Food and Agriculture Organization of the United Nations

World Health Organization

Biotech 01/06

Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology

Headquarters of the Food and Agriculture Organization of the United Nations (FAO)
German Room (C-269)
Viale delle Terme di Caracalla, 00100 Rome, Italy
22 – 25 January 2001

Topic 4: Sequence Homology and Allergen Structure

Wolf-Mienhard Becker
Division of Biochemical and Molecular Allergology
Research Center Borstel
Germany

Abstract

In the public discussion about genetically engineered food the strongest argument against it is the reproach that such food would elicit allergies. In 1996 the IFBC/ILSI set up a decision tree which under given conditions either excluded the hazard of eliciting an allergy by genetically engineered food or included it with a labelling requirement. One aspect of the discussion deals with the question of whether the sequence homology of eight contiguous amino acid residues in a suspect protein identical with a known allergen convicts the protein or acquits it of being hazardous if such sequence identity is not found.

Applying this ”eight amino acid match approach” (EAAM-approach) leads to the insight that conformational epitopes and contiguous parts of these epitopes after denaturation, and epitopes made up by sugar residues, are not identifiable by this procedure. Apart from that result identified linear epitopes of peanut or cod fish only consist of 6 or 4 contiguous amino acid residues which are essential for IgE binding. Thus the EAAM-approach would fail. The conclusion from that is that the EAAM-approach even including only six contiguous amino acids can only identify potential allergenic components but not rule them out. Since predicting or excluding allergenicity is a matter of immunology the epitope, the interface between chemical structure and the immune system, should come into focus. From that it becomes clear that the chemical structure is suitable but the most convincing tools are epitope receptors such as patients´IgE or monoclonal antibodies to test the allergenicity of the protein in question in the genetically engineered food. Since the maturation of the immune system cannot be predicted monitoring studies of immune responses in consumers should be undertaken after the genetically engineered food has reached the market.

Introduction

The identification of allergens by their chemical structure is an unrealizable dream. It overlooks the fact that allergens are defined by the misguided immune system. It is true that the specificity of the allergic reaction is strongly correlated with the chemical structure of the allergens but the development of these specificities depends primarily on the individually conditioned immune system of patients. The immune system is not able to recognize the whole structure of a macromolecule such as proteins or glycoproteins but smaller sections of the molecule called determinants or epitopes. However, it cannot be excluded that there exists an interdependence of the three-dimensional presentation of the epitope with the whole structure of the molecule. In other words a given epitope may display a different immunological behaviour depending upon the whole structure of the molecule. In principle the immune system reacts with two different kinds of epitopes, conformational ones, which are made up by the (3-dimensional) surface area of a molecule formed by discontinuous sections of the primary protein structure and continuous or linear epitopes built directly by the primary protein structure. Three compartments of the immune system deal with epitopes: B-cells and their secreted antibodies, T-cells via the T-cell-receptor and cells presenting molecules of the MHC. Whereas B-cells react to conformational and linear epitopes, T-cell epitopes exclusively are of linear nature. The MHC class I molecules react with 9mer peptides via anchor points in position 2 and 9 and MHC class II molecules react with peptides more heterogeneous in length via anchor points in the relative position 1,4, 6 and sometimes 9 (1).

Normally allergens, especially aeroallergens and pollen related food allergens possess conformational epitopes but in classical food allergens linear epitopes are found. For the major allergen of peanut Ara h 1 twenty three linear IgE-reactive epitopes (2), for Ara h 2 ten epitopes (3) and for Ara h 3 four epitopes are described (4). The identification of linear epitopes is

performed by overlapping synthetic peptides delivering the essential minimal peptide configuration in length, which is in some peanut allergen epitopes made up by six amino acid residues. The predominant feature of linear epitopes is their stability to denaturation, which is caused by unfolding of the protein through heating or detergents. In this context the heat stable cross-reacting carbohydrate determinants (CCD) have to be mentioned as a third kind of epitopes built up by carbohydrates or by carbohydrates together with few amino acid residues. Via asparagines a possible N-link to carbohydrate residues may be predictable by the NXS/T motif.

At present there is no algorithm available to identify conformational epitopes, since the number of known conformational epitopes is rare much less IgE-reactive epitopes of this kind. Experimentally it is very difficult to identify conformational epitopes, since up to now X-ray crystallography of the allergen- antibody-complex is needed for the epitope identification. A well-studied example in this regard is the egg lysozyme. These experiments were performed on egg lysozyme antibody (Fab) complexes for the first time. Three epitopes were identified (reviewed in (5)) by performing X-ray crystallography of complexes of egg lysozyme antibody and Fab antibody fragments. These discontinuous epitopes were located on the surface of the molecule and were approximately 750 Å² large. The amino acids that form the epitope recognized by the monoclonal antibody HyHEL-10 are residues 15-16, 20-21, 63, 73, 75, 89, 93, 96-98, and 100-102 (5). It cannot be ruled out that the monoclonal antibody HyHEL-10 reacts with the continuous sequence parts of the conformational epitope e.g. amino acid residues 96-98 and 100-102 if the protein is denatured.

