

7 Assessment of possible allergenicity (Proteins) in foods derived from recombinant-DNA plants

Food allergies

Food allergies are adverse reactions to an otherwise harmless food or food component and involve an abnormal response of the body's immune system to specific protein(s) in foods known as "allergens". True food allergies may involve several types of immunological response (Sampson and Burks, 1996).

The most common types of food allergies are mediated by allergen-specific immunoglobulin E (IgE) antibodies²¹. IgE-mediated reactions are known as immediate hypersensitivity reactions because symptoms occur within minutes to a few hours after ingestion of the offending food. IgE-mediated reactions can occur to pollens, mould spores, animal danders, insect venoms and other environmental stimuli as well as foods. IgE-mediated reactions affect perhaps 10 to 25 percent of the population in developed countries (Mekori, 1996).

Food allergies represent a small fraction of all allergic diseases, affecting less than 2.5 percent of the population in developed countries (Anderson, 1996). Infants and young children are more commonly affected by IgE-mediated food allergies than adults; the prevalence among infants under the age of 3 years may be as high as 5 to 8 percent (Bock, 1987; Sampson, 1990).

True food allergies also include cell-mediated reactions, which involve sensitized tissue-bound lymphocytes rather than antibodies (Sampson, 1990). In cell-mediated reactions, the onset of symptoms occurs more than 8 hours after ingestion of the offending food. The role of foods in cell-mediated reactions remains uncertain (Burks and Sampson, 1993) but coeliac disease²², also known as gluten-sensitive enteropathy, affects one in every 300 to 3 000 individuals in the population, depending upon the specific geographical region. Both IgE-mediated food allergies and gluten-sensitive enteropathy are treated with specific avoidance diets. Because in both cases the threshold dose is quite low, great care must be taken in the construction of safe and effective avoidance diets.

The Codex Alimentarius Commission has produced a list of the most common allergenic foods associated with IgE-mediated reactions on a worldwide basis, which includes peanuts, soybeans, milk, eggs, fish, crustacea, wheat and tree nuts. These commonly allergenic foods account for over 90 percent of all moderate to severe allergic reactions to foods, although an extensive literature search has revealed more than 160 foods associated with sporadic allergic reactions (Hefle *et al.*, 1996).

Allergic reactions to fresh fruits and vegetables, comprising the so-called oral allergy syndrome, are also rather common (Parker *et al.*, 1990), but these foods are not included on the Codex Alimentarius Commission list because the symptoms are typically mild and confined to the oropharyngeal region, and the allergens are unstable to heating and digestion. The list established by the Codex Alimentarius Commission also includes gluten-containing cereals (wheat, rye, barley, oats and spelt) that are implicated in the aetiology of gluten-sensitive enteropathy. Table 7.1 provides a summary of protein sequences of food allergens from foods of plant origin and their accession numbers for retrieving the sequence data from the relevant databases.

²¹ IgE, or immunoglobulin E, is a protein antibody that recognizes an allergen. IgE circulates in the blood and becomes fixed on the surface of specific cells (basophils and mast cells). When IgE on the cell surface binds to an allergen, this triggers the release of chemical mediators that provoke the symptoms associated with allergic reactions.

²² Gluten-sensitive enteropathy is a malabsorption syndrome characterized by body wasting, anaemia, diarrhoea and bone pain, along with other symptoms.

Almost all food allergens are proteins, although it is possible that other food components may act as haptens²³. Similarly, prolamin proteins from wheat, rye, barley, etc. are involved in the elicitation of gluten-sensitive enteropathy. While the crops from which staple foods are derived contain tens of thousands of different proteins, relatively few are allergenic. The distribution of these proteins varies throughout the plant and can be influenced by environmental factors, such as climate and disease stress. Conventional breeding removes diversity from or introduces protein diversity into the food supply, but has had little, if any, effect on the allergenic potential of our major foods.

