



Current codex guidelines for assessment of potential protein allergenicity

G.S. Ladics*

DuPont Experimental Station, Building 353, Wilmington, DE 19880, USA

ARTICLE INFO

Keywords:

Protein allergenicity
Genetically modified crops
CODEX guidelines
Weight-of-evidence

ABSTRACT

A rigorous safety assessment process exists for GM crops. It includes evaluation of the introduced protein as well as the crop containing such protein with the goal of demonstrating the GM crop is "as-safe-as" non-transgenic crops in the food supply. One of the major issues for GM crops is the assessment of the expressed protein for allergenic potential. Currently, no single factor is recognized as an identifier for protein allergenicity. Therefore, a weight-of-evidence approach, which takes into account a variety of factors and approaches for an overall assessment of allergenic potential, is conducted [Codex Alimentarius Commission, 2003. Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy, 30 June–5 July, 2003. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants, and Appendix IV, Annex on the assessment of possible allergenicity, pp. 47–60]. This assessment is based on what is known about allergens, including the history of exposure and safety of the gene(s) source; protein structure (e.g., amino acid sequence identity to human allergens); stability to pepsin digestion *in vitro* [Thomas, K. et al., 2004. A multi-laboratory evaluation of a common *in vitro* pepsin digestion assay protocol used in assessing the safety of novel proteins. Regul. Toxicol. Pharmacol. 39, 87–98]; an estimate of exposure of the novel protein(s) to the gastrointestinal tract where absorption occurs (e.g., protein abundance in the crop, processing effects); and when appropriate, specific IgE binding studies or skin prick testing. Additional approaches may be considered (e.g., animal models; targeted sera screening) as the science evolves; however, such approaches have not been thoroughly evaluated or validated for predicting protein allergenicity.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The prevalence of food allergy is approximately 1–2% in adults and 6–8% in children [Metcalf et al., 1996; Sampson, 1997]. A food allergy is a reaction of the immune system to an otherwise harmless food or food component. Most food allergies are mediated by IgE and are characteristic of type-I reactions. Food allergic reactions occur typically in atopic individuals who are genetically predisposed to allergy and who have been previously sensitized to the allergen [Sicherer, 2000]. Antigen specific IgE is produced and binds to the surface of mast cells and basophils distributed throughout the body. Subsequent exposure of sensitized subjects to the inducing antigen causes an allergic response (i.e., the release of chemical mediators such as histamine from basophils and mast cells with the typical clinical manifestation of itching, hives and/or gastrointestinal symptoms). The potential consequences of food allergy

can be serious with severe reactions (i.e., anaphylaxis) occurring in approximately three individuals per 100,000 per year [Burks and Sampson, 1997]. The only current treatment for food allergies is avoidance of the offending food [Smith and Munoz-Furlong, 1997].

Overall, approximately 90% of all food allergies are associated with a small number of specific proteins represented by eight major allergenic foods: peanuts, tree nuts, cow's milk, hen's eggs, fish, crustacea (e.g., shrimp), wheat, and soybeans. The prevalence of food allergy to specific foods can vary geographically (e.g., buckwheat in Asia, celery in Europe, walnut and pecan in the US) [Metcalf et al., 1996]. Approximately 200 of the hundreds of thousands of proteins in food are identified as food allergens [Day, 1996]. The remaining 10% of food allergies are caused by less commonly allergenic proteins or minor allergenic foods and affect a relatively small number of people [Hefle et al., 1996]. Importantly, food, whether developed by conventional means or through biotechnology, is a potential source of allergens.

Prior to marketing of foods derived from recombinant-DNA plants, these products are required to undergo an evaluation of the potential allergenic activity of the protein(s) that are produced from the introduced genes. The process by which the allergy

Abbreviations: Codex, Codex Alimentarius Commission; ELISA, enzyme-linked immunosorbent assay; FAO/WHO, Food and Agriculture Organization of the United Nations/World Health Organization; GM, genetically-modified.

* Tel.: +302 695 6852; fax: +302 695 3075.

