



Performing IgE serum testing due to bioinformatics matches in the allergenicity assessment of GM crops

Richard E. Goodman *

Food Allergy Research and Resource Program, Department of Food Science and Technology, University of Nebraska, 143 Food Industry Complex, Lincoln, USA

ARTICLE INFO

Keywords:

IgE
Allergenicity
Genetically modified
Bioinformatics

ABSTRACT

Proteins introduced into genetically modified (GM) organisms through genetic engineering must be evaluated for their potential to cause allergic disease under various national laws and regulations. The Codex Alimentarius Commission guidance document (2003) calls for testing of serum IgE binding to the introduced protein if the gene was from an allergenic source, or the sequence of the transferred protein has >35% identity in any segment of 80 or more amino acids to a known allergen or shares significant short amino acid identities. The Codex guidance recognized that the assessment will evolve based on new scientific knowledge. Arguably, the current criteria are too conservative as discussed in this paper and they do not provide practical guidance on serum testing. The goals of this paper are: (1) to summarize evidence supporting the level of identity that indicates potential risk of cross-reactivity for those with existing allergies; (2) to provide example bioinformatics results and discuss their interpretation using published examples of proteins expressed in transgenic crops; and (3) to discuss key factors of experimental design and methodology for serum IgE tests to minimize the rate of false negative and false positive identification of potential allergens and cross-reactive proteins.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

Varieties of genetically modified cotton, maize, soybeans and canola contain genes encoding specific proteins to dramatically improve farming efficiency by improving reduced tillage with herbicide tolerance or to reduce the need for chemical pesticides. These varieties are now grown on over 68 million hectares around the world (James et al., 2003). Potential new products are under development, such as Golden Rice, which has been modified to produce significant levels of the provitamin A (β -carotene), aim to improve the health of individuals who subsist on rice and have inadequate intake of alternative sources. Golden rice was modified by the introduction of genes encoding two enzymes from other sources (Hoa et al., 2003). Release for cultivation and the importation and trade of such crops follow years of development and testing, including specific measures to evaluate safety and reduce the possibility of unintended harm due to potential increases in allergenicity or toxicity (Goodman et al., 2005; National Academy of Sciences, USA, 2004; Konig et al., 2004).

Abbreviations: Aa, Amino acid; DBPCFC, Double-blind placebo-controlled food challenge; CTP, Chloroplast transit peptide; GM, Genetically modified; HDM, House dust mite; kU, Kilo-unit; PR, Pathogenesis related; SPT, Skin prick test.

* Tel.: +1 402 472 0452.

E-mail address: rgoodman2@unl.edu

International guidelines for the evaluation of the potential allergenicity of GM crops were developed by the Codex Alimentarius Commission of the Food and Agricultural Organization and World Health Organization of the United Nations (Codex, 2003). The end-point of the assessment is a conclusion as to the likelihood that the newly introduced protein is an allergen. The strategy is a multistep process. First is to evaluate the source of the gene as to whether it is a potential source of allergens causing food, airway or contact symptoms. If reports of allergy are common or severe serum IgE binding studies would be performed using serum from individuals allergic to the source. If the source is wheat, barley or rye an evaluation of the protein is necessary to ensure it does not elicit gluten-induced enteropathy. In all cases the sequence of the protein is compared to known allergenic proteins and if the identity is above the threshold, serum testing would be performed using sera from subjects allergic to the matched allergen. If the source is not commonly allergenic and the sequence does not match a proven allergen no serum testing is performed. Additionally the stability of the protein is tested in an in vitro pepsin digestion assay. Proteins of high stability are considered to have a greater risk of allergy. An additional consideration is the abundance in food fractions of the crop as highly abundant proteins are more likely to be food allergens. Development of the allergenicity assessment strategy was reviewed recently with a discussion of the strengths and weaknesses of the Codex guidance document

(Goodman et al., 2005; Goodman and Hefle, 2005). The Codex document is an overall guideline and it does not contain specific details about appropriate serum donor selection, assay design, target protein characterization or interpretation of results, which will be discussed here.

