

What Makes a Food Protein an Allergen?

Gary A. Bannon, PhD

Address

Monsanto, 800 N. Lindbergh Boulevard, St. Louis, MO 63167, USA.
E-mail: gary.a.bannon@monsanto.com

Current Allergy and Asthma Reports 2004, 4:43–46

Current Science Inc. ISSN 1529-7322

Copyright © 2004 by Current Science Inc.

Food allergens are almost always proteins, but not all food proteins are allergens. This one statement sums up the purpose of this article, defining the difference between an innocuous food protein and a food allergen. The simplest answer is that a food allergen has the ability to first elicit an IgE response, and then, on subsequent exposures, to elicit a clinical response to the same or similar protein. However, this simplistic answer avoids the more complex issues of defining the biochemical characteristics that allow a food protein to survive the extremes of food processing, escape the digestive enzymes of the human gastrointestinal tract, and interact with the immune system. More than 700 allergen sequences have been identified from food and nonfood sources. However, despite increasing knowledge of the structure and amino acid sequences of the identified allergens, only a few biochemical characteristics can be associated with food allergens. Food allergen characteristics, including abundance of the protein in the food; multiple, linear IgE binding epitopes; resistance of the protein to digestion and processing; and allergen structure are discussed, and the possible reasons they predispose some food proteins to become allergens are suggested.

Introduction

People experience adverse immunologically mediated reactions to allergens from a wide variety of sources, including animal dander, chemical additives, drugs, foods, fungi, insect stings, and pollens. The severity of the reaction can range from mildly irritating to life-threatening. No matter what the source of the allergen is, or the severity of the reaction, IgE mediates all type I immediate hypersensitivity reactions. This immunoglobulin is typically produced in response to a pathogen, but in sensitized individuals, it is produced in response to an otherwise innocuous foreign substance. For a person to experience an allergic reaction, three criteria must be met. The first criterion involves the individual's immune system. It must be capable of mounting an inappropriate IgE response to an otherwise innocuous foreign protein (*ie*, genetic contribution). Second, the individual must encounter the allergen under the appropriate circumstances (*ie*, environ-

mental contribution). Finally, the individual must encounter the allergen several times for a clinical reaction to occur. Most allergic reactions are not life-threatening. However, certain allergies, such as to food, insect stings, and drugs, are more likely than others to result in anaphylaxis [1,2•]. This review is limited to discussion of the unique characteristics exhibited by food allergens.

Food-allergy reactions affect 6% to 8% of children and 1% to 2% of the adult population [3,4]; they are elicited by a small subset of the food we consume. The most common food allergies known to affect children are IgE-mediated reactions to cow's milk, eggs, peanuts, soybeans, wheat, fish, and tree nuts [5]. Approximately 80% of all reported food allergies in children are due to peanuts, milk, or eggs. Although most childhood food allergies are outgrown, allergies to peanuts, tree nuts, and fish are rarely resolved in adulthood. In adults, the most common food allergies are to peanuts, tree nuts, fish, and shellfish.

Although the most commonly allergenic foods contain thousands of proteins, allergic reactions are typically elicited by a defined subset of proteins that are found in the food. Aalberse [6] classified food allergens into two groups, broadly based on their ability to only elicit clinical symptoms in previously sensitized individuals (incomplete food allergens) or their ability to sensitize and elicit reactions in susceptible individuals (complete food allergens). Incomplete food allergens comprise those proteins that can only elicit clinical symptoms due to their homology with another allergen. Complete food allergens comprise those proteins that can both sensitize and elicit clinical symptoms.

Identification and purification of allergens has been essential for the structural and immunologic studies necessary to understand how these proteins stimulate IgE antibody formation [7]. In the past several years, a number of allergens have been identified that stimulate IgE production and cause IgE-mediated disease in humans. However, despite increasing knowledge of the structure and amino acid sequences of the identified allergens, specific features associated with IgE antibody formation have not been fully determined [8]. To date, only a few distinct biochemical characteristics can be associated with proteins known to be allergenic. These characteristics include abundance of the protein in the food; multiple, linear IgE-binding epitopes; resistance of the protein to digestion and processing; and allergen structure. In the following paragraphs, each of these characteristics is examined individually for its contribution in enabling food proteins to become allergens.