E-mail address: Gregory.s.ladics@usa.dupont.com

assessment has been conducted has involved guidance from a number of expert scientific bodies, including the US FDA (1992), FAO/WHO (1996, 2000, 2001), and Codex (2003). The objectives of the recommendations provided by these expert bodies for assessing the potential allergenicity of foods derived from GM crops are two-fold: (1) protect allergic consumers from accidental exposure to allergens or cross-reactive proteins that may trigger an adverse reaction in those already allergic to such proteins and (2) protect the general population from risks associated with the introduction of genes encoding proteins that are likely to become food allergens.

The approach by which allergy assessment has usually been conducted involves the use of a step-wise, decision tree (Metcalfe et al., 1996; FAO/WHO, 2001). However, to date, no single factor has been recognized as the primary identifier for allergenicity. Therefore, a weight-of-evidence approach, which accounts for a variety of factors and experimental approaches for an overall assessment of the allergenic potential of the new protein, is recommended (Codex, 2003). These assessments are currently based on what is known about food allergens such as history of exposure and safety of the gene product(s) (i.e., Is the source of the gene for the new protein known to induce allergy?), molecular structures (e.g., *in silico* amino acid sequence homology and structural similarity comparison to known human allergens), physicochemical properties such as stability to pepsin digestion *in vitro*, and an estimate of the exposure of the novel protein(s) to the gastrointestinal tract, where absorption occurs (e.g., digestibility, protein abundance in the crop, and food/feed processing effects). Where appropriate, specific IgE binding studies using well-characterized allergic subject serum or skin prick testing can add to the assessment.

2. Amino acid sequence homology

One area that has seen significant progress over the last decade is the adaptation of *in silico* bioinformatic tools to evaluate the allergenicity potential of a novel protein. Most of the major food and respiratory allergens have been identified and cloned. Subsequently, the protein sequences for most, if not all of these allergens have been incorporated into various databases. As a result, novel proteins can be routinely screened for amino acid sequence homology with, and structural similarity to, known human allergens using an array of bioinformatic tools early on in the product development process. Importantly, bioinformatics allows one primary question to be asked: Is the novel protein an existing allergen? Bioinformatics also allows a secondary question to be asked: Is the novel protein likely to cross-react with an existing allergen? Bioinformatics is not intended to answer whether a novel protein will “become” an allergen. Knowledge of the protein structures responsible for inducing sensitization is still lacking even for known allergens.

Typically, sequence homology searches comparing the structure of novel proteins to known allergens in a database are conducted using various algorithms such as FASTA (Pearson and Lipman, 1988; Ladics et al., 2007) to predict overall structural similarities. As recommended by FAO/WHO (2001), IgE cross-reactivity between a novel protein and a known allergen is considered a possibility when there is more than 35% identity over a segment of 80 or greater amino acids. It should also be pointed out, however, that for cross-reactivity to occur, Aalberse (2000) has reported that a high degree of homology is needed, likely in excess of 50–60%, over significant spans of the target protein and allergen.

In addition, step-wise contiguous identical amino acid segment searches are also performed to identify amino acid sequences that may represent linear IgE binding epitopes. IgE binding epitopes,

however, have only been identified for a few allergens. Therefore, in the absence of an IgE binding epitope database, potential epitopes can be evaluated by producing all overlapping peptides of the allergens contained in a particular database and comparing them in a pair-wise manner to all same-size potential peptides of a novel protein using bioinformatic tools. Eight contiguous amino acid matches between a novel protein and a known allergen(s) are routinely used to identify sequences that may represent linear epitopes.