2. Allergenic Source of the Gene May Trigger Serum Testing

If the source of the gene transferred into the GM crop is a common airway allergen, food allergen or contact allergen (e.g. grass pollen, peanut or rubber tree, respectively), an evaluation is made as to the likelihood that appropriate serum donors may be found. Serum IgE binding tests would be performed to determine whether the protein is a likely cause of allergy, whether all allergens from the source have been identified and characterized or not. Commonly allergenic proteins would present a greater risk to the population as a greater number of people would be sensitized. The definition of common has not been explicitly defined. For foods a reasonable guideline would be food allergen labeling laws in various countries. Otherwise the literature and clinical allergy experts may be consulted in deciding whether such testing is required. Sources from the same genus or certainly the same species described as commonly allergenic in the literature may be viewed as equal. If a gene was transferred from annual bluegrass, *Poa annua*, it should be considered a common allergen even though most reactions described in the literature have tested *Poa pretensis*, Kentucky bluegrass as their proteins are likely to be cross-reactive. But less closely related organisms are less likely to be cross-reactive. The more specific consideration is whether the introduced protein is similar in sequence and therefore structure to any known allergen (described below). Whether serum testing would be triggered by the source of the gene or sequence identity, the principles of serum IgE testing are similar.

3. Identification of Potential Allergenic Proteins Through Bioinformatics

A bioinformatics comparison of proteins introduced into food crops through genetic engineering versus those of known allergens provides an efficient mechanism to identify those proteins that may present an increased risk of allergic reactions for individuals with existing allergies. The sequence comparison may be performed based on bioinformatics testing with key-word association with allergens from non-specialized public protein databases, but normally it would be performed specifically against a specialized allergen database such as <http://www.allergenonline.com> (Goodman et al., 2005). The goal of the comparison is to determine if there is already evidence the GM protein is an allergen, or if the sequence and structure is so similar to an allergen that it is likely to induce allergic cross-reactions.

Three comparative bioinformatics approaches have been used, or suggested for regulatory safety studies by various authors. One is a search for identical short (6–8) amino acid segments of the query protein that exactly match any segment of a known allergenic protein, the assumption being this is close to the smallest size of a potential IgE binding epitope. The second is an overall primary sequence comparison using FASTA or BLAST. The criterion recommended by FAO/WHO (2001) and Codex (2003) is 35% identity over any segment of 80 amino acids (aa) between the GM protein and any allergen. However, the empirical evidence indicates that clinical cross-reactivity is rare for proteins sharing less than 50% identity over their entire length (Aalberse, 2000). The third approach is based on three dimensional prediction programs or motif matching algorithms with the intent to enhance the likelihood of identifying potential conformational epitopes (Stadler and Stadler,

2003; Ivanciuc et al., 2003) and is intended to identify three dimensional similarities that might predict potential cross-reactivity. The utility of each of these approaches was debated in a bioinformatics workshop coordinated by HESI in early 2005 (Goodman, 2006). The unanimous opinion of the more than 40 participating scientists was that the short segment match (e.g. 6 to 8 aa) provides an unacceptably high rate of false positive matches (Hileman et al., 2002; Stadler and Stadler, 2003; Silvanovich et al., 2006), an undetermined rate of true positive matches and therefore the search for short identical segments (e.g., 6 or 8 aa) should be dropped from the allergenicity assessment process (Goodman, 2006). There was also general agreement that the most practical approach at this time is to compare the primary amino acid sequence identities to known allergens using either FASTA or BLAST as the most simple and generally predictive methods. While there was good agreement that identities of just greater than 35% over 80 or more aa (recommended by Codex in 2003) is quite conservative, the conclusion was that additional data a literature review or specific carefully controlled studies with a number of test cases may justify increasing the identity criterion. Currently there are few studies that demonstrate a correlation of co-reactivity for proteins sharing as little as 40% identity that are significantly cross-reactive based on in vitro inhibition (Goodman and Hefle, 2005). Little evidence has been presented to demonstrate the ability of three dimensional structural comparison programs to predict potential allergenicity for proteins that have not been analyzed by crystallography or NMR or for proteins that are not highly similar to known allergens. There seemed to be general agreement that further studies would be needed to test motif matching and three dimensional or surface plot predictions to evaluate the predictive value compared to the overall value of FASTA or BLAST algorithms. It is essential to note that the most efficient searches are performed against special databases of allergens (e.g. AllergenOnline.com), however because all databases have a lag-time in updating, a final step with a BLAST comparison to NCBI protein sequences with a restricted keyword delimiter of “allergen” has been recommended to catch any missed sequence matches (Goodman and Hefle, 2005). The end result of the bioinformatics search is a decision as to the need for specific serum testing and if so, the specific allergic population that should be used to collect serum samples (Goodman et al., 2005). But since appropriate serum testing is not trivial, correct interpretation of bioinformatics findings are important.

