve the immunoglobulin-like structure.[82,83] Another strategy to produce hypoallergens consists of modifying residues involved in IgE antibody binding, without affecting the fold of the allergen. In this case knowledge of the 3-dimensional structure of the allergen is required. The best way to precisely locate the epitope recognized by an antibody is the determination of the structure of the allergen in complex with the antibody. For example, the structure of Bet v 1 in complex with the Fab fragment of an mAb that interfered with IgE antibody binding identified a dominant epitope also involved in IgE cross-reactivity with homologous allergens.[84,85] The x-ray crystal structures of additional complexes revealed determinants of specificity and cross-reactivity.[47,86-91] Bla g 2 and Der p 1 have been extensively analyzed through the structures of 2 and 4 allergen-antibody complexes, respectively (Fig 3, *E* and *F*).[47,89-91] The structural basis of cross-reactivity was analyzed for mAb 4C1, which binds Der p 1 and Der f 1 in equivalent epitopes (Fig 3, *F*).[47] It is not possible to obtain the milligram amounts of pure native IgE antibodies required for x-ray crystallography given their polyclonality and low concentration in sera compared with IgG antibodies (0.05%). Therefore recombinant IgE mAb, which were derived from combinatorial libraries from allergic patients, were used in complexes with bovine milk Bos d 5 (β-lactoglobulin; Fig 3, *G*) and timothy grass pollen Phl p 2.[87,88] This strategy, combined with site-directed mutagenesis, is a powerful tool to identify the main residues involved in IgE antibody recognition and to produce hypoallergens.[47,85,92]

Molecular allergy diagnosis has been shown to improve patient diagnosis compared with exclusively using clinical history and skin prick tests with allergen extracts.[37,93-97] Structural analyses reveal relationships between homologous allergens and contribute to the design of panels of purified allergens for molecular diagnosis. Although a general rule suggests that cross-reactivity is likely among proteins that share a high degree of amino acid identity throughout the entire protein (>70%) and tends to be rare at less than 50% identity, exceptions do occur.[98-102] Predictions of cross-reactivity based on the overall homology among allergens can only be used as guidelines but lack precision. The 3-dimensional structure of allergens allows identification of solvent-exposed residues responsible for antibody recognition.[103] Ideally, the selection of representative molecules from highly cross-reactive groups of allergens and the inclusion of allergens with species-specific epitopes can simplify molecular diagnosis.[104,105]

LOOKING INTO THE FUTURE

In the future, an increase in structure-function studies is expected by combining methodologies from the disciplines of medicine and structural biology. Cryoelectron microscopy will enable study of larger complexes of antigens and antibodies without the need to crystallize them. Use of the free electron laser as an x-ray source will allow investigations using nanocrystals and decrease the chances of structural changes caused by radiation damage.[106,107] Nuclear magnetic resonance, which is currently only rarely used for structural studies of antigen-antibody complexes, is expected to have a higher impact on ligand screening and dynamic studies. As more structures become available, the ability to create informative homology models for hypothesis-driven research will be enhanced. Tighter

interactions between functional and structural studies are expected to affect more significantly the drug discovery process. The creation of a big picture and better reproducibility will be possible by using sophisticated database systems that will organize and analyze structural data in biomedical laboratories.[108]

X-ray crystallography has led to major contributions in clinical immunology and allergy. The 3-dimensional structures of molecules involved in immunologic processes and allergens, either alone or in complex with antibodies, have provided detailed information at the atomic level that reveals mechanisms of molecular interaction. Recombinant allergens are engineered to either preserve their native fold and amino acid composition or to have specific residues and/or the overall fold modified for reduced IgE reactivity, which might diminish side effects caused by increasing allergen doses administered during therapy. Recombinant allergens have already shown promising results in immunotherapy clinical trials. The structural information is being used to increase the accuracy of diagnosis and to design new forms of immunotherapy that will complement and improve current approaches to treat allergic disease.

We thank Jill Glesner for her assistance in preparation of tables and Dr Susanna Keller for providing inspiring literature.

REFERENCES

1. Friedrich W, Knipping P, Laue M. Interferenz-Erscheinungen bei Röntgenstrahlen. Sitz Bayer Akad Wiss 1912;303-22.
2. Crowfoot D, Bernal JD. X-ray photographs of crystalline pepsin. Nature 1934; 133:794-5.
3. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. Nature 1958;181:662-6.
4. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North AC. Structure of haemoglobin: a three-dimensional Fourier synthesis at 5.5-A. resolution, obtained by X-ray analysis. Nature 1960;185:416-22.
5. Blake CC, Koenig DF, Mair GA, North AC, Phillips DC, Sarma VR. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. Nature 1965;206:757-61.
6. Dauter Z, Jaskolski M, Wlodawer A. Impact of synchrotron radiation on macromolecular crystallography: a personal view. J Synchrotron Radiat 2010; 17:433-44.
7. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res 2000;28:235-42.
8. Coontz R, Fahrenkamp-Uppenbrink J, Lavine M, Vinson V. Crystallography at 100. Going from strength to strength. Introduction. Science 2014;343:1091.
9. Mueller GA, Edwards LL, Aloor JJ, Fessler MB, Glesner J, Pomés A, et al. The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. J Allergy Clin Immunol 2010;125:909-17.
10. Mueller GA, Gosavi RA, Krahn JM, Edwards LL, Cuneo MJ, Glesner J, et al. Der p 5 crystal structure provides insight into the group 5 dust mite allergens. J Biol Chem 2010;285:25394-401.
11. Mueller GA, Gosavi RA, Pomés A, Wunschmann S, Moon AF, London RE, et al. Ara h 2: crystal structure and IgE binding distinguish two subpopulations of peanut allergic patients by epitope diversity. Allergy 2011;66:878-85.
12. Mueller GA, Pedersen LC, Lih FB, Glesner J, Moon AF, Chapman MD, et al. The novel structure of the cockroach allergen Bla g 1 has implications for allergenicity and exposure assessment. J Allergy Clin Immunol 2013;132:1420-6.
13. Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, et al. Structure of the human histamine H1 receptor complex with doxepin. Nature 2011;475:65-70.
14. Fischer N, Neumann P, Konevega AL, Bock LV, Ficner R, Rodnina MV, et al. Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by C-corrected cryo-EM. Nature 2015;520:567-70.
15. Cymborowski M, Klimecka M, Chruszcz M, Zimmerman MD, Shumilin IA, Borek D, et al. To automate or not to automate: this is the question. J Struct Funct Genomics 2010;11:211-21.
16. Minor W, Tomchick D, Otwinowski Z. Strategies for macromolecular synchrotron crystallography. Structure 2000;8:R105-10.
17. Chruszcz M, Wlodawer A, Minor W. Determination of protein structures-a series of fortunate events. Biophys J 2008;95:1-9.
18. Wlodawer A, Minor W, Dauter Z, Jaskolski M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. FEBS J 2008;275:1-21.
19. Chruszcz M, Borek D, Domagalski M, Otwinowski Z, Minor W. X-ray diffraction experiment-the last experiment in the structure elucidation process. Adv Protein Chem Struct Biol 2009;77:23-40.
20. Dauter Z, Wlodawer A, Minor W, Jaskolski M, Rupp B. Avoidable errors in deposited macromolecular structures: an impediment to efficient data mining. IUCrJ 2014;1:179-93.
21. Schmidt A, Teeter M, Weckert E, Lamzin VS. Crystal structure of small protein crambin at 0.48 A resolution. Acta Crystallogr Sect F Struct Biol Cryst Commun 2011;67:424-8.
22. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes. Acta Crystallogr D Biol Crystallogr 2006;62: 859-66.
23. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr 2012; 68:352-67.
24. Zheng H, Chordia MD, Cooper DR, Chruszcz M, Muller P, Sheldrick GM, et al. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat Protoc 2014;9:156-70.
25. Zheng H, Hou J, Zimmerman MD, Wlodawer A, Minor W. The future of crystallography in drug discovery. Expert Opin Drug Discov 2014;9:125-37.
26. Scott CA, Peterson PA, Teyton L, Wilson IA. Crystal structures of two I-Ad-peptide complexes reveal that high affinity can be achieved without large anchor residues. Immunity 1998;8:319-29.
27. Kusano S, Kukimoto-Niino M, Satta Y, Ohsawa N, Uchikubo-Kamo T, Wakiyama M, et al. Structural basis for the specific recognition of the major antigenic peptide from the Japanese cedar pollen allergen Cry j 1 by HLA-DP5. J Mol Biol 2014;426:3016-27.
28. LaPorte SL, Juo ZS, Vaclavikova J, Colf LA, Qi X, Heller NM, et al. Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system. Cell 2008;132:259-72.
29. Romeo MJ, Agrawal R, Pomés A, Woodfolk JA. A molecular perspective on TH2-promoting cytokine receptors in patients with allergic disease. J Allergy Clin Immunol 2014;133:952-60.
30. Holdom MD, Davies AM, Nettleship JE, Bagby SC, Dhaliwal B, Girardi E, et al. Conformational changes in IgE contribute to its uniquely slow dissociation rate from receptor FcεRI. Nat Struct Mol Biol 2011;18:571-6.
31. Drinkwater N, Cossins BP, Keeble AH, Wright M, Cain K, Hailu H, et al. Human immunoglobulin E flexes between acutely bent and extended conformations. Nat Struct Mol Biol 2014;21:397-404.
32. Yuan D, Keeble AH, Hibbert RG, Fabiane S, Gould HJ, McDonnell JM, et al. Ca2+-dependent structural changes in the B-cell receptor CD23 increase its affinity for human immunoglobulin E. J Biol Chem 2013;288:21667-77.
33. Dhaliwal B, Yuan D, Pang MO, Henry AJ, Cain K, Oxbrow A, et al. Crystal structure of IgE bound to its B-cell receptor CD23 reveals a mechanism of reciprocal allosteric inhibition with high affinity receptor FcepsilonRI. Proc Natl Acad Sci U S A 2012;109:12686-91.
34. Botos I, Segal DM, Davies DR. The structural biology of Toll-like receptors. Structure 2011;19:447-59.
35. Bocskei Z, Groom CR, Flower DR, Wright CE, Phillips SE, Cavaggioni A, et al. Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. Nature 1992;360:186-8.
36. Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Joost van Neerven RJ, et al. X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. Nat Struct Biol 1996;3:1040-5.
37. Westman M, Lupinek C, Bousquet J, Andersson N, Pahr S, Baar A, et al. Early childhood IgE reactivity to pathogenesis-related class 10 proteins predicts allergic rhinitis in adolescence. J Allergy Clin Immunol 2015;135: 1199-206.e11.
38. Chapman MD, Pomés A, Breiteneder H, Ferreira F. Nomenclature and structural biology of allergens. J Allergy Clin Immunol 2007;119:414-20.
39. Radauer C, Nandy A, Ferreira F, Goodman RE, Larsen JN, Lidholm J, et al. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. Allergy 2014;69:413-9.
40. Wagner S, Radauer C, Bublin M, Hoffmann-Sommergruber K, Kopp T, Greisenegger EK, et al. Naturally occurring hypoallergenic Bet v 1 isoforms fail to induce IgE responses in individuals with birch pollen allergy. J Allergy Clin Immunol 2008;121:246-52.

