



Common structures of allergens

Les structures communes des allergènes

A. Pomés

INDOOR Biotechnologies, Inc., 1216 Harris Street, Charlottesville, VA 22903, USA

Abstract

Allergens can be classified into different families of proteins with common structural features, but there are not structures and functions in common to all allergens. High degree of amino acid sequence identity between homologous proteins reflects on common molecular surface patches that are the basis of allergenic cross-reactivity. Allergens need to be available for exposure (dose) to individuals susceptible (genetic predisposition) to develop allergies.

© 2008 Elsevier Masson SAS. All rights reserved.

Résumé

Les allergènes peuvent être classés en différentes familles de protéines partageant des similarités structurales. Par ailleurs, il n'y a pas de fonctions et de structures communes à tous les allergènes. La réactivité croisée allergénique des protéines homologues s'explique par le fort taux de similarité dans leurs séquences en acides aminés et des régions protéiques exposées. Leur présence dans l'environnement (dose) est un prérequis pour le développement d'allergies chez les individus génétiquement prédisposés.

© 2008 Elsevier Masson SAS. All rights reserved.

Keywords: Allergens; Molecular structure; Cross-reactivity; Biochemical function

Mots clés : Allergènes ; structures moléculaires ; Réactions croisées ; Fonction biochimique

1. Introduction

Allergens are proteins or glycoproteins lacking a common molecular structure or biochemical function that identifies them as allergens. Like nonallergenic proteins, allergens are structurally heterogeneous: some have only α -helices (Fel d 1, tropomyosins), others only β -sheets (Group 2 dust mite allergens), and most are constituted of both secondary structural elements. Structural variety reflects in variety of biochemical functions, and little is known about the relevance of function on allergenicity. Since the identification of several proteases among mite allergens (the cysteine protease Der p 1 and the serine proteases Der p 3, Der p 6 and Der p 9), proteolytic activity is the main biochemical function that has been associated to allergenicity. The increase of bronchial permeability by disruption of epithelial tight junctions, the promotion of IgE synthesis by cleavage of CD23 in activated B

cells and CD25 in T cells, and the induction of release of proinflammatory cytokines from bronchial epithelial cells, mast cells and basophils are among the mechanisms thought to contribute to the allergenicity of these proteolytic enzymes [1]. Interestingly, most allergens are not proteolytic enzymes. So, how can we distinguish allergens from nonallergenic proteins?

With advances in molecular biology, a better knowledge of the structure and biochemical or biological function of allergens has been acquired in the past 15 years. Since the first experimental determination of the tertiary structure of a clinically important inhalant allergen, Bet v 1, in 1996, ~50 allergen structures have been reported in the Protein Data Bank (<http://www.rcsb.org/pdb/home/home.do>), including few allergen–antibody complexes that allowed the precise mapping of epitopes on the allergen surface [2]. Why is interesting to know the tertiary structure of allergens? It allows the mapping of B and T cell epitopes, it provides information about possible functions of allergens, provides a better understanding of cross-reactivity, contributes to study mechanisms of allergenicity and allows the design of hypoallergenic variants for immunotherapy.

E-mail address: apomes@inbio.com.

In this manuscript, example of how function is determined by molecular structure and a review of four structure–function relationships among allergens are presented, followed by observations related to allergenic cross-reactivity that derive from the analysis of the structure of allergens.

2. How does molecular structure define function?

Molecular shape, as a result of the three dimensional structure of the protein, defines in part the biochemical function. For example, the linear shape of tropomyosin contributes to the structural function of this protein during muscle contraction, whereas, globular shapes with a pocket provide the structural environment for a catalytic site to proteolytic enzymes such as Der p 1 [3–5]. Some allergens have an internal cavity that is associated to the transport of a ligand. This is the case for rodent lipocalins that bind pheromones, or cat Fel d 1, for which the specific function of the ligand is unknown [6,7].

Flexibility is another determinant of function, and calcium binding proteins such as the cockroach allergen Bla g 6, which is a troponin C, have a regulatory function due to their capacity to modify the structure upon calcium binding [8].

Finally, specific amino acids determine allergen function. The cysteine protease Der p 1, as well as serine and aspartic proteases, has few amino acids involved in the catalytic mechanism of the enzyme. Specific motifs are also involved in enzymatic inhibition (such as the cysteine protease inhibitor motifs in Fel d 3 and Can f 1) [9,10]. Ion binding, as described for zinc bound to Bla g 2, may confer stability to the allergen, and contribute to its persistence in the environment and therefore to the development of asthma [11].