3.1. Bioinformatics analysis of potential GM products

A number of food crops have been transformed with genes encoding proteins that have not yet been completely evaluated or taken through a full regulatory review process. Some of those may not prove efficacious in field trials or for some other reason may never reach regulatory review, let alone become commercially available. Summary results from comparisons of four GM examples using published, not confirmed sequences are presented below along with recommendations for further actions. The search routines freely available at <http://www.AllergenOnline.com> were used with version 6.0 of the database. Default parameters were used and each sequence was tested by overall FASTA and using a sliding window of 80 amino acids.

Example 1: A globulin protein gene was transferred from *Amaranthus hypochondriacus* into potato to improve the amino acid and total protein content of potato (Chakraborty et al., 2000). The highest identity of the full-length 304 amino acid sequence (accession number AF491291.1) to any of the 1537 known or putative allergens of AllergenOnline version 6.0 was 22% to a latex allergen. The highest identity by FASTA using the 80 amino acid sliding window was 35.8% to the latex allergen (accession number CAA11042), a number that is marginally above the 35% identity threshold

recommended by Codex (2003). It is highly unlikely that this level of identity represents a real risk of allergic cross-reactivity, but some regulators might demand serum testing using donors who are allergic to latex.

Example 2: An 11S globulin gene from *A. hypochondriacus* was transferred into maize to improve the nutritional quality (Rascon-Cruz et al., 2004). The protein (amarantin) expressed in *A. hypochondriacus* is processed into an acidic and basic subunit with almost none of the proamarantin present, but in transgenic maize the proamarantin (53 kDa) is the dominant form and the 32 kDa acidic and 21 kDa basic subunits are almost absent unless digested by exogenous proteases. The developers performed a short amino acid matching routine recommended by FAO/WHO (2001) and found matches of six amino acid peptides to legumin-like allergenic proteins in buckwheat, hazelnut, soybean and peanut, matches of 7 amino acids to peanut allergen Ara h 3, a soybean allergen and buckwheat allergens and two eight amino acid matches to an 11S allergen of hazelnut (Sinagawa-Garcia et al., 2004). The authors examined published data of the allergenic seed storage proteins to determine whether any of the matched peptides were in known IgE epitopes. Since they were not, their conclusion was that the data did not demonstrate a significant risk of allergy. However, when the sequence was tested using the FASTA algorithms and the full-length sequence (53 kDa, accession number X82121) full-length protein identity matches of between 49% and 52% were found to 11S globulin allergens in cashew, buckwheat, Brazil nut and hazelnut. The highest 80 aa match was 70% to hazelnut protein (accession number AAL73404). This data indicates a fairly high likelihood of IgE cross-reactivity (Aalberse, 2000). Based on these results, serum IgE binding studies should be undertaken or the crop should not be released for commercial production.

Example 3: An alpha-amylase inhibitor was transferred from kidney bean into green peas in order to protect against losses due to a storage beetle pest (Schroeder et al., 1995). This protein is processed after the removal of a signal peptide, into an alpha- and a beta-subunit that are held together by non-covalent interactions along, primarily as a hetero tetramer (two alpha- and two beta-chains) and both subunits may be glycosylated at specific asparagine residues (Young et al., 1999). Peptide lengths after processing may differ slightly and complex glycan structures vary some, as the protein is expressed in green peas (Prescott et al., 2005). Comparison of the sequence of the intact 223 aa precursor (full-length minus the signal peptide, accession # AAA33769) using the AllergenOnline website with version 6.0 identified low sequence identity overall matches and matches greater than 35% identity over 80 aa windows with a soybean lectin prepeptide (accession # AAA33983 at 39% overall, 44% best 80 aa window), peanut agglutinin precursor (accession # AAB22817 at 35% overall, 41% best 80 aa window) and peanut mannose binding protein (accession # AAA74576 at 28% overall, 37% best 80 aa window). While the results clearly demonstrate evolutionary homology, questions remain regarding the relevance of the matches. We recently reviewed all the entries in our AllergenOnline version 6 database, evaluating published data on the allergenicity of each protein entry (unpublished). Shibasaki et al. (1992) tested human IgE binding and leukocyte histamine release to both soybean lectin and peanut agglutinin and were not able to demonstrate specific binding or histamine release to either. Zavasal and Krauz (1985) found that soybean lectin was not able to precipitate human IgE (no binding), did not elicit positive skin prick tests, but did cause rat mast cell degranulation only when the rats were sensitized and had high IgE levels. That suggests soybean lectin may act on rat IgE, but not human IgE. Burks et al. (1994) demonstrated very low level IgE binding to a partially purified peanut agglutinin. However, the ability of soluble agglutinin to inhibit IgE binding