The 2001 FAO/WHO consultation has suggested using a six amino acid match for this type of analysis. The use of a contiguous amino acid match of <8, however, occurs too commonly between unrelated proteins and therefore, is not a reliable criterion for predicting allergenic potential (Hileman et al., 2002; Stadler and Stadler, 2003; Silvanovich et al., 2006; Ladics et al., 2006). Many random matches that are very unlikely to indicate potential IgE epitopes are observed using such a short sequence. For example, Hileman et al. (2002) examining corn proteins reported that using a seven or less contiguous amino acid match resulted in a high rate of false positives (e.g., 84% when using six contiguous amino acids, a clear overestimate for corn, which is considered a minor source of food allergy). Furthermore, in order to cross-link IgE receptors on mast cells and induce an intracellular signal and subsequent release of various mediators (e.g., histamine), two IgE binding epitopes on the same protein are necessary. Therefore, a single match of contiguous amino acids (e.g., 6, 7, or 8) observed between a novel protein and an allergen may not be clinically relevant.

3. *In vitro* pepsin digestion assay

The first published application of an *in vitro* pepsin digestion assay to address the question of food allergen stability was by Astwood et al. (1996). In this study, a number of food allergens or their peptide fragments were shown to be resistant to digestive proteolysis *in vitro*. Since this initial report, there have been several studies repeating the pepsin digestion assay for a variety of proteins (Buchanan et al., 1997; Kenna and Evans, 2000; Okunuki et al., 2001; Fu, 2002). A number of variations in SGF assay parameters have been reported and include: differences in the pH of the assay, the purity of the pepsin, the pepsin to target protein ratio, the target protein purity, and finally, the method of detection. Recently, Thomas et al. (2004) established a standardized protocol for evaluating the *in vitro* digestibility of proteins in SGF in the context of an interlaboratory study. Although a correlation between resistance to pepsin digestion and allergenic potential has been proposed (Astwood et al., 1996) the relationship is not absolute (Fu et al., 2002). There are examples of proteins in food that are not digestible and do not illicit food allergy. Likewise, it cannot be concluded that allergenic food proteins are necessarily more resistant to digestion. There is recent work in kiwi that suggests that these unstable allergens may be protected from pepsin digestion by components of the food matrix (Polovic et al., 2007). Therefore, measurement of protein digestibility should not be regarded as a stand-alone endpoint for the safety assessment of novel proteins. Instead, a weight-of-evidence approach, which accounts for a variety of factors and experimental approaches for an overall assessment of the allergenic potential of the new protein should be utilized, as no single factor has been recognized as the primary identifier for allergenicity (Codex, 2003).

4. Specific IgE sera screening studies

IgE cross-reactivity between a novel protein and a known allergen is currently considered a possibility when there is 35% or

greater identity over 80 or more amino acids or a match of 8 or greater contiguous identical amino acids (Metcalfe et al., 1996; FAO/WHO, 2001; Codex, 2003). To investigate the potential protein cross-reactivity further, specific IgE screening studies with sera from subjects allergic to the identified allergen are conducted. IgE-specific Western blots and ELISAs are typically used to identify allergenic proteins. Conducting 'scientifically meaningful' IgE serum screens, however, are problematic due in part to the lack of standardized methods and the availability of well-characterized human serum samples from a sufficient number of patients since the incidence of individuals with specific allergies is quite low (i.e., 1–2% of adults and 4–6% of children) (Goodman and Leach, 2004). In the case of a major allergen (i.e., one to which more than 50% of individuals sensitive to that material respond in IgE-specific immunoassays), FAO/WHO (2001) indicated that a minimum of 8 relevant sera is required to achieve a 99% certainty that a new protein is not an allergen. FAO/WHO (2001) further indicated that a minimum of 24 relevant sera might be required to achieve a 99% certainty with regard to a minor allergen. Such quantities of well-characterized sera may not be available for testing purposes. Standard criteria are also needed for selection of allergic individuals (Eigenmann and Sampson, 1997). It is also imperative that during assay development, conditions be optimized (e.g., detection methods, blocking reagents) in order to establish assay sensitivity and specificity and to minimize false-positive IgE binding.