41. Wallner M, Hauser M, Himly M, Zaborsky N, Mutschlechner S, Harrer A, et al. Reshaping the Bet v 1 fold modulates T(H) polarization. J Allergy Clin Immunol 2011;127:1571-8.
42. Radauer C, Bublin M, Wagner S, Mari A, Breiteneder H. Allergens are distributed into few protein families and possess a restricted number of biochemical functions. J Allergy Clin Immunol 2008;121:847-52.
43. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res 2014;42:D222-30.
44. Llewellyn R, Eisenberg DS. Annotating proteins with generalized functional linkages. Proc Natl Acad Sci U S A 2008;105:17700-5.
45. Martin AC, Orengo CA, Hutchinson EG, Jones S, Karmirantzou M, Laskowski RA, et al. Protein folds and functions. Structure 1998;6:875-84.
46. Chruszcz M, Chapman MD, Vailes LD, Stura EA, Saint-Remy JM, Minor W, et al. Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. J Mol Biol 2009;386:520-30.
47. Chruszcz M, Pomés A, Glesner J, Vailes LD, Osinski T, Porebski PJ, et al. Molecular determinants for antibody binding on group 1 house dust mite allergens. J Biol Chem 2012;287:7388-98.
48. Pomés A, Chapman MD, Vailes LD, Blundell TL, Dhanaraj V. Cockroach allergen Bla g 2: structure, function, and implications for allergic sensitization. Am J Respir Crit Care Med 2002;165:391-7.
49. Gustchina A, Li M, Wünschmann S, Chapman MD, Pomés A, Wlodawer A. Crystal structure of cockroach allergen Bla g 2, an unusual zinc binding aspartic protease with a novel mode of self-inhibition. J Mol Biol 2005;348:433-44.
50. Hindley J, Wünschmann S, Satinover SM, Woodfolk JA, Chew FT, Chapman MD, et al. Bla g 6: a troponin C allergen from *Blattella germanica* with IgE binding calcium dependence. J Allergy Clin Immunol 2006;117:1389-95.
51. Verdino P, Westritschnig K, Valenta R, Keller W. The cross-reactive calcium-binding pollen allergen, Phl p 7, reveals a novel dimer assembly. EMBO J 2002;21:5007-16.
52. Chruszcz M, Maleki SJ, Majorek KA, Demas M, Bublin M, Solberg R, et al. Structural and immunologic characterization of Ara h 1, a major peanut allergen. J Biol Chem 2011;286:39318-27.
53. Maleki SJ, Kopper RA, Shin DS, Park CW, Compadre CM, Sampson H, et al. Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation. J Immunol 2000;164:5844-9.
54. Nicolaou N, Murray C, Belgrave D, Poorafshar M, Simpson A, Custovic A. Quantification of specific IgE to whole peanut extract and peanut components in prediction of peanut allergy. J Allergy Clin Immunol 2011;127:684-5.
55. Dang TD, Tang M, Choo S, Licciardi PV, Koplin JJ, Martin PE, et al. Increasing the accuracy of peanut allergy diagnosis by using Ara h 2. J Allergy Clin Immunol 2012;129:1056-63.
56. Klemans RJ, van Os-Medendorp H, Blankestijn M, Bruijnzeel-Koomen CA, Knol EF, Knulst AC. Diagnostic accuracy of specific IgE to components in diagnosing peanut allergy: a systematic review. Clin Exp Allergy 2015;45:720-30.
57. Chruszcz M, Chapman MD, Osinski T, Solberg R, Demas M, Porebski PJ, et al. *Alternaria alternata* allergen Alt a 1: a unique beta-barrel protein dimer found exclusively in fungi. J Allergy Clin Immunol 2012;130:241-7.
58. Randall TA, Perera L, London RE, Mueller GA. Genomic, RNAseq, and molecular modeling evidence suggests that the major allergen domain in insects evolved from a homodimeric origin. Genome Biol Evol 2013;5:2344-58.
59. Hamiaux C, Maddumage R, Middleditch MJ, Prakash R, Brummell DA, Baker EN, et al. Crystal structure of kiwellin, a major cell-wall protein from kiwifruit. J Struct Biol 2014;187:276-81.
60. Kaiser L, Gronlund H, Sandalova T, Ljunggren HG, van Hage-Hamsten M, Achour A, et al. The crystal structure of the major cat allergen Fel d 1, a member of the secretoglobin family. J Biol Chem 2003;278:37730-5.
61. Kaiser L, Velickovic TC, Badia-Martinez D, Adedoyin J, Thunberg S, Hallen D, et al. Structural characterization of the tetrameric form of the major cat allergen Fel d 1. J Mol Biol 2007;370:714-27.
62. Derewenda U, Li J, Derewenda Z, Dauter Z, Mueller GA, Rule GS, et al. The crystal structure of a major dust mite allergen Der p 2, and its biological implications. J Mol Biol 2002;318:189-97.
63. Park BS, Song DH, Kim HM, Choi BS, Lee H, Lee JO. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. Nature 2009;458:1191-5.
64. Trompette A, Divanovic S, Visintin A, Blanchard C, Hegde RS, Madan R, et al. Allergenicity resulting from functional mimicry of a Toll-like receptor complex protein. Nature 2009;457:585-8.
65. Karp CL. Guilt by intimate association: what makes an allergen an allergen? J Allergy Clin Immunol 2010;125:955-60.
66. Eisenbarth SC, Piggott DA, Huleatt JW, Visintin I, Herrick CA, Bottomly K. Lipopolysaccharide-enhanced, toll-like receptor 4-dependent T helper cell type 2 responses to inhaled antigen. J Exp Med 2002;196:1645-51.
67. Piggott DA, Eisenbarth SC, Xu L, Constant SL, Huleatt JW, Herrick CA, et al. MyD88-dependent induction of allergic Th2 responses to intranasal antigen. J Clin Invest 2005;115:459-67.
68. Redecke V, Hacker H, Datta SK, Fermin A, Pitha PM, Broide DH, et al. Cutting edge: activation of Toll-like receptor 2 induces a Th2 immune response and promotes experimental asthma. J Immunol 2004;172:2739-43.
69. Shakib F, Ghaemmaghami AM, Sewell HF. The molecular basis of allergenicity. Trends Immunol 2008;29:633-42.
70. Chapman MD, Wunschmann S, Pomés A. Proteases as Th2 adjuvants. Curr Allergy Asthma Rep 2007;7:363-7.
71. Schulz O, Laing P, Sewell HF, Shakib F. Der p I, a major allergen of the house dust mite, proteolytically cleaves the low-affinity receptor for human IgE (CD23). Eur J Immunol 1995;25:3191-4.
72. Hewitt CR, Brown AP, Hart BJ, Pritchard DI. A major house dust mite allergen disrupts the immunoglobulin E network by selectively cleaving CD23: innate protection by antiproteases. J Exp Med 1995;182:1537-44.
73. King C, Brennan S, Thompson PJ, Stewart GA. Dust mite proteolytic allergens induce cytokine release from cultured airway epithelium. J Immunol 1998;161:3645-51.
74. Wan H, Winton HL, Soeller C, Tovey ER, Gruenert DC, Thompson PJ, et al. Der p 1 facilitates transepithelial allergen delivery by disruption of tight junctions. J Clin Invest 1999;104:123-33.
75. Huang S, Rutkowsky JM, Snodgrass RG, Ono-Moore KD, Schneider DA, Newman JW, et al. Saturated fatty acids activate TLR-mediated proinflammatory signaling pathways. J Lipid Res 2012;53:2002-13.
76. Flower DR. The lipocalin protein family: structure and function. Biochem J 1996;318:1-14.
77. Vance SJ, McDonald RE, Cooper A, Smith BO, Kennedy MW. The structure of latherin, a surfactant allergen protein from horse sweat and saliva. J R Soc Interface 2013;10:20130453.
78. Bublin M, Eiwegger T, Breiteneder H. Do lipids influence the allergic sensitization process? J Allergy Clin Immunol 2014;134:521-9.
79. Passalacqua G, Pasquali M, Ariano R, Lombardi C, Giardini A, Baiardini I, et al. Randomized double-blind controlled study with sublingual carbamylated allergoid immunotherapy in mild rhinitis due to mites. Allergy 2006;61:849-54.
80. Cromwell O, Hafner D, Nandy A. Recombinant allergens for specific immunotherapy. J Allergy Clin Immunol 2011;127:865-72.
81. Meyer W, Narkus A, Salapatek AM, Hafner D. Double-blind, placebo-controlled, dose-ranging study of new recombinant hypoallergenic Bet v 1 in an environmental exposure chamber. Allergy 2013;68:724-31.
82. Smith AM, Chapman MD, Taketomi EA, Platts-Mills TA, Sung SS. Recombinant allergens for immunotherapy: a Der p 2 variant with reduced IgE reactivity retains T-cell epitopes. J Allergy Clin Immunol 1998;101:423-5.
83. Takai T, Yokota T, Yasue M, Nishiyama C, Yuuki T, Mori A, et al. Engineering of the major house dust mite allergen Der f 2 for allergen-specific immunotherapy. Nat Biotechnol 1997;15:754-8.
84. Mirza O, Henriksen A, Ipsen H, Larsen JN, Wissenbach M, Spangfort MD, et al. Dominant epitopes and allergic cross-reactivity: complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet v 1. J Immunol 2000;165:331-8.
85. Spangfort MD, Mirza O, Ipsen H, Van Neerven RJ, Gajhede M, Larsen JN. Dominating IgE-binding epitope of Bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. J Immunol 2003;171:3084-90.
86. Padavattan S, Schirmer T, Schmidt M, Akdis C, Valenta R, Mittermann I, et al. Identification of a B-cell epitope of hyaluronidase, a major bee venom allergen, from its crystal structure in complex with a specific Fab. J Mol Biol 2007;368:742-52.
87. Niemi M, Jylha S, Laukkanen ML, Soderlund H, Makinen-Kiljunen S, Kallio JM, et al. Molecular interactions between a recombinant IgE antibody and the beta-lactoglobulin allergen. Structure 2007;15:1413-21.
88. Padavattan S, Flicker S, Schirmer T, Madritsch C, Randow S, Reese G, et al. High-affinity IgE recognition of a conformational epitope of the major respiratory allergen Phl p 2 as revealed by X-ray crystallography. J Immunol 2009;182:2141-51.
89. Li M, Gustchina A, Alexandratos J, Wlodawer A, Wunschmann S, Kepley CL, et al. Crystal structure of a dimerized cockroach allergen Bla g 2 complexed with a monoclonal antibody. J Biol Chem 2008;283:22806-14.
90. Li M, Gustchina A, Glesner J, Wunschmann S, Vailes LD, Chapman MD, et al. Carbohydrates contribute to the interactions between cockroach allergen Bla g 2 and a monoclonal antibody. J Immunol 2011;186:333-40.