Table 7.1. Food allergen protein sequences of plant origin¹

<i>Species</i>	<i>Common name</i>	<i>Allergen</i>	<i>Synonym/function</i>	<i>Accession²</i>
<i>Arachis hypogaea</i>	Peanuts	<i>Ara h 1</i>	Clone P41b	L34402
			Clone 5A1	L33402
			Clone P17	L38853
		Peanut lectin	Agglutinin	S14765
<i>Bertholletia excelsa</i>	Brazil nut	<i>Ber e 1</i>	2S albumin (BE2S1 gene)	X54490
<i>Brassica juncea</i>	Leaf mustard	<i>Bra j 1E-L</i>	2S albumin large chain	S35592
		<i>Bra j 1E-S</i>	2S albumin small chain	S35591
<i>Carica papaya</i>	Papaya	Papain		M15203
<i>Glycine max</i>	Soybean	Glycinin	A1aBx subunit	X02985
			A2B1a subunit	Y00398
			A3B4 subunit	M10962
			G1 subunit	X15121
			G2 subunit	X15122
			G3 subunit	X15123
		beta-Conglycinin	alpha-subunit	X17698
			CG4 subunit	S44893
		Soy lectin	Soy agglutinin	K00821
		Kuntz trypsin inhibitor	KTi-s subtype	X80039
			KTi-a subtype	X64447
			KTi-b subtype	X64448
<i>Hordeum vulgare</i>	Barley	<i>Hor v 1</i>	alpha-amylase/trypsin inhibitor	S26197
		<i>Hor v 1</i>	alpha-amylase/trypsin inhibitor	P32360
<i>Malus domestica</i>	Apple	<i>Mal d 1</i>	Profilin	X83672
<i>Oryza sativa</i>	Rice	RAP	Rice allergenic protein	X66257
		RAG1	Rice allergen 1	D11433
		RAG2	Rice allergen 2	D11434
		RAG5	Rice allergen 3	D11430
		RAG14	Rice allergen 14	D11432
		RAG17	Rice allergen 17	D11431
<i>Phaseolus vulgaris</i>	Kidney bean	PR-1	Pathogenesis related protein 1	S11929
		PR-2	Pathogenesis related protein 2	S11930
<i>Sinapis alba</i>	White mustard	<i>Sin a 1.1</i>	2S albumin/amylase inhibitor	S54101
		<i>Sin a 1.2</i>	2S albumin/amylase inhibitor	PC1247
<i>Triticum aestivum</i>		WGA	Wheat germ agglutinin A	M25536
		WGA	Wheat germ agglutinin D	M25537
<i>Triticum durum</i>	Pasta wheat	WGA	Wheat germ agglutinin	J02961
<i>Triticum turgidum</i>	Poulard wheat	16K allergen	alpha-amylase inhibitor	S19296

¹ Adapted from Metcalfe et al. (1996).

² Public domain databases: GenBank/EM BL/Genpept ver 86.0, SWISSPROT ver 30, PIR ver 41.

²³ Haptens are small molecules that may interact with body proteins or food proteins and cause these proteins to become allergenic.



Allergenicity potential of foods derived from recombinant-DNA plants

Potential allergenicity is a concern with proteins introduced into the human diet through food derived from recombinant-DNA plants, especially when there is no history of their consumption, where the source cannot be readily identified, or when they are recombined versions of proteins from different sources. The current allergenicity assessment approach is presented in the Annex “Assessment of possible allergenicity” of the Codex Guideline (see Appendix 2). As there is no definitive test that can be relied upon to predict allergic responses in humans to a newly expressed protein, the Codex recommends that an integrated, stepwise, case-by-case approach be used in the assessment of possible allergenicity of newly expressed proteins. This approach takes into account the evidence derived from several types of information and data because no single criterion is sufficiently predictive.

In addition to the Annex, the Codex Guideline outlines approaches to allergenicity assessment in paragraphs 41–43.