In addition to direct binding studies, inhibition Western blot assays and ELISAs should also be conducted to demonstrate the specificity of any identified IgE binding and to further minimize false-positive findings. These studies involve holding the solid-phase protein target constant, while varying the concentration of soluble inhibitors (e.g., pure specific proteins, extracts, or non-related control proteins) mixed with allergic subject sera to generate an inhibition curve in a quantitative assay. Alternatively, such studies can be conducted in a qualitative or semi-quantitative fashion by employing a fixed concentration of inhibitor (Goodman and Leach, 2004; Scheurer et al., 1999). Following the incubation of allergic subject serum with the various inhibitors, the disappearance of previously observed IgE binding indicates a lack of specificity/affinity of the IgE for the target protein. Conducting inhibition studies is important, as the broad range of IgE binding that sometimes occurs may not be clinically relevant (Taylor and Hefle, 1996; Bindslev-Jensen and Poulsen, 1997; Pasini et al., 2000), particularly for cross-reactivity associated with carbohydrate determinants (Vieths et al., 2002).

5. Endpoints requiring further evaluation/validation

5.1. Animal models to evaluate potential protein allergenicity

A number of investigators have recently begun to develop models for predicting or ranking the potential allergenicity of food proteins in several different species. Several models have been proposed using rats (Atkinson et al., 1996; Knippels et al., 1998; Knippels et al., 2000), mice (Dearman et al., 2000; Dearman and Kimber, 2001), swine (Helm et al., 2002), or dogs (Buchanan et al., 1997). Most of these models are based on the assessment of induced antibody (i.e., IgE or IgG) responses and the frequency of responders in the test groups. Presently, none of the animal models have been tested with a wide range of allergens and non-allergens, and there is a lack of data on the reproducibility and predictive value (sensitivity and specificity) of any of the models. A number of relevant questions still remain and include: (1) what is the most appropriate endpoint or design for an animal model? (2) What constitutes the measurement of a positive allergic response? (3) What is the most appropriate route of exposure or form of protein

to test (i.e., isolated pure protein vs. protein in the food matrix). Thus, no animal model has been extensively evaluated and validated.

5.2. Targeted serum screens

The Joint FAO/WHO (2001) consultation recommended using targeted serological screening. Targeted sera screening was recommended in order to help identify potentially cross-reactive proteins using sera from patients allergic to substances broadly related (i.e., dicots, monocots, fungi/yeast, vertebrates, invertebrates) to the source of the gene. As noted above, an issue of critical importance to any sera screening study is the availability of well characterized, quality human sera from a sufficient number of patients. Furthermore, the positive and negative predictive value of the targeted sera screening method has not been demonstrated and the method may have no clinical significance. For example, some types of IgE binding found in these assays may not be clinically relevant (affinity, single vs. multiple epitopes) and thus can lead to a high degree of false positives or false negatives (Taylor and Hefle, 1996; Bindslev-Jensen and Poulsen, 1997; Pasini et al., 2000). In addition, the occurrence of cross-reactive IgE antibodies is not always correlated with clinical food allergy, particularly for cross-reactivity for carbohydrate determinants (Vieths et al., 2002). Thus, the predictive value or clinical relevance of using targeted serum screens has not been established, and was identified by Codex (2003) as one technique requiring further development and validation before being implemented.

In summary, the available allergenicity tests, in the context of a 'weight of evidence approach', provide assurance of the safety of new proteins associated with the agronomic traits introduced into GM crops and food products derived from those crops. As the science and technology advance, additional approaches (e.g., animal models, targeted sera screens) may be utilized; however, to date such approaches have not been extensively evaluated and validated in regard to sensitivity, specificity, and reproducibility in predicting protein allergenicity.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