to crude peanut extract, or purified Ara h 2 was less than 1:1000 and 1:10,000, respectively, leaving serious doubts on the relevance of agglutinin to peanut allergy. No published studies were found indicating allergic responses or IgE binding to peanut mannose binding protein. Based on the review, the soybean lectin and peanut mannose binding protein was removed from the AllergenOnline database version 7.0 (available online since January 2007). Peanut agglutinin remains in as a putative allergen based on the weak IgE binding in the Burks study. The reviews of the scientific data regarding the allergenicity of the sequences that match the aAI sequence do not support a need to perform serum testing to evaluate potential cross-reactivity. However, the differences noted in complex glycan structure between of the protein as expressed in kidney bean (Young et al., 1999) compared to green pea (Prescott et al., 2005) may merit further investigation.

Example 4: Three genes were transferred into rice to generate the improved version of Golden Rice 2 (Paine et al., 2005). The purpose was to make GM rice capable of producing biologically relevant levels of provitamin A to correct the dietary deficiencies in rice-based societies (<http://www.goldenrice.org/index.html>). The sequences of the proteins produced from the three introduced genes have been tested against the AllergenOnline version 6.0 database by overall FASTA and a sliding window of 80 amino acids (available as a PDF at <http://www.allergenonline.com>). None of the three proteins matched any known or putative allergen at 35% or more identity.

Two additional reasons for performing serum tests of GM crops have been proposed in addition to sequence matches (FAO/WHO, 2001; Codex, 2003). One is when the gene is taken from a source that is known as a common cause of allergy, even if the sequence of the protein does not match a known allergen. The utility of such testing was demonstrated in the case when Pioneer-Hibred introduced a gene from Brazil nut into soybeans to enhance the nutritional value for animal feed. Since allergens from Brazil nuts were not known, serum and skin prick tests were conducted using subjects with allergy to Brazil nut, to ensure that the protein was not an allergen and the results were clearly positive (Nordlee et al., 1996). Because of those results the potential product was not submitted for regulatory review, nor was it ever marketed. The second was a proposal for a targeted serum test using serum from subjects allergic to organisms broadly related to the source of the GM gene (FAO/WHO, 2001). The rationale was that such tests might uncover cross-reactivity that could not be predicted based on bioinformatics using currently available data as some real allergens have not been studied. Since protein structure is broadly conserved, the proponents reasoned there might be individuals who are sensitized to evolutionarily conserved proteins with a taxonomic order or class. However, based on the rare occurrence of cross-reactivity of proteins across broad taxonomic groups (families or higher) there is little reason to believe such testing would enhance the assessment, and in fact, due to the technical imperfections of serum testing, is more likely to increase the background false positive IgE detection rate (Goodman and Hefle, 2005). Further, the decision to do serum testing is based on the combination of evaluating the source of the gene (allergenic or not) and sequence similarity. In combination these considerations markedly reduce the risk of introducing an important allergen and based on that understanding, and expected low predictive value of broadly targeted serum testing the Codex panel did not follow the FAO/WHO (2001) recommendation.

4. General Considerations for Serum Testing

Regardless of the bioinformatics method used to match similarities or the scoring criteria, once it is determined that serum testing