91. Osinski T, Pomés A, Majorek KA, Glesner J, Offermann LR, Vailes LD, et al. Structural analysis of Der p 1-antibody complexes and comparison with complexes of proteins or peptides with monoclonal antibodies. J Immunol 2015; In press.
92. Glesner J, Wunschmann S, Li M, Gustchina A, Wlodawer A, Himly M, et al. Mechanisms of allergen-antibody interaction of cockroach allergen Bla g 2 with monoclonal antibodies that inhibit IgE antibody binding. PLoS One 2011;6:e22223.
93. Sastre J, Landivar ME, Ruiz-Garcia M, Andregnette-Rosigno MV, Mahillo I. How molecular diagnosis can change allergen-specific immunotherapy prescription in a complex pollen area. Allergy 2012;67:709-11.
94. Stringari G, Tripodi S, Caffarelli C, Dondi A, Asero R, Di Rienzo BA, et al. The effect of component-resolved diagnosis on specific immunotherapy prescription in children with hay fever. J Allergy Clin Immunol 2014;134:75-81.
95. Tripodi S, Frediani T, Lucarelli S, Macri F, Pingitore G, Di Rienzo BA, et al. Molecular profiles of IgE to *Phleum pratense* in children with grass pollen allergy: implications for specific immunotherapy. J Allergy Clin Immunol 2012;129:834-9.
96. Le TM, Bublin M, Breiteneder H, Fernandez-Rivas M, Asero R, Ballmer-Weber B, et al. Kiwifruit allergy across Europe: clinical manifestation and IgE recognition patterns to kiwifruit allergens. J Allergy Clin Immunol 2013;131:164-71.
97. Lieberman JA, Glaumann S, Batelson S, Borres MP, Sampson HA, Nilsson C. The utility of peanut components in the diagnosis of IgE-mediated peanut allergy among distinct populations. J Allergy Clin Immunol Pract 2013;1:75-82.
98. Aalberse RC. Structural features of allergenic molecules. Chem Immunol Allergy 2006;91:134-46.
99. Hazebrouck S, Ah-Leung S, Bidat E, Paty E, Drumare MF, Tilleul S, et al. Goat's milk allergy without cow's milk allergy: suppression of non-cross-reactive epitopes on caprine beta-casein. Clin Exp Allergy 2014;44:602-10.
100. Saarelainen S, Rytkonen-Nissinen M, Rouvinen J, Taivainen A, Auriola S, Kauppinen A, et al. Animal-derived lipocalin allergens exhibit immunoglobulin E cross-reactivity. Clin Exp Allergy 2008;38:374-81.
101. D'Avino R, Bernardi ML, Wallner M, Palazzo P, Camardella L, Tuppo L, et al. Kiwifruit Act d 11 is the first member of the ripening-related protein family identified as an allergen. Allergy 2011;66:870-7.
102. Bublin M, Kostadinova M, Radauer C, Hafner C, Szepfalusi Z, Varga EM, et al. IgE cross-reactivity between the major peanut allergen Ara h 2 and the nonhomologous allergens Ara h 1 and Ara h 3. J Allergy Clin Immunol 2013;132:118-24.
103. Power TD, Ivanciuc O, Schein CH, Braun W. Assessment of 3D models for allergen research. Proteins 2013;81:545-54.
104. Aalberse RC, Akkerdaas J, van Ree R. Cross-reactivity of IgE antibodies to allergens. Allergy 2001;56:478-90.
105. Mueller GA, Pedersen LC, Glesner J, Edwards LL, Zakzuk J, London RE, et al. Analysis of glutathione S-transferase allergen cross-reactivity in a North American population: relevance for molecular diagnosis. J Allergy Clin Immunol 2015;136:140-50, e1-7.
106. Kupitz C, Basu S, Grotjohann I, Fromme R, Zatsepin NA, Rendek KN, et al. Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. Nature 2014;513:261-5.
107. McSweeney S, Fromme P. Crystallography: sources of inspiration. Nature 2014;505:620-1.
108. Zimmerman MD, Grabowski M, Domagalski MJ, Maclean EM, Chruszcz M, Minor W. Data management in the modern structural biology and biomedical research environment. Methods Mol Biol 2014;1140:1-25.