3. Structure-function relationships among allergens

3.1. Same structure-same function

Most mammalian allergens are lipocalins: mouse Mus m 1, rat Rat n 1, dog Can f 1 and Can f 2, cow Bos d 2 and Bos d 5, cat Fel d 4 and horse Equ c 1 and Equ c 2. The cockroach allergen Bla g 4 is also a lipocalin [12]. Despite sharing low sequence homology (20–30%), allergen lipocalins have the same overall fold: an anti-parallel β -barrel with a C-terminal α -helix. The β -barrel encloses an internal cavity and exposes a ligand binding site in a hydrophobic pocket. Lipocalins are secreted or excreted into the skin or urine and act as ligand binding proteins, such as the pheromone transporters Mus m 1 and Rat n 1 [6]. The Group 2 mite allergens is another example of allergens with the same structure, a immunoglobulin-like fold similar to the LPS binding protein MD-2, and a function related to lipid transport [13,14]. Similarly, tropomyosins from invertebrates (cockroach, mite, snails and seafood) share ~80% amino acid identity, the same function and a common fold: two parallel α -helices wound around each other [3].

3.2. Same structure-different function

Some allergens, despite sharing the same structure, have different functions due to differences in specific amino acids involved in the function. Bla g 2, for example, is an aspartic protease homolog from the German cockroach *Blattella germanica*, which has the same tertiary structure as active aspartic proteases (pepsin, chymosin, renin...), but is inactive due to important amino acid substitutions in the catalytic site [11,15]. On the other hand, some active aspartic proteases are also allergens. For example, 28% of patients with allergic bronchopulmonary aspergillosis showed IgE binding to the aspartic protease rAsp f 10 (WHO/IUIS Allergen nomenclature website: www.allergen.org). Airway symptoms reported in rennet producing plants have also been associated with occupational exposure to aspartic proteases from fungi (*Rhizomucor miehei*, *Cryphonectria parasitica*) [16]. These enzymes, capable of clotting milk, are frequently used for cheese production.

3.3. Different structure-same function

This is the most uncommon case of structure–function relationship, probably resulting from a convergent evolution of two structurally different proteins into performing the same function. As mentioned above, the cat cystatin Fel d 3 and the dog lipocalin Can f 1 have different tertiary structures, but share a cysteine protease motif that could confer them the same function. Similarly, the salivary lipocalin Von Ebner's gland protein has a cysteine protease motif that has been proven to act as the natural inhibitors cystatins [1,9,10,17].

3.4. Different structure-different function

Most proteins with different structures have different functions, and the same applies to allergens. There is a wide range of allergen functions including proteolytic and non-proteolytic enzymes, enzyme inhibitors, binding proteins, structural proteins, storage proteins and regulatory proteins. It has recently been reported that most plant and animal food and pollen allergens belong to a restricted number of protein superfamilies.

Animal food allergens were classified into three main families (tropomyosins, EF-hand proteins and caseins), most plant food allergens were grouped into just four (prolamin, cupin, profilin, Bet v 1 superfamily) and the major pollen allergen families are constituted by expansins, profilins and calcium-binding proteins [18,19,20]. Analysis of the evolutionary relationships of allergenic and nonallergenic members of these families, combined with information about the quantitative availability of the proteins for human exposure, will provide a wider perspective on determinants of allergenicity.

4. Molecular structure and allergenic cross-reactivity

The definition of cross-reactivity is based on immunologic recognition by antibodies [21], therefore proteins sharing the

same fold but not a common surface will not show allergenic cross-reactivity. Proteins sharing the same overall fold or tertiary structure can have an amino acid sequence homology as low as ~20–30%. In this case, the protein 'framework' or group of amino acids that contribute to the overall fold is conserved, whereas the molecular surface will most likely lack many common patches of amino acids, and allergenic cross-reactivity will be absent. Lipocalin allergens are a good example of such a low homology with conserved tertiary structure.

Contrarily, allergen cross-reactivity occurs when two proteins share high degree of amino acid sequence identity (~80%), and therefore have common amino acid surface patches. Group 2 dust mite allergens provide a good example of how the degree of identity reflects on allergenic cross-reactivity. Der p 2, Der f 2 and Eur m 2 share ~83% amino acid sequence identity and are cross-reactive. However, the storage mite allergens Lep d 2, Tyr p 2 and Gly d 2, which show extensive cross-reactivity among them, share only ~40% identity with Der p 2 and do not cross-react with this allergen [14,22]. Similarly, the high homology (~80% amino acid identity) among tropomyosins from invertebrates (cockroach, mite, snails and seafood) is the basis of antigenic cross-reactivity. However, the lack of cross-reactivity between vertebrate tropomyosins and the homologous invertebrate allergens presumably is explained by the low homology (~45%) that exists between proteins from both groups.

The affinity of IgE antibodies towards cross-reacting allergens may be different, and cross-inhibition experiments allow distinction between primary (sensitizing) versus incomplete or nonsensitizing allergen. Bet v 1 and Mal d 1 illustrate both kinds of allergens, respectively [23]. Cross-reactivity between proteins without structural similarity is unlikely and, in the field of allergy, this situation has not been found to result in relevant biological effects [21].