is required a variety of issues must be addressed to maximize the utility of testing. Allergic sensitization (development of antigen-specific IgE) and elicitation (induction of a specific allergic response) relative to any one allergen vary dramatically across the global population. It is clear that an objective assessment must take into consideration the diverse nature of the allergic subjects and allergens to evaluate the potential risk to those most at-risk. The focus must be on subjects who are most likely to have been sensitized to the protein under review, or who are likely sensitized to a very similar protein. Serum testing in the context of the safety evaluation is intended to identify proteins that are already allergens or are sufficiently similar to an allergen that an individual sensitized to the similar protein may suffer an allergic cross-reaction when exposed to the new protein. The focus is on IgE mediated reactions and it is clear that allergen-specific IgE binding is necessary but not sufficient to cause an allergic reaction. It takes at least two IgE epitopes on an individual protein to cross-link high affinity IgE receptors on mast cells and there is some broad range of affinity of IgE binding necessary for optimum triggering. The nature of the epitopes is important as well. Specific IgE will bind to sequential amino acid structures (linear or sequential epitopes), to discontinuous amino acid structures (conformational epitopes) or to specific sugar structures on an asparagine-linked (N-linked) glycan. Different assay formats are required to improve the likelihood of identifying binding to discontinuous epitopes compared to sequential epitopes. The relevance of IgE binding to glycans is debated, but has been demonstrated to be inconsequential for many proteins (van Ree, 2002; Ballmer-Weber et al., 2001). It is important to consider the diagnostic predictive value of specific-IgE and even skin prick tests when designing serum tests for the allergenicity assessment. Based on controlled studies it turns out that the predictive value of different commercial *in vitro* IgE assays as well as skin test reagents can produce results that vary markedly (Sampson and Ho, 1997). Results and discussion from a number of investigators should be interpreted as a need for careful validation of reagents and assays, selection of well characterized serum donors as test subjects and the need for care in selecting positive threshold values in performing predictive allergenicity assessments (Plebani et al., 1998; Dolan, 2003; Purohit et al., 2005; Linneberg et al., 2006).

The number of potentially allergic subjects must be taken into account as it directly relates to the potential pool of relevant serum donors. The incidence of allergy to any single allergenic source may be as high as 0.5 to 2% for fish, milk, eggs and peanuts, depending on whether children or adults are considered (Sampson, 2004), but is probably much lower than 0.1% for many common foods including peas, rice, maize, beef, chicken although hard data are not available. Proteins are the targets of IgE antibody binding, either directly to specific structures made up of sequential or non-contiguous amino acids, or to specific glycan structures on some glycoproteins. Complex foods include hundreds to thousands of different proteins, yet it is common to find from one to less than ten proteins in any given food being responsible for most (or all) allergic reactions. Most of the allergenic proteins fall into a few structural families of proteins although each of the structural families also contains apparently non-allergenic proteins as well (Jenkins et al., 2005). For instance in peanuts, one of the most allergenic of foods, there are 3 major allergenic protein types: the 2S albumins (conglutins) including Ara h 2, Ara h 6 and Ara h 7, the vicilin Ara h 1 (a cupin) and the glycinins Ara h 3 and Ara h 4 (Scurlock and Burks, 2004; Koppelman et al., 2005) and all are abundant seed storage proteins. Ara h 1, 2 and 3 are thought to be responsible for most reactions to peanuts, based on multiple IgE binding studies. It should be recognized that it would not be possible to differentiate between IgE binding to Ara h 2 rather than Ara h 6 or Ara h 3 rather than Ara h 4 in most studies as the proteins are so sim-

ilar that even inhibition studies with pure proteins may not be definitive. A small number of other peanut proteins have been implicated in allergy of a small number of peanut allergic subjects: Ara h 5 (a profilin), Ara h 8 (a Bet v 1 homolog), oleosin and lipid transfer protein. Maize seems to have only one major allergen, a lipid transfer protein (Pastorello et al., 2003) although there is at least one report that a different corn protein may cause allergic reactions to dietary corn (Pasini et al., 2002).

Three important conclusions may be drawn from this evidence relative to a strategy for assessing the potential allergenicity of any specific protein including those introduced into a food via genetic engineering. First, that only a very restricted number of proteins in an allergenic source are responsible for most allergic reactions making the choice of serum donors extremely important. Second, the risk of allergy due to different proteins is highly graded. There are a few very potent and common allergens and a somewhat larger number of apparently less potent allergens, and a huge number of proteins that apparently do not cause allergic reactions. Third, because of the low incidence of allergy to any specific protein, it is extremely important to make sure that the right serum donor population is selected if testing is required or the likelihood of finding specific IgE binding will be low.