Abundance

One explanation for why abundant food proteins become allergens can be explained by the idea that the immune system is more likely to encounter these proteins rather than those representing only a small percentage of the total protein ingested. A review of the major allergens in the commonly allergenic foods seems to confirm the importance of this characteristic in making a food protein an allergen. For example, the major allergen of shrimp, Pen a 1, comprises 20% of the total shrimp-tail protein [9]. Also, the major peanut allergens, Ara h 1 and Ara h 2, comprise 2% and 6% of the total protein in peanut [10]. In fact, the major allergens of most of the commonly allergenic foods, including milk, eggs, peanuts, soy, tree nuts, and wheat, represent major protein constituents (>1%) of their respective foods [10]. However, there are exceptions to this rule. An example in which a major allergen does not represent a significant portion of the food is Gal d 1, the major allergen of codfish. Gal d 1 has been identified as parvalbumin and is not a dominant protein in cod muscle [11]. Although abundance might not be a universal characteristic of all food allergens, it seems to be a predisposing factor that, when coupled with other biochemical characteristics, could produce a food allergen.

Multiple, Linear IgE-binding Epitopes

The requirement that IgE be cross-linked on effector cells to release the mediators of allergic symptoms dictates that there be at least two high-affinity, IgE-binding epitopes on a single allergen. Two categories of IgE-binding epitopes, linear and conformational, are generally considered to occur in food allergens. Conformational epitopes occur when either the secondary or the tertiary structure of the allergen is required before IgE will bind. In contrast, linear epitopes only require the primary amino acid sequence of the allergen for IgE to bind. Whereas conformational IgE-binding epitopes are prevalent and important to the etiology of aeroallergen-mediated allergic reactions, linear epitopes are important to food allergens mainly because the immune system encounters them only after they have been partially denatured and digested by the human gastrointestinal (GI) tract. Therefore, the linear IgE-binding epitopes of food allergens have garnered more attention than the less prevalent conformational epitopes. However, Vila *et al.* [12] have shown that, in comparison with tolerant patients, milk-allergic children with persistent symptoms have a significantly higher ratio of specific IgE antibodies to linear epitopes than to native α - and β -casein proteins. These results illustrate the importance of linear IgE-binding epitopes in persistent food allergy and the potential predictive value of conformational epitopes.

Most linear IgE-binding epitopes of food allergens have been mapped using overlapping peptides and serum IgE from patients with documented hypersensitivity to the food in question. For example, the linear IgE-binding epitopes for the major peanut allergens Ara h 1, Ara h 2, and Ara h 3 were mapped using this technology. Twenty-one different linear

IgE-binding epitopes were identified throughout the length of the Ara h 1 molecule [13]. Ten IgE-binding epitopes were identified in Ara h 2 and four in Ara h 3 using the same methods [14,15]. Linear, IgE-binding epitopes have been mapped for other major food allergens, including those in milk [16,17], soy [18], shrimp [19], and cashew [20]. In general, the epitopes that were determined in this manner ranged in length from 6 to 15 amino acids, but there was no obvious sequence motif shared by all peptides. This characteristic makes it difficult to predict new IgE-binding epitopes, and only those that share sequence identity with already determined IgE epitopes can be identified.

Resistance to Denaturation and Digestion

Resistance to denaturation and digestion is thought to be an important characteristic of food allergens because the longer a significant portion of the protein remains intact, the more likely it is to encounter cells of the immune system. A common method for assessing protein stability was to use in vitro models of digestion whose main purpose has been to assess the availability of amino acids for nutritional purposes [21–24]. One of these digestion models was adapted to evaluate the allergenic potential of dietary proteins [25]. In this model, stability to digestion by pepsin, the primary digestive enzyme of the stomach, was used as a criterion for distinguishing food allergens from safe, nonallergenic dietary proteins. It should be stressed that this in vitro pepsin digestion model was not designed to mimic human digestion and, therefore, cannot predict the half-life of a protein in vivo. Rather, this assay was conceived as a means to determine the relative stability of a protein to the extremes of pH and pepsin protease encountered in the mammalian stomach. Using this assay, many of the major food allergens were tested and found to be stable, whereas nonfood allergens were unstable [25,26]. Since these initial reports, there have been numerous studies repeating the pepsin digestion assay on these major food allergens [27]. In general, the original findings that these allergens were stable to pepsin digestion relative to nonallergen proteins were confirmed, but the length of time that either the whole protein or fragments of the allergen were stable did not always agree [28]. The most likely explanation for this quantitative difference is due to subtle changes in the pepsin digestibility assay or in the method by which the proteins of interest were detected. For example, changes to enzyme concentration, pH, protein purity, and method of detection could have large effects on the interpretation of any in vitro assay [29].