Allergenicity assessment strategy

The initial steps in assessing the possible allergenicity of any newly expressed protein are the determination of the source of the introduced protein, any significant similarity between the amino acid sequence of the protein and that of known allergens, and its structural properties, including, but not limited to, its susceptibility to enzymatic degradation, heat and/or acid and enzymatic treatment.

As there is no single test that can predict the likely human IgE response to oral exposure, the first step in the characterization of newly expressed proteins should be the comparison of the amino acid sequence and certain physicochemical characteristics of the newly expressed protein with those of established allergens using a weight of evidence approach (see Box 7.1 for an outline of some important parameters used). This will require the isolation of any newly expressed proteins from the recombinant-DNA plant or the synthesis or production of the substance from an alternative source, in which case the material should be shown to be structurally, functionally and biochemically equivalent to that produced in the recombinant-DNA plant. Particular attention should be paid to the choice of the expression host, because the post-translational modifications allowed by different hosts (i.e. eukaryotic vs prokaryotic systems) may have an impact on the allergenic potential of the protein.

It is important to establish whether the source is known to cause allergic reactions. Genes derived from known allergenic sources should be assumed to encode an allergen unless scientific evidence demonstrates otherwise.

CODEX GUIDELINE PARAGRAPH 41. When the protein(s) resulting from the inserted gene is present in the food, it should be assessed for potential allergenicity in all cases. An integrated, stepwise, case-by-case approach used in the assessment of the potential allergenicity of the newly-expressed protein(s) should rely upon various criteria used in combination (since no single criterion is sufficiently predictive on either allergenicity or non-allergenicity). As noted in paragraph 20, the data should be obtained using sound scientific methods. A detailed presentation of issues to be considered can be found in the Annex to this document.

CODEX GUIDELINE PARAGRAPH 42. The newly expressed proteins in foods derived from recombinant-DNA plants should be evaluated for any possible role in the elicitation of gluten-sensitive enteropathy, if the introduced genetic material is obtained from wheat, rye, barley, oats, or related cereal grains.

CODEX GUIDELINE PARAGRAPH 43. The transfer of genes from commonly allergenic foods and from foods known to elicit gluten-sensitive enteropathy in sensitive individuals should be avoided unless it is documented that the transferred gene does not code for an allergen or for a protein involved in gluten-sensitive enteropathy.

Box 7.1. Important parameters used in the assessment of allergenicity

Source of the protein

As part of the database supporting the safety of foods derived from recombinant-DNA plants, any reports of allergenicity associated with the donor organism should be described. Allergenic sources of genes are defined as those organisms for which reasonable evidence of IgE-mediated oral, respiratory or contact allergy is available. Knowledge of the source of the introduced protein allows the identification of tools and relevant data to be considered in the allergenicity assessment. These include the availability of sera for screening purposes, documentation of the type, severity and frequency of allergic reactions, the structural characteristics and amino acid sequence of the protein, and the physiochemical and immunological properties (if available) of known allergenic proteins from that source.

Amino acid sequence homology

The purpose of a sequence homology comparison is to assess the extent to which a newly expressed protein is similar in structure to a known allergen. This information may suggest whether the protein has allergenic potential. Sequence homology searches should be performed to compare the structure of all newly expressed proteins with all known allergens. Searches should be conducted using various algorithms such as FASTA or BLASTP²⁴ to predict overall structural similarities. Strategies such as stepwise contiguous identical amino acid segment searches may also be performed to identify sequences that may represent linear epitopes. The size of the contiguous amino acid search should be based on a scientifically justified rationale in order to minimize the potential for false negative or false positive results²⁵. Validated search and evaluation procedures should be used in order to produce biologically meaningful results.

IgE cross-reactivity between the newly expressed protein and a known allergen should be considered a possibility when there is more than 35 percent identity in a segment of 80 or more amino acids (FAO/WHO, 2001), or when other scientifically justified criteria are met. All the information resulting from the sequence homology comparison between the newly expressed protein and known allergens should be reported to allow a case-by-case scientifically based evaluation.