The method(s) chosen to measure antigen-specific IgE binding are also important in determining the utility of any testing. Experiences in the development of appropriate diagnostic assays have demonstrated that care is needed in the preparation, characterization and standardization of the protein test material, the design of the assay and selection of appropriate positive and negative controls. Lack of care may lead to higher probability of false negative or false positive results that might lead to the unnecessary rejection of a product or risk to the allergic consumers. Other critical factors are the specificity of the IgE detection reagents, blocking of non-specific IgE binding and definition of limits of sensitivity for detection as well as evaluation of relevant positive levels in the specific assay (Goodman and Leach, 2004).

5. Protein Test Materials

Allergenic proteins differ in the structure of the epitope (discontinuous, sequential or carbohydrate) and the abundance of epitopes. Those characteristics influence the likelihood of allergic reactions and can dramatically influence the IgE binding in an IgE binding assay. The method of assay will directly influence the probability of detecting relevant IgE binding to the target allergens. The specificity and strength of binding of the IgE antibody to the epitope is graded, not absolute. IgE can bind to a structure more or less similar to the biologically relevant allergenic epitope with low affinity that is unlikely to cause an allergic reaction. Since there is no way to know *a priori* what types of epitopes are important for any single allergen (except whether the protein includes a glycan), appropriate conditions and test materials must be selected to evaluate all types of antibody-epitope interactions if there were justification for testing to evaluate IgE binding. The ideal situation would be to obtain highly purified protein in native conformation from the source material that will be used on food, but that is not always possible. It is important to understand the protein of interest. Is it a simple monomeric protein, homo-multimeric or heteromeric? Does it contain inter-molecular or intra-molecular disulfide bonds? How abundant is it in the GM crop? Is it glycosylated? That knowledge is useful for selecting the right test material and for choosing the best IgE binding assay format.

For example, Golden Rice version 2 was modified by inserting one gene from maize (*Zea mays*), phytoene synthase and one from a bacterium (*Erwinia uredovora*), carotene desaturase, which was fused to the chloroplast-transit peptide of the ribulose-1,5-bis-

phosphate decarboxylase gene of the garden pea (*Pisum sativum*) in order to produce high levels of β -carotene (Paine et al., 2005). An additional marker gene, phosphomannitol isomerase from another bacterium (*Escherichia coli*) was also inserted for rapid selection of successfully transformed plants (Paine et al., 2005). Enzymes such as these are expressed at relatively low levels in plant tissue and it would be extremely difficult to purify sufficient native highly purified protein from the food crop, such as rice, in order to perform a number of tests. In some cases it might be necessary to produce the protein in a heterologous source such as *E. coli*, SF8 insect cells or tobacco in order to meet test needs. If the test is designed for one of those proteins it may not be possible to purify the protein in sufficient quantities. While those proteins are simple monomers, the chloroplast transit peptide (CTP) should be cleaved from the carotene desaturase. The CTP is probably degraded immediately and does not stay associated with the remainder of the protein. If the product was a 2S albumin it would be important to look at disulfide bonds. Some 2S albumins are maintained as a single polypeptide chain with intramolecular disulfide bonds. Other 2S albumins are cleaved, but the peptides remain linked due to pre-formed disulfide bonds. If heterologously produced protein is tested, it may be instructive to test IgE binding to a whole plant extract containing the GM protein, compared to IgE binding to a non-GM, but genetically equivalent variety in addition to the heterologously produced protein.

In other cases enough pure protein could be purified from the GM plant. The alpha-amylase inhibitor from kidney bean that was transferred into garden peas to reduce losses from storage beetle pests is expressed at sufficiently high concentration to purify protein for a variety of tests (Morton et al., 2000). If there were reasons to perform serum testing it should be possible to use plant-produced purified protein in this case.

Whether using test material from the GM crop, or from a heterologous source it is important to characterize the material to ensure that the protein is representative of the material that would be in the food and that it be folded properly to test for potential discontinuous (conformational) epitopes. Characterization will include sequencing of the DNA insert, or mRNA transcribed from the insert. It will usually also include demonstration of expected size (by SDS-PAGE and or mass measurement) and at least partial amino acid sequence (N-terminal sequencing and/or MALDI-TOF or MALDI-MS-MS). If the sequence indicates a potential asparagine-linked (N-linked) glycosylation site and the gene encodes a signal peptide, glycosylation or lack of glycosylation will need to be demonstrated by staining and/or mass analysis. If the protein has biochemical or bioactive activity, that should be demonstrated to verify conformational structure.