Allergens from less commonly allergenic foods have also been studied using the pepsin digestion assay. For example, allergens from fruit and vegetable foods have been tested for their ability to survive pepsin digestion. Allergens from these sources are typically classified as incomplete (cross-reactive) allergens because they share significant structural homology with another allergen. Good examples of this are the allergens of fruit that share significant sequence homology with aeroallergens. The Bet v 1 homologous allergens of apple

(Mal d 1), pear (Pyr c 1), apricot (Pru ar 1), and cherry (Pru av 1) are, in general, labile to enzymatic digestion [27]. The pepsin-sensitive, cross-reactive proteins typically cause localized symptoms of the oropharynx, including tingling or swelling of the lips, tongue, or glottis, and are less likely to cause systemic or GI symptoms. The lack of stability of this class of food allergens might reflect the difference in sensitization routes of food allergens (GI tract) versus aero-allergens (respiratory system). This logic appears to have been confirmed, at least for milk and wheat allergy. Buchanan *et al.* [30,31] have shown that when stability of the major allergens from these foods is disrupted by reduction of disulfide bonds, the allergens were strikingly sensitive to pepsin digestion and lost their allergenicity, as determined by their level of ability to provoke skin test and gastrointestinal symptoms in previously sensitized dogs.

Allergen Structure

Although allergen stability has been demonstrated for a variety of food allergies, there is little known about why these proteins have the ability to resist degradation. The observation that many of the food allergens are proteins containing intramolecular disulfide bonds that might be important to their allergenicity [32] has led to the assumption that protein structure might be an important factor in an allergen's ability to resist denaturation. This assumption was tested using an allergen in milk, β -lactoglobulin (BLG). BLG contains two disulfides [33] that, when disrupted by site-directed mutagenesis, changed its structure [34] and the accessibility of IgG- and IgE-binding epitopes [35,36]. The disruption of disulfide bonds in BLG also had an impact on the sensitivity of this protein to digestion with pepsin and its overall allergenicity with respect to skin test responses and gastrointestinal symptoms in a sensitized dog model of food allergy [31].

Sen *et al.* [37] have shown that the overall structure of another food allergen, Ara h 2, also is dramatically changed when disulfide bonds are reduced. There are eight cysteine residues that could form up to four disulfide bonds in Ara h 2 [38]. Interestingly, the Ara h 2 structures are not completely randomized when the disulfide bonds are reduced but instead are predominated by a β -pleated sheet and β -turn configuration. These results indicate that the Ara h 2 molecule is a very ordered protein, even if its disulfide bonds are not intact. In addition, only the reduced form of Ara h 2 becomes susceptible to rapid digestion with pepsin, chymotrypsin, or trypsin, indicating a reduction in its overall allergenicity.

The structure of peanut allergen, Ara h 1, might also contribute to its stability. Ara h 1 tertiary structure consists of two sets of opposing anti-parallel β -sheets in Swiss roll topology joined by an interdomain linker. The terminal regions of the molecule, which contain many of the IgE-binding epitopes, consist of α -helical bundles containing three helices each [39•]. When Ara h 1 was incubated with digestive enzymes, numerous protease-resistant fragments containing IgE-bind-

ing epitopes were observed. Although there are numerous protease digestion sites throughout the length of the protein, the structure might be so compact that potential cleavage sites are inaccessible until the protein is denatured. Furthermore, the formation of a homotrimer might also afford the molecule protection from protease digestion and denaturation. These results indicated that the quaternary structure of the Ara h 1 protein plays an important role in protecting these epitopes from digestive enzymes [40].

Other investigators have attempted to identify protein structure motifs that are common to known allergens. Ivancic *et al.* [41] have proposed the use of the three-dimensional (3-D) structure of allergens and bioinformatic tools to help in identifying similar structures in other proteins to aid in the identification of potential allergens. A major difficulty in this approach is that only a few food allergens have been studied in enough detail that their 3-D structure has been determined. This limits the usefulness of such an approach. It remains to be determined if certain protein structures or motifs can be associated with a protein's allergenic potential.

Conclusions

So, what is the answer to the question posed in the title of this manuscript? Obviously, a food protein that is abundant and resistant to digestion and processing has the potential to become a food allergen. But will all food proteins with these characteristics become allergens? The answer to this question is an unqualified no. The reason that all food proteins with these characteristics are not allergens is that the development of an allergic reaction is a complex process involving a receptive immune system (*ie*, a genetic predisposition to propagation of an IgE response), the protein being presented with the correct inducers (*ie*, signals that elicit a Th2 response), and a protein with the appropriate biochemical characteristics to survive the GI tract. Although there are a few studies that have explored the genetics of allergic individuals [42•] and the signals that elicit an IgE response [43], the science in these areas is still maturing. We are left with what we know best—the biochemistry of the offending protein. This might not be the most satisfying answer, but it is an excellent start.