References

- Aalberse, R.C., 2000. Structural biology of allergens. *J. Allergy Clin. Immunol.* 106, 228–238.
- Astwood, J.D., Leach, J.N., Fuchs, R.L., 1996. Stability of food allergens to digestion in vitro. *Nat. Biotechnol.* 14, 1269–1273.
- Atkinson, H.A.C., Johnson, I.T., Gee, J.M., Grigoriadou, F., Miller, K., 1996. Brown Norway rat model of food allergy: effect of plant components on the development of oral sensitization. *Food Chem. Toxicol.* 34, 27–32.
- Bindslev-Jensen, C., Poulsen, L.K., 1997. In vitro diagnostic methods in the evaluation of food hypersensitivity. In: Metcalfe, D.D., Sampson, H.A., Simon, R.A. (Eds.), *Adverse Reactions to Foods and Food Additives*. Blackwell Science, Cambridge, MA, pp. 72–74.
- Buchanan, B.B. et al., 1997. Thiodioxin-linked mitigation of allergic responses to wheat. *PNAS* 94, 5372–5377.
- Burks, A.W., Sampson, H.A., 1997. Anaphylaxis and food allergy. In: Metcalfe, D.D., Sampson, H.A., Simon, R.A. (Eds.), *Food Allergy: Adverse Reactions to Foods and Food Additives*. Blackwell Science, Cambridge, MA, pp. 25–245.
- Codex Alimentarius Commission, 2003. *Alinorm 03/34: Joint FAO/WHO Food Standard Programme, Codex Alimentarius Commission, Twenty-Fifth Session, Rome, Italy, 30 June–5 July, 2003. Appendix III, Guideline for the conduct of food safety assessment of foods derived from recombinant-DNA plants, and Appendix IV, Annex on the assessment of possible allergenicity*, pp. 47–60.
- Day, P.R., 1996. The biology of plant proteins. *Crit. Rev. Food Sci. Nutr.* 36, 39–67.
- Dearman, R.J., Caddick, H., Basketter, D.A., Kimber, I., 2000. Divergent antibody isotype responses induced in mice by systemic exposure to proteins: a comparison of ovalbumin with bovine serum albumin. *Food Chem. Toxicol.* 38, 351–360.
- Dearman, R.J., Kimber, I., 2001. Determination of protein allergenicity: studies in mice. *Toxicol. Lett.* 120, 181–186.