Additional test materials that are often used include the GM crop material (e.g. seed, herbaceous material, or processed food fraction such as flour), non-transgenic crop material of equivalent genetic background, material from heterologous expression system without the transgene, material from the organism that served as the source of the gene. In the case where the transgene is similar to a known allergen, the allergenic material containing the matched allergen should also be tested.

6. Assay design for IgE binding to GM proteins

Since there is limited published experience on testing specific IgE binding to any GM product, we must turn to well designed studies of non-GM materials that have been used to identify new allergens, or diagnose allergic sensitization. In addition, there are two very recent examples have been performed using methods and overall designs that might be followed (Ladics et al., 2005; Hoff et al., 2007) and the following discussion is meant to review prob-

lem areas and technical issues that should be considered in designing appropriate assays in the future.

6.1. Serum donors

Serum simply must be obtained from individuals with well-documented, clinically-confirmed allergy to the allergen or allergen source of interest. For test subjects the donors should have clear objective symptoms within 20 minutes of exposure to the allergen (e.g. acute urticaria, angioedema, bronchial asthma, diarrhea, emesis or anaphylaxis). A positive skin prick test to an extract of the source of the gene is useful to confirm biologically active IgE to the gene donor, although without corresponding historical symptoms may not be meaningful. If the source of the gene or the source of the bioinformatics-matched, cross-reactive protein is a food, double-blind, placebo-controlled food challenge (DBPCFC) or in some cases single blinded or open challenge testing with the source material is useful to confirm allergies. An in vitro IgE test to the source of the gene, or for proteins sequence matched to an allergen, of the allergenic protein is useful to demonstrate the presence of relevant IgE (e.g. immunoblotting, ELISA or commercial IgE test). It is important to choose donors with significant levels of serum IgE, not just marginally positive by statistical measure. Two classes of negative control donors should be used. One includes individuals without allergy to the source material, but clear allergic symptoms and high specific IgE to an irrelevant allergen. Additionally non-allergic donors should be used to demonstrate background levels of IgE binding in the assay. Individuals with unusually high total IgE (e.g. >2,000 IU/L), such as those with systemic lupus erythematosus or severe atopic dermatitis, should be excluded from these predictive allergy studies as the specificity of their IgE is often broad and unrelated to allergic disease (Elkayam et al., 1995; Laske and Niggemann, 2004), possibly leading to false positive results. While high specific IgE correlates well with allergy to cow's milk and eggs for individuals with atopic dermatitis, it does not correlate very well for wheat, soybean and rice (Hamada et al., 2005; Ikematsu et al., 2006). That could be due in part to irrelevant in vitro IgE cross-reactivity that is "specific" for common N-linked glycan structures, termed CCD (van Ree et al., 2000; van Ree, 2002). Some studies have demonstrated limited biological reactivity of IgE to the carbohydrate determinants and an apparent high rate of false positive detections relative to allergic diagnosis, especially for food allergy (Lauer et al., 2004). The HESI workshop participants, including key experts (B.K. Ballmer-Weber, L. Poulsen, R. van Ree and S. Vieths) suggested taking steps to exclude those with high IgE to carbohydrates in selecting serum donors for the allergenicity assessment, or at least identifying which donors bind glycan. If they are included it may be possible and advisable to determine whether any IgE binding to a glycan structure on a GM protein is biologically relevant using in vitro inhibition and basophil histamine release assays to evaluate relevance (Lauer et al., 2004; Ballmer-Weber et al., 2005). An interesting finding from Laske and Niggemann (2004) that has been found in other studies is that the negative predictive value of a specific IgE test is usually higher than the positive predictive value. In other words a specific serum IgE test result that is negative is more likely to represent a true negative clinical status than a positive result is to reflect a true allergic status, assuming the target antigen was appropriately represented in the assay.

6.2. Statistical Design

The number of serum donors that may be found and recruited for a given study will often be below the ideal number for highly predictive outcomes. Binomial probability theory may be used to establish the ideal sample number needed for prediction at various