References and Recommended Reading

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. The Allergy Report: **Anaphylactic and Anaphylactoid**, vol 3. Milwaukee, WI: The American Academy of Allergy, Asthma, & Immunology, Inc., 2000:115–130.
2. Metcalfe DD, Sampson HA, Simon RA: *Food Allergy: Adverse Reactions to Foods and Food Additives*, edn 3. Malden, MA: Blackwell Publishing; 2003.
3. Sampson HA: **Food allergy. Part 1: immunopathogenesis and clinical disorders.** *J Allergy Clin Immunol* 1999, **103**:717–728.

Excellent review of clinical and basic science of food allergy.

4. Burks AW, Sampson HA: **Food allergies in children.** *Curr Probl Pediatr* 1993, 23:230–252.
 5. Hefle SL, Nordlee JA, Taylor SL: **Allergenic foods.** *Crit Rev Food Sci Nutr* 1996, 36(Suppl):S69–S89.
 6. Aalberse RC: **Food allergens.** *Environ Tox Pharm* 1997, 4:55–60.
 7. Anderson JA, Sogn DD: **Adverse Reactions to Foods.** Publication No. 84-2442. Bethesda, MD: National Institutes of Health; 1984:1–6.
 8. Stanley JS, Bannon GA: **Biochemistry of food allergens.** *Clin Rev Allergy Immunol* 1999, 17:279–291.
 9. Daul CB, Slattery M, Reese G, Lehrer SB: **Identification of the major brown shrimp (*Penaeus aztecus*) as the muscle protein tropomyosin.** *Int Arch Allergy Immunol* 1994, 105:49–55.
 10. Metcalfe DD, Astwood JD, Townsend R, *et al.*: **Assessment of the allergenic potential of foods derived from genetically engineered crop plants.** *Crit Rev Food Sci Nutr* 1996, 36:S165–S186.
 11. Elsayed S, Bennich H: **The primary structure of allergen M from cod.** *Scand J Immunol* 1975, 4:203–208.
 12. Vila L, Beyer K, Jarvinen KM, *et al.*: **Role of conformational and linear epitopes in the achievement of tolerance in cow's milk allergy.** *Clin Exp Allergy* 2001, 31:1599–1606.
 13. Burks AW, Shin D, Cockrell G, *et al.*: **Mapping and mutational analysis of the IgE-binding epitopes on Ara h 1, a legume vicilin protein and a major allergen in peanut hypersensitivity.** *Eur J Biochem* 1997, 245:334–339.
 14. Stanley JS, King N, Burks AW, *et al.*: **Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2.** *Arch Biochem Biophys* 1997, 342:244–253.
 15. Rabjohn P, Helm EM, Stanley JS, *et al.*: **Molecular cloning and epitope analysis of the peanut allergen Ara h 3.** *J Clin Invest* 1999, 103:535–542.
 16. Busse PJ, Jarvinen KM, Vila L, *et al.*: **Identification of sequential IgE-binding epitopes on bovine alpha(s2)-casein in cow's milk allergic patients.** *Int Arch Allergy Immunol* 2002, 129:93–96.
 17. Jarvinen KM, Chatchatee P, Bardina L, *et al.*: **IgE and IgG binding epitopes on alpha-lactalbumin and beta-lactoglobulin in cow's milk allergy.** *Int Arch Allergy Immunol* 2001, 126:111–118.
 18. Helm RM, Cockrell G, Connaughton C, *et al.*: **A soybean G2 glycinin allergen. 2. Epitope mapping and three-dimensional modeling.** *Int Arch Allergy Immunol* 2000, 123:213–219.
 19. Ayuso R, Lehrer SB, Reese G: **Identification of continuous, allergenic regions of the major shrimp allergen Pen a 1 (tropomyosin).** *Int Arch Allergy Immunol* 2002, 127:27–37.
 20. Wang F, Robotham JM, Teuber SS, *et al.*: **Ana o 1, a cashew (*Anacardium occidentale*) allergen of the vicilin seed storage protein family.** *J Allergy Clin Immunol* 2002, 110:160–166.
 21. Marquez UM, Lajolo FM: **Composition and digestibility of albumin, globulins, and glutelins from *Phaseolus vulgaris*.** *J Agric Food Chem* 1981, 29:1068–1074.
 22. Nielson SS: **Degradation of bean proteins by endogenous and exogenous proteases: a review.** *Cereal Chem* 1988, 65:433–442.