Sequence homology searches have certain limitations. In particular, comparisons are limited to the sequences of known allergens in publicly available databases and the scientific literature. There are also limitations in the ability of such comparisons to detect non-contiguous epitopes capable of binding specifically with IgE antibodies.

A negative sequence homology result indicates that a newly expressed protein is not a known allergen and is unlikely to be

cross-reactive with known allergens. A result indicating the absence of significant sequence homology should be considered along with the other data outlined under this strategy in assessing the allergenic potential of newly expressed proteins. Further studies should be conducted as appropriate (see also Specific serum screening, below). A finding of positive sequence homology indicates that the newly expressed protein is likely to be allergenic. If the product is to be considered further, it should be assessed using serum from individuals sensitized to the identified allergenic source.

Pepsin resistance

Resistance to pepsin digestion has been observed in several food allergens; thus a correlation exists between resistance to digestion by pepsin and allergenic potential²⁶. Therefore, the resistance of a protein to degradation in the presence of pepsin under appropriate conditions indicates that further analysis should be conducted to determine the likelihood of the newly expressed protein being allergenic. The establishment of a consistent and well validated pepsin degradation protocol may enhance the utility of this method. However, it should be taken into account that a lack of resistance to pepsin does not exclude the possibility that the newly expressed protein could be a relevant allergen. Although the pepsin resistance protocol is recommended, it is recognized that other enzyme susceptibility protocols exist. Alternative protocols may be used where adequate justification is provided²⁷.

Specific serum screening

For those proteins that originate from a source known to be allergenic, or that have sequence homology with a known allergen, testing in immunological assays should be performed where sera are available. Sera from individuals with a clinically validated allergy to the source of the protein can be used to test the specific binding to IgE class antibodies of the protein in *in vitro* assays. A critical issue for testing will be the availability of human sera from sufficient numbers of individuals²⁸. In addition, the quality of the sera and the assay procedure need to be standardized to produce a valid test result. For proteins from sources not known to be allergenic, and which do not exhibit sequence homology to a known allergen, targeted serum screening may be considered where such tests are available, as described in the final paragraph below.

In the case of a newly expressed protein derived from a known allergenic source, a negative result in *in vitro* immunoassays may not be considered sufficient, but should prompt additional testing, such as the possible use of skin test and *ex vivo* protocols²⁹. A positive result in such tests would indicate a potential allergen.

²⁴ FASTA is a computer program, based on the method of W. Pearson and D. Lipman (*Proc. Natl. Acad. Sci. USA*, 85: 2444–2448, 1988), that searches for similarities between one sequence (the query) and any group of sequences (the database) (<http://fasta.bioch.virginia.edu/>). The BLAST (basic local alignment search tool) program uses a strategy based on matching sequence fragments by employing a powerful statistical model, developed by S. Karlin and S. Altschul (*Proc. Natl. Acad. Sci. USA*, 87: 2264–2268, 1990), to find the best local alignments. BLASTP is the NCBI BLAST program for comparing a protein query sequence to a protein database. The original BLAST program was developed at NCBI (<http://www.ncbi.nih.gov/BLAST/>). There is a separate BLAST distribution called WU-BLAST available from Washington University (<http://blast.wustl.edu/>).

²⁵ It is recognized that the 2001 FAO/WHO consultation suggested moving from eight to six identical amino acid segments in searches. The smaller the peptide sequence used in the stepwise comparison, the greater the likelihood of identifying false positives; conversely, the larger the peptide sequence used, the greater the likelihood of false negatives, thereby reducing the utility of the comparison (FAO/WHO, 2001).

²⁶ The method outlined in the United States Pharmacopoeia (1995) was used in the establishment of the correlation (Astwood *et al.* 1996).

²⁷ Report of the Joint FAO/WHO Expert Consultation on allergenicity of foods derived from biotechnology (FAO/WHO, 2001): Section 6.4 Pepsin resistance.