DETERMINATION OF X-RAY CRYSTAL STRUCTURES OF BIOLOGICAL MOLECULES

Proteins are first purified from the natural source or from an *in vitro* expression system in high amounts (5-10 mg) to be used in crystallization experiments. An inherent variability associated with some natural or recombinant proteins is one of the main difficulties to overcome and might be a consequence of different factors that will be illustrated here with allergens as examples.

MOLECULAR POLYMORPHISMS

A common source of allergen variability is the presence of molecular polymorphisms that can impair crystallization if copurified from the natural source. As a guideline, isoallergens share amino acid identity of greater than 67%, and each isoallergen might have multiple forms of closely similar sequences (>90%) that are designated as variants or isoforms. This guideline was established by the WHO/IUIS Allergen Nomenclature Sub-Committee (www.allergen.org).[E1] The heterogeneity generated by the presence of these polymorphisms, especially by isoallergens, can impair crystallization, although not always. For example, despite a large number of polymorphisms reported for group 1 mite allergens (24 variants for Der p 1 and 10 for Der f 1), only 1 variant of each (Der p 1.0105 and Der f 1.0101) could be crystallized and used to solve the structures.[E2]

GLYCOSYLATION

Molecular heterogeneity can result from glycosylation through addition of carbohydrates of different lengths to the side chains of proteins. Removal of carbohydrates by means of enzymatic digestion followed by purification of the deglycosylated allergen often is not efficient for production of high amounts of protein required for crystallography. Instead, N-glycosylation can be prevented by expressing recombinant proteins in *Escherichia coli*, which are not glycosylated, or by modifying the N-glycosylation sites to prevent glycosylation when expressing proteins in the yeast eukaryotic system *Pichia pastoris*, which tend to be hyperglycosylated. Such amino acid substitutions were introduced in Bla g 2 (N93Q), recombinant pro–Der p 1 (N132E), and recombinant Der p 1 (N52Q) to favor crystallization.[E3-E5]

PROTEOLYSIS

Molecular heterogeneity caused by proteolysis is a common occurrence, especially when proteolytic enzymes (including allergens themselves) are present in the allergen source. An example is Der p 1, a cysteine protease that is expressed as a proenzyme and becomes enzymatically active after maturation, involving removal of 80 amino acids that cover the active site. The first structure of this allergen was recombinant C114A pro–Der p 1 (prevented from spontaneous maturation by mutating the active site cysteine).[E4] A second structure was determined with active, mature recombinant Der p 1.[E5] In this case allergen dimerization took place during crystal growth, and fortuitously, a loop from one molecule sterically blocked the catalytic cleft in the second molecule.[E6] In the most recently determined structures of native Der p 1 and Der f 1, the catalytic cysteine was found to be oxidized, which might have contributed to the success of crystallization by rendering the molecule inactive.[E6]

FRAGMENTATION

Some proteins break down into fragments that preserve their allergenic potential. The cockroach allergen Bla g 1 is one of them. This allergen is formed by approximately 100 consecutive amino acid repeats resulting from duplication and mutation of an original domain present only in other insect proteins involved in digestive or detoxifying functions.[E7,E8] Bla g 1 breaks into fragments containing different numbers of repeats. The basic structural unit of Bla g 1 was the first to be determined for this group of proteins and revealed a novel fold containing 2 consecutive repeats with a large internal hydrophobic cavity that contains lipids (Fig 3, *D*). The allergen is formed by several linked consecutive units. Despite fragmentation, the basic unit of Bla g 1 preserves the fold that exposes the epitopes involved in IgE antibody binding.