A factor to be taken into account when analyzing cross-reactivity is the homology of the exogenous allergen with endogenous proteins. With rare exceptions such as beta-casein, IgE antibodies are raised against epitopes that are non-common to homologous endogenous proteins [24]. In other words, amino acids in common between two allergens will not necessarily be involved in cross-reactivity if they are also present in the endogenous human homolog. For animal food proteins, for example, the evolutionary relationships of each of the above mentioned allergen superfamilies showed that in general, proteins with a sequence identity to a human homolog above approximately 62% were rarely allergenic [18].

5. Final remarks

Important information has derived from structural studies of allergens which helps to assess allergenic cross-reactivity and to design therapeutic and diagnostic tools. In the next years, extensive analysis of structural features involved in antibody binding epitopes by cocrystallization of allergen–antibody complexes will contribute to understand IgE and IgG antibody responses. It has recently been suggested that allergens are proteins which are exposed to the immune system and lack

bacterial homologues [25]. A lot of work has been done, but there are still plenty of open avenues to explore in order to fully understand what makes a protein an allergen.

References

- [1] Pomés A, Smith AM, Grégoire C, Vailes LD, Arruda LK, Chapman MD. Functional properties of cloned allergens from dust mite, cockroach and cat - are they relevant to allergenicity? *Allergy Clin Immunol Int* 2001;13:162–9.
- [2] 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.
- [3] Reese G, Ayuso R, Lehrer SB. Tropomyosin: an invertebrate pan-allergen. *Int Arch Allergy Immunol* 1999;119:247–58.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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 secretoglobulin family. *J Biol Chem* 2003;278:37730–5.
- [8] 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.
- [9] Ichikawa K, Vailes LD, Pomés A, Chapman MD. Molecular cloning, expression and modelling of cat allergen, cystatin (Fel d 3), a cysteine protease inhibitor. *Clin Exp Allergy* 2001;31:1279–86.
- [10] Konieczny A, Morgenstern JP, Bizinkauskas CB, Lilley CH, Brauer AW, Bond JF, et al. The major dog allergens, Can f 1 and Can f 2, are salivary lipocalin proteins: cloning and immunological characterization of the recombinant forms. *Immunology* 1997;92:577–86.
- [11] 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.
- [12] Flower DR. The lipocalin protein family: structure and function. *Biochem J* 1996;318(Pt 1):1–14.
- [13] Keber MM, Gradisar H, Jerala R. MD-2 and Der p 2 - a tale of two cousins or distant relatives? *J Endotoxin Res* 2005;11:186–92.
- [14] Smith AM, Benjamin DC, Hozic N, Derewenda U, Smith WA, Thomas WR, et al. The molecular basis of antigenic cross-reactivity between the group 2 mite allergens. *J Allergy Clin Immunol* 2001;107:977–84.
- [15] Wünschmann S, Gustchina A, Chapman MD, Pomés A. Cockroach allergen Bla g 2: an unusual aspartic proteinase. *J Allergy Clin Immunol* 2005;116:140–5.
- [16] Jensen A, Dahl S, Sherson D, Sommer B. Respiratory complaints and high sensitization rate at a rennet-producing plant. *Am J Ind Med* 2006;49:858–61.
- [17] van't Hof W, Blankenvoorde MF, Veerman EC, Amerongen AV. The salivary lipocalin von Ebner's gland protein is a cysteine proteinase inhibitor. *J Biol Chem* 1997;272:1837–41.
- [18] Jenkins JA, Breiteneder H, Mills EN. Evolutionary distance from human homologs reflects allergenicity of animal food proteins. *J Allergy Clin Immunol* 2007;120:1399–405.
- [19] Radauer C, Breiteneder H. Evolutionary biology of plant food allergens. *J Allergy Clin Immunol* 2007;120:518–25.
- [20] Radauer C, Breiteneder H. Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. *J Allergy Clin Immunol* 2006;117:141–7.

- [21] Aalberse RC, Akkerdaas J, van RR. Cross-reactivity of IgE antibodies to allergens. *Allergy* 2001;56:478–90.
- [22] Gafvelin G, Johansson E, Lundin A, Smith AM, Chapman MD, Benjamin DC, et al. Cross-reactivity studies of a new group 2 allergen from the dust mite *Glycyphagus domesticus*, Gly d 2, and group 2 allergens from *Dermatophagoides pteronyssinus*, *Lepidoglyphus destructor*, and *Tyrophagus putrescentiae* with recombinant allergens. *J Allergy Clin Immunol* 2001;107:511–8.
- [23] Aalberse RC. Assessment of allergen cross-reactivity. *Clin Mol Allergy* 2007;5:2–7.
- [24] Bernard H, Negroni L, Chatel JM, Clement G, del-Patient K, Peltre G, et al. Molecular basis of IgE cross-reactivity between human beta-casein and bovine beta-casein, a major allergen of milk. *Mol Immunol* 2000;37:161–7.
- [25] Emanuelsson C, Spangfort MD. Allergens as eukaryotic proteins lacking bacterial homologues. *Mol Immunol* 2007;44:3256–60.