- Eigenmann, P.A., Sampson, H.A., 1997. Adverse reactions to foods. In: Kaplan, A.P. (Ed.), *Allergy*. W.B. Saunders, Co., Philadelphia, PA, pp. 542–572.
- FAO/WHO, 1996. Biotechnology and Food Safety Report of a Joint FAO/WHO Consultation, Rome, Italy, 30 September–4 October 1996. FAO Food and Nutrition Paper 61. Food and Agriculture Organisation of the United Nations, Rome.
- FAO/WHO, 2000. Safety aspects of genetically modified foods of plant origin. Report of a Joint FAO/WHO Expert Consultation on Foods Derived from Biotechnology, Geneva, Switzerland. Food and Agriculture Organization of the United Nations, Rome.
- 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, January 22–25, 2001, Rome, Italy.
- Fu, T.J., 2002. Digestion stability as a criterion for protein allergenicity assessment. *Ann. NY Acad. Sci.* 964, 99–110.
- Fu, T.J., Abbott, U., Hatzos, C., 2002. Digestibility of food allergens and non-allergenic proteins in simulated gastric and intestinal fluids – a comparative study. *J. Agr. Food Chem.* 50, 7154–7160.
- Goodman, R.E., Leach, J.N., 2004. Assessing the allergenicity of proteins introduced into genetically modified crops using specific human IgE assays. *J. AOAC Int.* 87 (6), 1423–1432.
- Hefle, S.L., Nordlee, J.A., Taylor, S.L., 1996. Allergenic foods. *Crit. Rev. Food Sci. Nutr.* 36, 69–89.
- Helm, R.M. et al., 2002. A neonatal swine model for peanut allergy. *J. Allergy Clin. Immunol.* 109, 135–142.
- Hileman, R.E., Silvanovich, A., Goodman, R.E., Rice, E.A., Holleschak, G., Astwood, J.D., Hefle, S.L., 2002. Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* 128, 280–291.
- Kenna, J.G., Evans, R.M., 2000. Digestibility of proteins in simulated gastric fluid. *Toxicologist* 54 (1), 7–9 (Abstract 666).
- Knippels, L.M.J., Penninks, A.H., Spanhaak, S., Houben, G.F., 1998. Oral sensitization to food proteins: a Brown Norway rat model. *Clin. Exp. Allergy* 28, 368–375.
- Knippels, L.M.J. et al., 2000. Comparison of antibody responses to hen's egg and cow's milk-proteins in orally sensitized rats and human patients. *Allergy* 55, 251–258.
- Ladics, G.S., Bardina, L., Cressman, R.F., Mattsson, J.L., Sampson, H.A., 2006. Lack of cross-reactivity between the *Bacillus thuringiensis* derived protein Cry1F in maize grain and dust mite Der p 7 protein with human sera positive for Der p 7-IgE. *Regul. Toxicol. Pharmacol.* 44 (2), 136–143.
- Ladics, G.S., Bannon, G.A., Silvanovich, A., Cressman, R.F., 2007. Comparison of conventional FASTA identity searches with the 80 amino acid sliding window FASTA search for the elucidation of potential identities to known allergens. *Mol. Nutr. Food Res.* 51 (8), 985–998.
- Metcalfe, D.D., Astwood, J.D., Townsend, R., Sampson, H.A., Taylor, S.L., Fuchs, R.L., 1996. Assessment of the allergenic potential of foods from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* 36 (S), 165–186.
- Okunuki, H. et al., 2001. Increased digestibility of two products in genetically modified food (CP4-EPSPS and CryIAb) after preheating. *J. Food Hyg. Soc. Jpn.* 43, 68–73.
- Pasini, G., Simonato, B., Giannattasio, M., Gemignani, C., Curioni, A., 2000. IgE binding to almond proteins in two CAP-FEIA-negative patients with allergic symptoms to almond as compared to three CAP-FEIA-false-positive subjects. *Allergy* 55, 955–958.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological comparison. *Proc. Natl. Acad. Sci. USA* 85, 2440–2448.
- Polovic, N., Blanus, M., Gavrovic-Jankulovic, M., Burazer, et al., 2007. *Clin. Exp. Allergy* 37, 764–771.
- Sampson, H.A., 1997. Immediate reactions to foods in infants and children. In: Metcalfe, D.D., Sampson, H.A., Simon, R.A. (Eds.), *Food Allergy: Adverse Reactions to Foods and Food Additives*. Blackwell Science, Cambridge, MA, pp. 169–182.
- Scheurer, S., Son, D.Y., Boehm, M., Karamloo, F., Franke, S., Hoffman, A., Hausteiner, D., Vieths, S., 1999. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. *Mol. Immunol.* 36, 155–167.
- Sicherer, S.H., 2000. Determinants of systemic manifestations of food allergy. *J. Allergy Clin. Immunol.* 106, 251–257.
- Silvanovich, A., Nemeth, M.A., Song, P., Herman, R., Tagliani, L., Bannon, G.A., 2006. The value of short amino acid sequence matches for prediction of protein allergenicity. *Toxicol. Sci.* 90 (1), 252–258.
- Smith, L.J., Munoz-Furlong, A., 1997. The management of food allergy. In: Metcalfe, D.D., Sampson, H.A., Simon, R.A. (Eds.), *Food Allergy: Adverse Reactions to Foods and Food Additives*. Blackwell Science, Cambridge, MA, pp. 431–444.
- Stadler, M.B., Stadler, B.M., 2003. Allergenicity prediction by protein sequence. *FASEB J.* 17, 1141–1143.
- Taylor, S.L., Hefle, S.L., 1996. Will genetically modified foods be allergenic? *J. Allergy Clin. Immunol.* 107, 765–771.
- Thomas, K. et al., 2004. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* 39, 87–98.
- US FDA, 1992. Statement of policy: foods derived from new plant varieties: notice. *Federal Register* 57, pp. 22984–23005.
- Vieths, S., Scheurer, S., Ballmer-Weber, B., 2002. Current understanding of cross-reactivity of food allergens and pollen. *Ann. NY Acad. Sci.* 964, 47–68.