AGGREGATION-PRECIPITATION

Protein solubility at a high protein concentration (typically >3 mg/mL) is one of the usual prerequisites for crystallization. However, some proteins will not reach high concentrations without precipitating or aggregating. For example, proteins expressed in *E coli* do not always fold properly and might precipitate in inclusion bodies inside cells. A technique based on successive dialysis of these denatured proteins into buffers of different ionic strength helps them to refold into their native conformation. This has been applied efficiently for dust mite group 2 allergens.[E9,E10] Another successful approach for proteins that did not express well alone or precipitated during purification is the use of fusion partners. Small allergens have shown an improved capacity to crystallize when expressed as fusions with a larger protein that stabilizes the construct and might contribute to crystal lattice formation. This approach has been successfully applied to the expression and crystallization of allergens fused to maltose-binding protein (MBP) and a variant of green fluorescent protein (GFP; ie, MBP–Der p 7, MBP–Ara h 2, and GFP–Bla g 1).[E8,E11,E12] The fused proteins might be optimized for different purposes. For example, the MBP was designed with surface entropy-reducing mutations to enhance crystallization.[E13,E14] Another advantage of the fused proteins is to facilitate purification because of their affinity for certain column resins (MBP binds amylose affinity resin, GST binds glutathione S-transferase columns, and the polyhistidine tag attached to GFP binds nickel agarose resin). Fusion proteins can be designed to allow isolation of the allergen by using a specific protease, such as tobacco etch virus protease.[E8]

MOLECULAR FLEXIBILITY

Some allergens have a regulatory function that involves molecular flexibility, which is determined by calcium binding to EF-hand motifs.[E15,E16] Nuclear magnetic resonance, which analyzes structures in solution, might be a more appropriate method for analysis of the dynamic structural changes.[E17] Nevertheless, x-ray crystallography provided useful data for these regulatory proteins once certain conformations were stabilized with the appropriate calcium concentrations.[E18]

ADDITIONAL MOLECULAR CHALLENGES FOR CRYSTALLOGRAPHY

Some proteins, such as receptors, are embedded in membranes, and their isolation requires specialized techniques, such as the use

of detergents, or expression of only their extracellular fragments. They are usually complex macromolecules involving multiple subunits. Examples of these molecules involved in allergic disease include cytokines and TLRs and B- and T-cell receptors (Fig 2).

REFERENCES

E1. Radauer C, Nandy A, Ferreira F, Goodman RE, Larsen JN, Lidholm J, et al. Update of the WHO/IUIS Allergen Nomenclature Database based on analysis of allergen sequences. Allergy 2014;69:413-9.

E2. Chruszcz M, Pomés A, Glesner J, Vailes LD, Osinski T, Porebski PJ, et al. Molecular determinants for antibody binding on group 1 house dust mite allergens. J Biol Chem 2012;287:7388-98.

E3. Gustchina A, Li M, Wünschmann S, Chapman MD, Pomés A, Wlodawer A. Crystal structure of cockroach allergen Bla g 2, an unusual zinc binding aspartic protease with a novel mode of self-inhibition. J Mol Biol 2005;348:433-44.

E4. Meno K, Thorsted PB, Ipsen H, Kristensen O, Larsen JN, Spangfort MD, et al. The crystal structure of recombinant proDer p 1, a major house dust mite proteolytic allergen. J Immunol 2005;175:3835-45.

E5. de Halleux S, Stura E, VanderElst L, Carlier V, Jacquemin M, Saint-Remy JM. Three-dimensional structure and IgE-binding properties of mature fully active Der p 1, a clinically relevant major allergen. J Allergy Clin Immunol 2006; 117:571-6.

E6. Chruszcz M, Pomés A, Osinski T, Majorek KA, Glesner J, Minor W, et al. Structural analysis reveals molecular basis for interactions of group 1 allergens with species specific and cross-reactive antibodies. J Allergy Clin Immunol 2013;131:AB15.

E7. Pomés A, Melén E, Vailes LD, Retief JD, Arruda LK, Chapman MD. Novel allergen structures with tandem amino acid repeats derived from German and American cockroach. J Biol Chem 1998;273:30801-7.

E8. Mueller GA, Pedersen LC, Lih FB, Glesner J, Moon AF, Chapman MD, et al. The novel structure of the cockroach allergen Bla g 1 has implications for allergenicity and exposure assessment. J Allergy Clin Immunol 2013;132: 1420-6.

E9. Derewenda U, Li J, Derewenda Z, Dauter Z, Mueller GA, Rule GS, et al. The crystal structure of a major dust mite allergen Der p 2, and its biological implications. J Mol Biol 2002;318:189-97.

E10. Nishiyama C, Yuuki T, Usui Y, Iwamoto N, Okumura Y, Okudaira H. Effects of amino acid variations in recombinant Der fII on its human IgE and mouse IgG recognition. Int Arch Allergy Immunol 1994;105:62-9.

E11. Mueller GA, Edwards LL, Aloor JJ, Fessler MB, Glesner J, Pomés A, et al. The structure of the dust mite allergen Der p 7 reveals similarities to innate immune proteins. J Allergy Clin Immunol 2010;125:909-17.

E12. Mueller GA, Gosavi RA, Pomés A, Wunschmann S, Moon AF, London RE, et al. Ara h 2: crystal structure and IgE binding distinguish two subpopulations of peanut allergic patients by epitope diversity. Allergy 2011;66:878-85.

E13. Derewenda ZS. Rational protein crystallization by mutational surface engineering. Structure 2004;12:529-35.

E14. Cormack BP, Bertram G, Egerton M, Gow NA, Falkow S, Brown AJ. Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans*. Microbiology 1997;143:303-11.

E15. Valenta R, Hayek B, Seiberler S, Bugajska-Schretter A, Niederberger V, Twardosz A, et al. Calcium-binding allergens: from plants to man. Int Arch Allergy Immunol 1998;117:160-6.

E16. Hindley J, Wünschmann S, Satinover SM, Woodfolk JA, Chew FT, Chapman MD, et al. Bla g 6: a troponin C allergen from *Blattella germanica* with IgE binding calcium dependence. J Allergy Clin Immunol 2006;117:1389-95.

E17. Henzl MT, Sirianni AG, Wycoff WG, Tan A, Tanner JJ. Solution structures of polcalcin Phl p 7 in three ligation states: Apo-, hemi-Mg2+-bound, and fully Ca2+-bound. Proteins 2013;81:300-15.

E18. Vinogradova MV, Stone DB, Malanina GG, Karatzaferi C, Cooke R, Mendelson RA, et al. Ca(2+)-regulated structural changes in troponin. Proc Natl Acad Sci U S A 2005;102:5038-43.

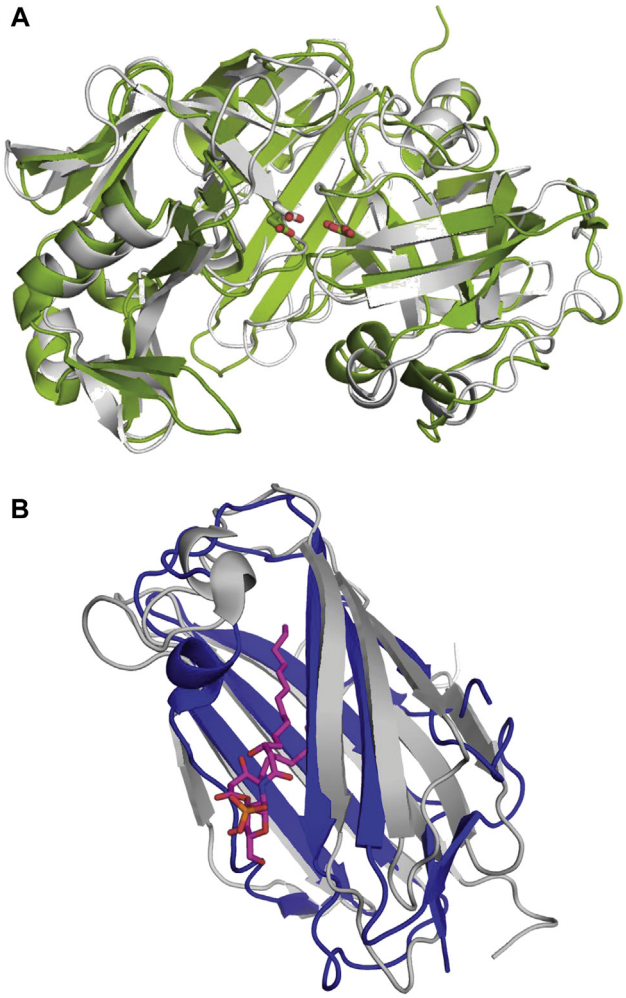

FIG E1. Structural alignment of Bla g 2 and Der p 2 with the respective homologous porcine pepsin and human MD-2. **A,** Bla g 2 *(green)* aligned to a structure of pepsin determined in 1990 at 1.80 Å. Aspartates 32 and 215 in the catalytic site of pepsin are at a closer distance than the equivalent aspartates in Bla g 2 (indicated in stick representation). **B,** Der p 2 *(blue)* aligned to MD-2. LPS bound to the MD-2 internal cavity is in *fuchsia*. Protein Data Bank accession codes are 3LIZ, 4PEP, 1KTJ, and 2E59, respectively. Pepsin and MD-2 are displayed in *gray*.