 23. Zikakis JP, Rzuclido SJ, Biasotto NO: **Persistence of bovine milk xanthine oxidase activity after gastric digestion in vivo and in vitro.** *J Dairy Science* 1977, 60:533–541.
 24. Petschow BW, Talbott RD: **Reduction in virus-neutralizing activity of a bovine colostrum immunoglobulin concentrate by gastric acid and digestive enzymes.** *J Pediatr Gastroenterol Nutr* 1994, 19:228–235.
 25. Astwood JD, Leach JN, Fuchs RL: **Stability of food allergens to digestion in vitro.** *Nat Biotechnol* 1996, 14:1269–1273.
 26. Astwood JD, Fuchs RL: **Allergenicity of foods derived from transgenic plants.** *Monogr Allergy* 1996, 32:105–120.
 27. Besler M, Steinhart H, Paschke A: **Stability of food allergens and allergenicity of processed foods.** *J Chromatogr B Biomed Sci Appl* 2001, 756:207–228.
 28. Fu TJ, Abbott UR, Hatzos C: **Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid: a comparative study.** *J Agric Food Chem* 2002, 50:7154–7160.
 29. Bannon GA, Goodman RE, Leach JN, *et al.*: **Digestive stability in the context of assessing the potential allergenicity of food proteins.** *Comments on Toxicology*, 2002, 8:271–285.
 30. Buchanan BB, Adamidi C, Lozano RM, *et al.*: **Thioredoxin-linked mitigation of allergic responses to wheat.** *Proc Natl Acad Sci U S A* 1997, 94:5372–5377.
 31. del Val G, Yee BC, Lozano RM, *et al.*: **Thioredoxin treatment increases digestibility and lowers allergenicity of milk.** *J Allergy Clin Immunol* 1999, 103:690–697.
 32. Lehrer SB, Horner WE, Reese G: **Why are some proteins allergenic? Implications for biotechnology.** *Crit Rev Food Sci Nutr* 1996, 36:553–564.
 33. Brownlow S, Morais-Cabral JH, Cooper R, *et al.*: **Bovine beta-lactoglobulin at 1.8 Å resolution: still an enigmatic lipocalin.** *Structure* 1997, 5:481–495.
 34. Peitsch MC: **ProMod and Swiss-Model: Internet-based tools for automated comparative protein modelling.** *Biochem Soc Trans* 1996, 24:274–279.
 35. Ball G, Shelton MJ, Walsh BJ, *et al.*: **A major continuous allergenic epitope of bovine beta-lactoglobulin recognized by human IgE binding.** *Clin Exp Allergy* 1994, 24:758–764.
 36. Kaminogawa S, Shimizu M, Ametani A, *et al.*: **Monoclonal antibodies as probes for monitoring the denaturation process of bovine beta-lactoglobulin.** *Biochim Biophys Acta* 1989, 998:50–56.
 37. Sen MM, Kopper R, Pons L, *et al.*: **Protein structure plays a critical role in peanut allergen stability and may determine immunodominant IgE-binding epitopes.** *J Immunol* 2002, 169:882–887.
 38. Stanley JS, King N, Burks AW, *et al.*: **Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2.** *Arch Biochem Biophys* 1997, 342:244–253.
 39. • Shin DS, Compadre CM, Maleki SM, *et al.*: **Biochemical and structural analysis of the IgE binding sites on Ara h1, an abundant and highly allergenic peanut protein.** *J Biol Chem* 1998, 273:13753–13759.
- A good example of the studies required to fully characterize a food-allergen protein.
40. Maleki SJ, Kopper RA, Shin DS, *et al.*: **Structure of the major peanut allergen Ara h 1 may protect IgE-binding epitopes from degradation.** *J Immunol* 2000, 164:5844–5849.
 41. Ivanciuc O, Schein CH, Braun W: **SDAP: database and computational tools for allergenic proteins.** *Nucleic Acids Res* 2003, 31:359–362.
 42. • Sicherer SH, Furlong TJ, Maes HH, *et al.*: **Genetics of peanut allergy: a twin study.** *J Allergy Clin Immunol* 2000, 106(1 Pt 1):53–56.
- One of only a few examples in which the inheritance of a food allergy has been studied.
43. Herrick CA, Bottomly K: **To respond or not to respond: T cells in allergic asthma.** *Nat Rev Immunol* 2003, 3:405–412.