Facts and discussion

In the public discussion about genetically engineered food the strongest argument against it is the reproach that such food would elicit allergies. In 1996 the IFBC/ILSI set up a decision tree which under given conditions either excluded the hazard of eliciting an allergy by genetically engineered food or predicted it with a labelling requirement (6).

This proposal provoked a lot of discussion. One point of the discussion dealt with question whether the sequence homology of eight contiguous amino acid residues in the suspect protein identical with a known allergen confirm the suspicion or acquits the protein of being hazardous in the case of a negative result.

Thus, similarity searches comparing novel food proteins with sequences of known allergens can be indicative of an allergenic potential. However the major criterion for excluding an allergenic material is the assumption in the strategy paper that ”Immunologically significant sequence identity requires a match of at least eight contiguous identical amino acids”. Although this corresponds to the minimal size of T-cell epitopes it is not sufficient to define immunological significance for B-cell epitopes (reviewed by S. Vieths in (7)). The following examples may disprove the above statement of B-cell level.

An lgE-reactive hexadeca-peptide derived from the sequence of the major codfish allergen contained 2 tetrapeptides (DEDK, DELK) separated by a 6-residues spacer which were essential for IgE binding (8). As demonstrated by mutant peptides the spacer was irrelevant for the IgE-binding capacity. Thus, this epitope region contained stretches of only 4 consecutive residues which determined immunological significance. Interestingly, these results were obtained by Elysayed and Apold during the first detailed epitope analysis performed on a food allergen in 1983. Ara h 1, the major peanut allergen, can be taken as another example. By scanning with synthetic peptides, its minimal epitope size was determined as 6 consecutive residues (2). In Ara h 2 (3) identified the 6mer peptide **RDPYSP** as major epitope. Performing a homology search

with the FASTA3 algorithm in the **SWALL** data base, interestingly you did not find Ara h 2 but RNAse T and Laminin with a 100 % match of the peptide with these proteins. Matching the proteins RNAse T and Laminin with Ara h 2 a match of at least eight contiguous identical amino acids is not given. From the theoretical point of view this approach of the strategy paper clearly fails in this example. Testing these proteins on cross-reactivity with IgE of peanut allergic patients would deliver strong arguments to judge the validity of this approach.

The authors of the 'strategy' also state that *'this approach is limited in that it cannot identify discontinuous conformational epitopes'* ' This is true but it represents an important pitfall since protein **structures** and not **sequences** are being conserved due to a common structure-based function. This can be illustrated by looking at the plant pan-allergen profilin which has been identified as the most important minor pollen allergen by R. Valenta and his coworkers (9;10). Since profilins are present in all eucaryotic cells, he also tested IgE cross-reactivities between birch pollen profilin and human profilin. lndeed, IgE from a birch pollen allergic patient cross-reacted with human profilin and human profilin also caused histamine release in birch pollen allergic patients (9).

A sequence alignment of the 2 homologues profilins from birch pollen and man reveals that the longest stretch of matching amino acids contains 4 residues. Despite an overall sequence identity below 30 %, the tertiary structures and functions of both proteins are very similar (11). So in the case of foods which are consumed in the native state, structure based cross-reactivity can be overlooked by this approach. Finally, it has to be noticed that small peptides of 3-4 residues can resemble discontinuous epitopes (12). This was although shown for egg lysozyme (s. above) These so called 'mimotopes' are again not covered by the approach of the 'assessment strategy'.

The following conclusions result from the data referred above: First, when stating that *"failure to find a match of eight contiguous residues anywhere among known allergen amino acid sequences suggests that there is little probability that the introduced protein could possess a shared linear epitope with known allergens' one* has to be aware that at the B-cell level, smaller sequential structures can be essential for reactivity. Thus, for a more reliable safety assessment it is suggested that the level of suspicion be reduced to six identical consecutive residues. In addition, overall similarities should be calculated from end-to-end alignments to identify proteins with homologous structures. Furthermore overall similarity und screening of similar biological function and functional domains should be included in this approach. From Ara h 2 we learn that not all important allergens are available in the protein data bases. Third, albeit only a few numbers of epitopes have been characterised so far, IgE reactive epitopes have to be included in this approach. Homologous sequences of epitopes are of more relevance than a match of at least eight contiguous identical amino acids with immunological irrelevance. Since epitopes are the interface between chemical structure of an allergen and the immune system this approach should be focussed on epitopes. In this context the third kind of epitopes mentioned above, the carbohydrate epitopes should take in consideration. We know from □-amylase that this allergen and protein is glycosylated, when expressed in eucaryotic plants and immunologically active but not in E. coli. This leads to the important question, whether the expression system of the transgenic plant has an influence on the expression and presentation of the protein in question in this organism. However, this question is not answerable by homology screening but immunologically by testing with the IgE of allergic patients. And now it becomes clear that homology screening with at least eight contiguous identical amino acids may help to find potential allergenic compounds in a restricted way but cannot exclude allergenicity of the protein in question if the identification of such homology failed. Immunological problems must be solved by applying immunological methods.