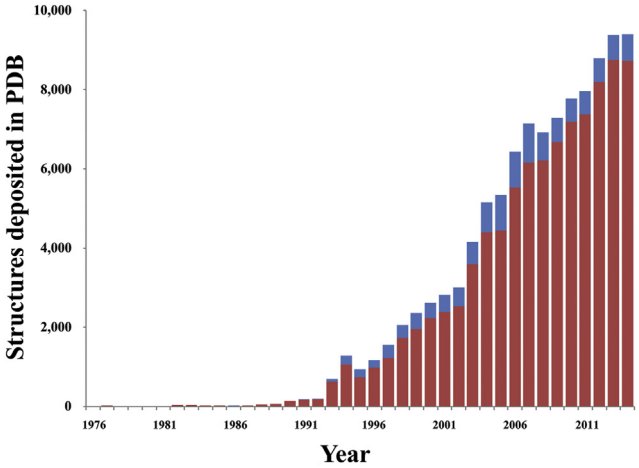

FIG E2. Exponential increase in the number of structures deposited to the Protein Data Bank *(PDB)* between 1976 and 2014. All the structures (*red* and *blue*) and the structures determined by using x-ray crystallography *(red)* are shown.

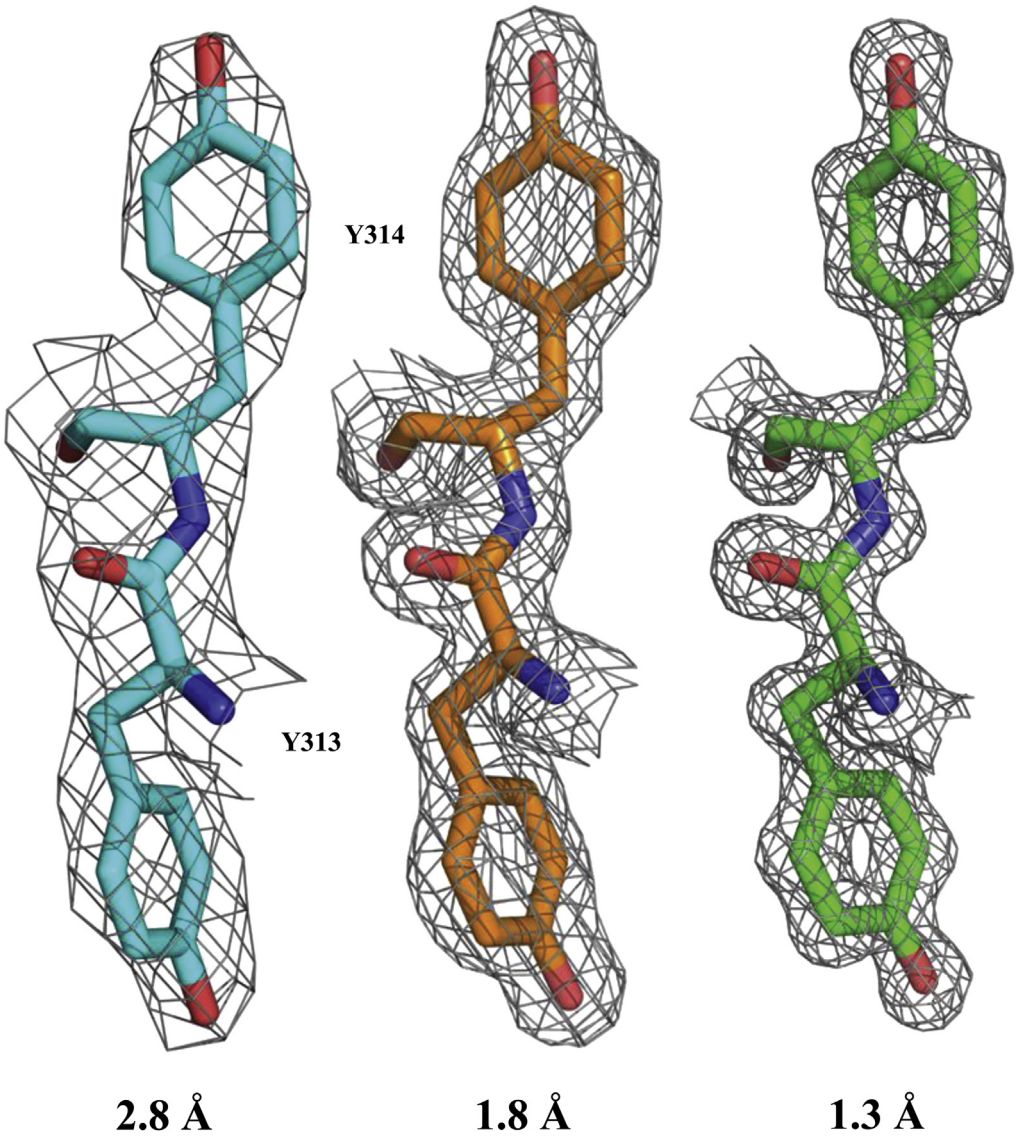

FIG E3. Resolution. Electron density maps calculated for x-ray crystal structures determined at higher resolution are much more detailed, allowing a more precise definition of the molecules under study. An example is provided by the maps calculated for 2 amino acids in 3 structures of Bla g 2 determined at lower *(left)* to higher *(right)* resolution. The corresponding structures are 2NR6, 3LIZ, and 1YG9.