²⁸ According to the Joint Report of the FAO/WHO Expert Consultation on allergenicity of foods derived from biotechnology (FAO/WHO, 2001) a minimum of eight relevant sera is required to achieve 99 percent certainty that the new protein is not an allergen, in the case of a major allergen. Similarly, a minimum of 24 relevant sera is required to achieve the same level of certainty in the case of a minor allergen. It is recognized that these quantities of sera may not be available for testing purposes.

²⁹ An *ex vivo* procedure is described as testing for allergenicity performed using cells or tissue culture from allergic human subjects (FAO/WHO, 2001).

The level of exposure to the newly expressed protein and the effects of relevant food processing will contribute towards an overall conclusion about the potential human health risk. In this regard, the nature of the food product intended for consumption should be taken into consideration in determining the types of processing that would be applied and its effects on the presence of the protein in the final food product.

As scientific knowledge and technology evolves, other methods and tools may be considered in assessing the allergenicity potential of newly expressed proteins as part of the assessment strategy. These methods should be scientifically sound and may include targeted serum screening (i.e. the assessment of protein binding to IgE in sera of individuals with clinically validated allergic responses to broadly related categories of foods), the development of international serum banks, use of animal models, and examination of newly expressed proteins for T-cell epitopes and structural motifs associated with allergens.

References

- Anderson, J.A. 1996. Allergic reactions to foods. *Critical Rev. Food Sci. Nutr.*, 36: S19–S38.
- Astwood, J.D., Leach, J.N. & Fuchs, R.L. 1996. Stability of food allergens to digestion *in vitro*. *Nature Biotech.*, 14: 1269–1273.
- Bock, S.A. 1987. Prospective appraisal of complaints of adverse reactions to foods in children during the first three years of life. *Paediatrics* 79: 683–688.
- Burks, A.W. & Sampson, H. 1993. Food allergies in children. *Curr. Prob. Paediatrics* 23: 230–252.
- FAO/WHO. 2001. *Evaluation of allergenicity of genetically modified foods*. Report of a joint FAO/WHO Expert Consultation on allergenicity of foods derived from biotechnology. 22–25 January 2001. Rome, Italy, Food and Agriculture Organization of the United Nations.
- Hefle, S.L., Nordlee, J.A. & Taylor, S.L. 1996. Allergenic foods. *Critical Rev. Food Sci. Nutr.*, 36: S69–S89.
- Mekori, Y.A. 1996. Introduction to allergic disease. *Critical Rev. Food Sci. Nutr.*, 36: S1–S18.
- Metcalf, D.D., Astwood, J.D., 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. *Critical Rev. Food Sci. Nutr.*, 36: S165–S186.
- Parker, S.L., Leznoff, A., Sussman, G.L., Tarlo, S.M. & Kronl, M. 1990. Characteristics of patients with food-related complaints. *J. Allergy Clin. Immunol.*, 86: 503–511.
- Sampson, H.A. 1990. Immunologic mechanisms in adverse reactions to foods. *Immunol. Allergy Clin. N. Am.* 11: 701–706.
- Sampson, H.A. & Burks, A.W. 1996. Mechanisms of food allergy. *Ann. Rev. Nutr.* 16: 161–177.

Additional resources

- International Food Biotechnology Council and International Life Sciences Institute Allergy and Immunology Institute. 1996. Allergenicity of foods produced by genetic modification. F.M. Clydesdale, ed. *Crit. Rev. Food Sci. Nutr.*, 36.
- OECD. 1997. *Safety assessment of new foods: results of an OECD survey of serum banks for allergenicity testing, and use of databases*. Paris, Organization for Economic Cooperation and Development (OECD).
[http://www.oilis.oecd.org/oilis/1997doc.nsf/LinkTo/NT00000C6A/\\$FILE/JT00121603.PDF](http://www.oilis.oecd.org/oilis/1997doc.nsf/LinkTo/NT00000C6A/$FILE/JT00121603.PDF)
- Taylor, S. 2002. *Topic 13: Allergenicity*. Joint FAO/WHO Expert Consultation on foods derived from biotechnology. Geneva, WHO/FAO ●