Conclusion

1. The match of at least eight contiguous identical amino acids of the protein in question with a known allergen is only useful for identifying potentially allergenic components
2. This approach is not usable for identifying conformational or parts of conformational epitopes nor epitopes made up of carbohydrates
3. Specific allergen databases should be used for comparison of sequences
4. Both global and local alignment tools as well as screening for homologous structures or domains should be applied.
5. Structure information about IgE reactive epitopes and specific reagents identifying them are the additional tools to identify potentially allergenic components. This will improve the chance of preventing genetically engineered foods from including allergenic components

Suggestions

Research is needed to map all the epitopes of known allergens and to develop monoclonal antibodies against them.

Since the immunological reaction against newly introduced proteins in food is complex and not predictable the elicited immunological reaction should be tested in animals. In the event genetically engineered food reaches the market the elicited immunological reaction in developing antibodies should be observed.

Acknowledgement:

The author is grateful to Dr. S. Vieths (Paul-Ehrlich-Institute, Langen, Germany) for critical comments and fruitful discussions.

References

1. Marsh SGE, Parham P, Barber LD (2000): The HLA Facts Book. Academic Press, San Diego pp 61-72.
2. Burks AW, Shin D, Cockrell G, Stanley JS, Helm RM, Bannon GA (1997). 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 245:334-339.
3. Stanley JS, King N, Burks AW, Huang SK, Sampson H, Cockrell G, Helm RM, West CM, Bannon GA (1997). Identification and mutational analysis of the immunodominant IgE binding epitopes of the major peanut allergen Ara h 2. Arch Biochem Biophys 342: 244-253.
4. Rabjohn P, Helm EM, Stanley JS, West CM, Sampson HA, Burks AW, Bannon GA.(1999). Molecular cloning and epitope analysis of the peanut allergen Ara h 3. Clin Invest 103: 535-542.
5. Padlan EA, in van Regenmortel MHV (Editor) Structure of Antigens,Vol 2, CRC Press, Boca Raton, FL,1992, p. 29.

6. Metcalfe DD, Astwood JD, Townsend R, Sampson H A, Taylor S L, Fuchs R L (1996). Assessment of the allergenic potential of foods derived from genetically engineered crop plants. Crit Rev Food Sci Nutr, Special Supplement, 36: S165-5186.

7. Vieths S (1997). Allergenic potential of genetically modified plant foods- how reliable is the proposed assessment strategy? In Proceedings of the International Symposium on novel foods regulation in the European Union -integrity of the process of safety evaluation BGVV (edit.), Berlin, Germany, pp. 295

8. Elsayed S, Apold J, Holen E, Vik H, Floorvaag E, Dybendal T (1991). The structural requirements of epitopes with IgE binding capacity demonstrated by three major allergens from fish, egg and tree pollen. Scand J Lab Invest 1991; 51 (Suppl 204): 17-31.

9. Valenta, R., Duchene, M., Pettenburger, K., Sillaber, S., Valent, P., Bettelheim, P., Breitenbach, M., Rumpold, H., Kraft D., Scheiner, O. (1991). Identification of profilin as a novel pollen allergen.; IgE auto-reactivity in sensitized individuals. Science 253: 557-560.

10. Valenta R, Duchene M, Ebner C, Valent P, Sillaber C, Deviller P, Ferreira F, Teijkl M, Edelmann H. Kraft D, Scheiner O (1992). Profilins constitute a novel family of functional plant pan-allergens. J Exp Med 175: 377-385.

11. Fedorov AA, Ball T, Mahoney NM, Valenta R, Almo SC (1997). The molecular basis for allergen cross-reactivity: crystal structure and IgE-epitope mapping of birch pollen profilin. Structure 5: 33-45.

12. Geysen HM, Rodda SJ, Mason TJ, Tribbick G, Schoofs PG (1987). Strategies for epitope analysis using peptide synthesis. J Immunol Methods 102:259-274.