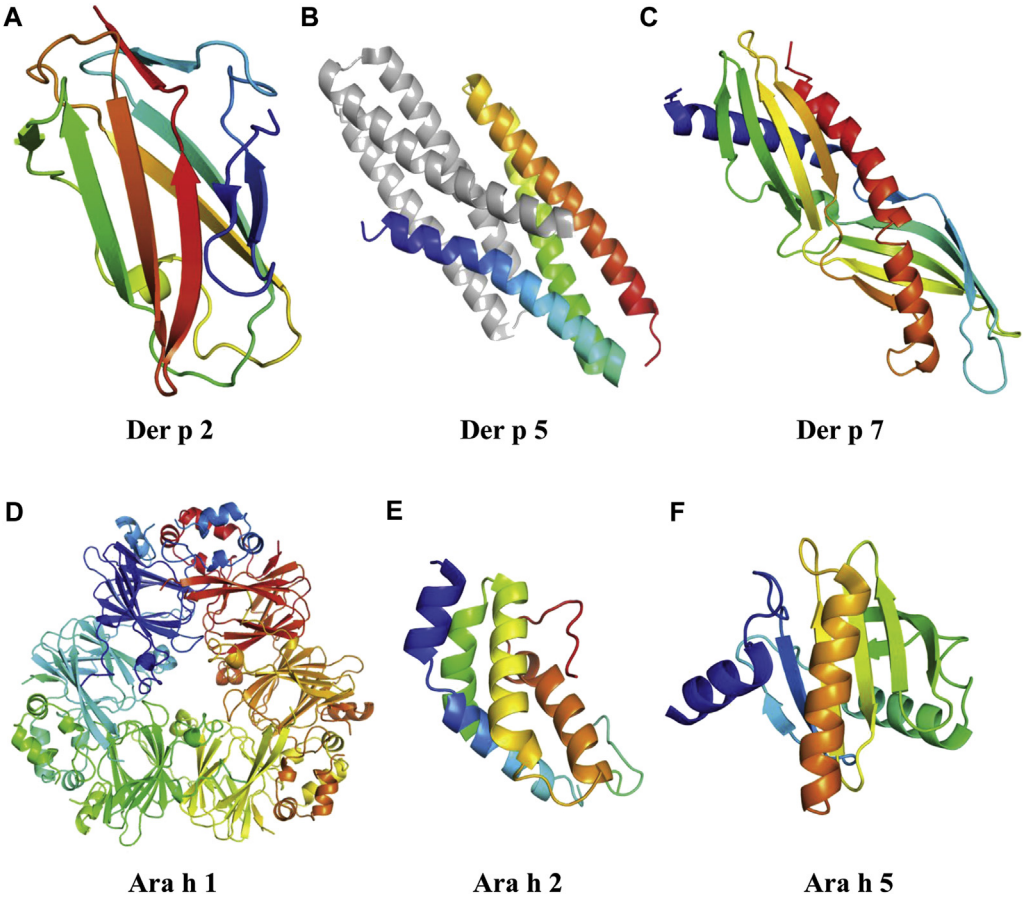

FIG E4. Allergen structures from mite and peanut allergens. **A,** Der p 2 (1KTJ). **B,** Der p 5 (3MQ1). **C,** Der p 7 (3H4Z). **D,** Ara h 1 (3S7I). **E,** Ara h 2 (3OB4). **F,** Ara h 5 (4ESP). Allergens are shown in rainbow colors from the N-terminus *(blue)* to C-terminus *(red)*.

TABLE E1. Tertiary structures of inhalant allergens

Allergen	Species of origin	Function/structure	PDB ID codes of 3D models based on x-ray data								PDB ID codes of 3D models based on NMR data		
Inhaled, indoor													
Blo t 19	*Blomia tropicalis*	Antimicrobial peptide homologue									2MFJ		
Blo t 12	*Blomia tropicalis*	Chitin-binding protein									2MFK		
Der f 1	*Dermatophagoides farinae*	Cysteine protease	3D6S	3RVV*									
Der p 1	*Dermatophagoides pteronyssinus*	Cysteine protease	1XKG	2AS8(m)	3F5V	3RVW*	3RVX*	4PP1*	4PP1*	4PP2*			
Der f 13	*Dermatophagoides farinae*	Fatty acid binding protein									2A0A		
Bla g 5	*Blattella germanica*	Glutathione S-transferase	5Q5R										
Blo t 8	*Blomia tropicalis*	Glutathione S-transferase	4Q5N										
Der p 8	*Dermatophagoides pteronyssinus*	Glutathione S-transferase	4Q5Q										
Bla g 1	*Blattella germanica*	Gut protein that carries lipids	4JRB										
Bla g 2	*Blattella germanica*	Inactive aspartic protease	1YG9(m)	2NR6(m)*	3LIZ(m)*	4RLD(m)							
Der f 2	*Dermatophagoides farinae*	Lipid-binding protein	1XWV	2F08							1AHK	1AHM	1WRF
Der p 2	*Dermatophagoides pteronyssinus*	Lipid-binding protein	1KTJ(m)								1A9V(m)		
Bla g 4	*Blattella germanica*	Lipocalin	3EBK	4N7D									
Bos d 2	*Bos domesticus*	Lipocalin	1BJ7	4WFU	4WFV								
Can f 2	*Canis familiaris*	Lipocalin	3L4R										
Can f 4	*Canis familiaris*	Lipocalin	4ODD										
Mus m 1	*Mus musculus*	Lipocalin	1MUP	1JV4							1DF3		
Per a 4	*Periplaneta americana*	Lipocalin	3EBW										
Rat n 1	*Rattus norvegicus*	Lipocalin	2A2G	2A2U									
Der f 7	*Dermatophagoides farinae*	LPS-binding protein-like	3UV1										
Der p 7	*Dermatophagoides pteronyssinus*	LPS-binding protein-like	3H4Z										
Blo t 5	*Blomia tropicalis*	Structural 3-helix bundle									2JMH	2JRK	2MEY
Blo t 21	*Blomia tropicalis*	Structural 3-helix bundle									2LM9		
Der p 5	*Dermatophagoides pteronyssinus*	Structural 3-helix bundle	3MQ1										
Fel d 1	*Felis domesticus*	Uteroglobin	1PUO(m)	1ZKR(m)	2EJN(m)								
Inhaled, outdoor													
Zea m 1	*Zea mays*	β-Expansin	2HCZ										
Heb v 2	*Hevea brasiliensis*	β-1,3-Glucanase	4HPG	4IIS									
Ole e 9	*Olea europaea*	β-1,3-Glucanase (C-terminus)									2JON		
Asp f 11	*Aspergillus fumigatus*	Cyclophilin: peptidyl-propyl isomerase	2C3B										
Art v 1	*Artemisia vulgaris*										2KPY		

(Continued)

TABLE E1. *(Continued)*

Allergen	Species of origin	Function/structure	PDB ID codes of 3D models based on x-ray data												PDB ID codes of 3D models based on NMR data		
		Defensin fold with polyproline domain															
Phl p 1	*Phleum pratense*	Expansin	1N10														
Phl p 2	*Phleum pratense*	Grass group II/III	1WHO	1WHP	2VXQ*										1BMW		
Phl p 3	*Phleum pratense*	Grass group II/III	3FT1	3FT9											2JNZ		
Chi t 1	*Chironomus thummi thummi*	Hemoglobin	1ECO														
Hev b 6	*Hevea brasiliensis*	Hevein precursor	1Q9B	1WKX											1T0W	1HEV	
Equ c 1	*Equus caballus*	Lipocalin	1EW3														
Asp f 6	*Aspergillus fumigatus*	Manganese superoxide dismutase	1KKC														
Cla h 8	*Cladosporium herbarum*	Mannitol dehydrogenase	3GDF	3GDG													
Asp f 1	*Aspergillus fumigatus*	Mitogillin	1AQZ														
Phl p 4	*Phleum pratense*	Oxidoreductase (berberine bridge enzyme)	3TSJ	3TSH	4PVE	4PVJ	4PVK	4PWB	4PWC								
Bet v 1	*Betula verrucosa*	Pathogenesis-related protein (PR-10)	1BV1	1QMR(m)	1FM4	1LLT(m)	1FSK*	3K78	4QIP	4MNS	4BTZ	4BK7	4BKC	4BKD	1BTV	1B6F	
Jun a 1	*Juniperus ashei*	Pectate lyase	1PXZ														
Bet v 4	*Betula verrucosa*	Polcalcin (Ca binding protein)													1H4B		
Che a 3	*Chenopodium album*	Polcalcin (calcium-binding protein)	2OPO														
Phl p 7	*Phleum pratense*	Polcalcin (calcium-binding protein)	1K9U												2LVK	2LVJ	2LVI
Ara t 8	*Arabidopsis thaliana*	Profilin	1A0K	3NUL													
Bet v 2	*Betula verrucosa*	Profilin	1CQA														
Hev b 8	*Hevea brasiliensis*	Profilin (Latex)	1G5U														
Equ c 3	*Equus caballus*	Serum albumin	3V08														
Asp o 21	*Aspergillus oryzae*	TAKA-amylase A	2TAA	6TAA	7TAA												
Alt a 1	*Alternaria alternata*	Unknown	3V0R	4AUD													
Amb t 5	*Ambrosia trifida*	Unknown													1BBG	2BBG	3BBG
Ole e 6	*Olea europaea*	Unknown													1SS3		
Phl p 6	*Phleum pratense*	Unknown	1NLX														
Phl p 5	*Phleum pratense*	Unknown: Phl p 5b	1L3P														

The codes refer to the structure of molecules that are not necessarily the same polymorphism reported in the WHO/IUIS Allergen Nomenclature database.
3D, Three-dimensional; *(m)*, modified or mutated molecule; *NMR*, nuclear magnetic resonance; *PDB*, Protein Data Bank.
*Allergen in complex with an antibody fragment.

TABLE E2. Tertiary structures of food, venom, and contact allergens

Allergen	Species of origin	Function/structure	PDB ID codes of 3D models based on x-ray data										PDB ID codes of 3D models based on NMR data	
Ingested, food														
Ara h 3	*Arachis hypogaea*	11S globulin (legumin-type, glycinin, cupin)	3C3V											
Gly m 6	*Glycine max*	11S globulin (legumin-type, glycinin)	1FXZ	1OD5	2D5H	2D5F								
Pru du 6	*Prunus dulcis*	11S globulin (legumin-type, amandin)	3FZ3											
Ara h 2	*Arachis hypogaea*	2S albumin (conglutin)	3OB4											
Ara h 6	*Arachis hypogaea*	2S albumin (conglutin)											1W2Q	
Ber e 1	*Bertholletia excelsa*	2S albumin (sulfur-rich seed storage)											2LVF	
Bra n 1	*Brassica napus*	2S albumin (sulfur-rich seed storage)	1PNB											
Ric c 1	*Ricinus communis*	2S albumin (sulfur-rich seed storage)											1PSY	
Ara h 1	*Arachis hypogaea*	7S globulin (vicilin-type, cupin)	3S7E	3S7I	3SMH									
Gly m 5	*Glycine max*	7S globulin (vicilin-type, β-conglycinin)	1IPJ	1IPK	1UIJ(m)									
Pis s 1	*Pisum sativum*	7S globulin (vicilin-type)	1PNB	3KSC										
Tria a 18	*Triticum aestivum*	Agglutinin isolectin 1	4AML	2X3T	2UVO	2CWG	1WGC	2WGC	7WGA	9WGA				
Bos d 4	*Bos domesticus*	α-Lactalbumin	1F6R	1F6S	2G4N	1HFZ	1HFX	1HFY						
Mus a 5	*Musa acuminata*	β-1,3-Glucanase	2CYG											
Bos d 5	*Bos domesticus*	β-Lactoglobulin	1GX8	1GX9	1GXA	2AKQ	1BSO	1UZ2(m)	2R56*					
Act d 1 homolog	*Actinidia chinensis*	Cysteine protease	2ACT	1AEC										
Act c 5	*Actinidia chinensis*	Kiwellin	4PMK											
Gal d 4	*Gallus domesticus*	Lysozyme†	1LYZ	1H6M(m)	1YQV*	1FDL*	1MLC*	3HFM*	2A2Y*	1DQJ*			1GXV	1GXX
Pru p 3	*Prunus persica*	Nonspecific lipid transfer protein	2ALG	2B5S										
Tri a 14	*Triticum aestivum*	Nonspecific lipid transfer protein	1BWO										1CZ2	1GH1
Zea m 14	*Zea mays*	Nonspecific lipid transfer protein	1FK0	1FK1	1FK2	1FK3	1FK4	1FK5	1FK6	1FK7	1MZL	1MZM	1AFH	
Gal d 2	*Gallus domesticus*	Ovalbumin	1JTI(m)	1OVA	1UHG(m)									
Gal d 3	*Gallus domesticus*	Ovotransferrin	1RYX	2D3I	1OVT	1AIV	1TFA	1NNT	1IEJ	1N04				
Cyp c 1	*Cyprinus carpio*	Parvalbumin (beta; Carp)	4CPV	5CPV										
Gad m 1	*Gadus morhua*	Parvalbumin (beta; Cod)											2MBX	
Act d 11	*Actinidia deliciosa*	Pathogenesis-related protein (PR-10 [ripening protein, MLP/RRP])	4IGV	4IHR	4IGW	4IGX	4IGY	4IH0	4IH2					
Api g 1	*Apium graveolens*	Pathogenesis-related protein (PR-10)	2BK0											
Ara h 8	*Arachis hypogaea*	Pathogenesis-related protein (PR-10)	4MAP	4MA6	4M9W	4M9B								

(Continued)

TABLE E2. *(Continued)*

Allergen	Species of origin	Function/structure	PDB ID codes of 3D models based on x-ray data				PDB ID codes of 3D models based on NMR data	
Dau c 1	*Daucus carota*	Pathogenesis-related protein (PR-10)	2WQL					
Fra a 1	*Fragaria ananassa*	Pathogenesis-related protein (PR-10)					2LPX	
Gly m 4	*Glycine max*	Pathogenesis-related protein (PR-10)					2K7H	
Pru av 1	*Prunus avium*	Pathogenesis-related protein (PR-10)					1H2O(m)	1E09
Vig r 6	*Vigna radiata*	Pathogenesis-related protein, cytokinin-specific binding protein (CSBP)	2FLH	3C0V				
Ara h 5	*Arachis hypogaea*	Profilin	4ESP					
Bos d 6	*Bos domesticus*	Serum albumin	3V03	4F5S	4JK4			
Ani s 5	*Anisakis simplex*	SXP/RAL-2 family protein					2MAR	
Act d 2	*Actinidia deliciosa*	Thaumatin-like protein	4BCT					
Mal d 2	*Malus domestica*	Thaumatin-like protein	3ZS3					
Mus a 4	*Musa acuminata*	Thaumatin-like protein	1Z3Q					
Pru av 2	*Prunus avium*	Thaumatin-like protein	2AHN					
Injected								
Ves v 5	*Vespula vulgaris*	Antigen 5	1QNX					
Arg r 1	*Argas reflexus*	Histamine-binding salivary lipocalin (calycin)	2X45	2X46				
Api m 2	*Apis mellifera*	Hyaluronidase	1FCQ	1FCU	1FCV	2J88*		
Ves v 2	*Vespula vulgaris*	Hyaluronidase	2ATM					
Api m 4	*Apis mellifera*	Melittin	2MLT				1BH1(m)	
Api m 1	*Apis mellifera*	Phospholipase A2	1POC					
Aed a 2	*Aedes aegypti*	Salivary antigen	3DXL	3DY9	3DYE	3DZT		
Sol i 2	*Solenopsis invicta*	Transport of hydrophobic ligands	2YGU					
Sol i 3	*Solenopsis invicta*	Ves v 5-like	2VZN					
Other routes of exposure (infection through skin)								
Asc s 13	*Ascaris suum*	Glutathione S-transferase	4Q5F					
Mala s 1	*Malassezia sympodialis*	β-Propeller fold	2P9W					
Mala s 6	*Malassezia sympodialis*	Cyclophilin	2CFE					
Mala s 13	*Malassezia sympodialis*	Thioredoxin	2J23					

The codes refer to the structure of molecules that are not necessarily the same polymorphism reported in the WHO/IUIS Allergen Nomenclature database.
3D, Three-dimensional; *(m)*, modified or mutated molecule; *NMR*, nuclear magnetic resonance; *PDB*, Protein Data Bank.
*Allergen in complex with an antibody fragment.
†Structures selected from multiple PDB entries (>540 for hen egg lysozyme).