specific confidence levels, the likelihood that a single individual from the allergic-specific population will test positive to the test material (i.e., pass/fail) if some simple assumptions are made about the prevalence of sensitivity to a specific protein in the sample (Metcalf et al., 1996). For simplicity an allergic incidence of 20% and a 95% level of confidence were assumed to choose the sample number (Metcalf et al., 1996). The 20% incidence was considered as representing a minor or moderate frequency allergen within a given allergenic source. With those assumptions Metcalf et al. (1996) calculated that a negative result from 14 specifically allergic subjects would provide 95% level of confidence that the negative finding was correct. If you assume 50% incidence, the level of certainty increases from 95% to 99.9%. The 50% incidence was assumed for major allergens such as Ara h 2 of peanut. Metcalf et al. (1996) recognized that for less commonly allergic sources such as rice or corn, it is extremely unlikely that 14 specifically allergic individuals would be identified and they suggest a minimum of five allergic subjects in such cases, realizing that there would only be a 95% probability of finding a major allergen (reactive in more than 50% of corn or rice allergic individuals) at that sampling level. The FAO/WHO expert panel consultation recommended higher subject numbers, indicating that 24 subjects would need to be tested and results negative for minor allergenic proteins in order to achieve 95% confidence that the negative finding was true (FAO/WHO, 2001). The FAO/WHO panel did recognize that in many cases a sufficient number of subjects are unlikely to be identified to allow testing with 24 allergic subjects. They also recognized that in most cases it is better to use a few clearly allergic subjects rather than a larger number of marginally allergic subjects and therefore recommended subjects be selected who have allergen-specific IgE scores of greater than 10 kU/L. While the value is somewhat arbitrary, the point is to select test subjects who have strong positive IgE binding to the relevant control (source or sequence matched allergen) to improve the chance of identifying significant IgE binding to the GM test protein.

A serum volume sufficient to perform multiple tests is extremely useful for regulatory studies (e.g. 2 and 5 mL per donor). Direct IgE binding studies with immunoblotting and ELISA will usually require at least 0.2 to 0.5 mL per complete assay. Inhibition tests will often be needed and there may be a need to perform repeated testing, increasing the needed sample volume (Hoff et al., 2007).

In almost every country there is a legal requirement to provide written information to potential serum donors so that they understand the purpose of the study that their serum will be used in and any risks associated with serum collection or testing to confirm allergies by skin prick tests or food challenges. The patient or legal guardian (e.g. parent), must provide written consent that is clearly tied directly to information about the purpose of the study and any potential risks and benefits to the individual donor. The study design and information forms must be in the native language of the serum donor and are usually reviewed by an unbiased committee to ensure patient safety prior to recruitment of subjects.

7. Experimental IgE Binding Methods

A variety of test methods may be used to measure antigen-specific IgE. Depending on the target, the availability of appropriate sera and the characteristics of the protein one or more of the methods are likely to work better than others. The most basic test methods may be considered to be derivations of immunoblotting or immunospotting. However, most of the time immunoblots are performed after the protein target is separated from other potential targets by electrophoresis then fixed onto either a nitrocellulose or nylon membrane followed by incubation with allergic donor

serum and immunodetection of bound IgE using some form of a labeled specific anti-IgE. In immunospotting techniques the proteins are not separated but are adsorbed to a surface followed by IgE binding and detection in a similar fashion. Variations include radioallergosorbent test (RAST), enzyme linked immunosorbent assay (ELISA) or dot blot with the protein or extract of proteins bound to a solid phase without separation, followed by incubation with allergic serum, then detection by an anti-IgE by detecting radioactive decay, enzymatically driven color development or light emission. Other variations of immunoassay are possible, but all depend on the high affinity interaction of IgE from allergic sera with the specific protein.

7.1. Immunoblots

Proteins in extracts of GM crops or most foods may be separated under an electric field in a polyacrylamide gel (PAGE) with the strong denaturing detergent to disrupt hydrophobic interactions, sodium dodecyl-sulfate (SDS) and a reducing agent which disrupts disulfide bonds (dithiothreitol, DTT or 2-mercaptoethanol, β -ME) and tested for IgE binding (Goodman and Leach, 2004). The protein migration distance is inversely proportional to their apparent molecular mass. The advantage of this system is that the protein can be separated from others in an extract and that the added information of apparent mass is identified along with potential IgE-binding. However, sometimes discontinuous or conformational epitopes are not adequately represented for binding because the proteins were denatured (SDS) and or reduced (DTT or mercaptoethanol), especially if disulfide bonds are important for secondary epitopes. Non-reducing conditions, or even native conditions may be used to enhance the probability of finding conformational epitopes although the proteins may not separate as well. An alternative method is to separate the proteins based on surface exposed acidic or basic amino acids using isoelectric focusing (IEF). The separated proteins are transferred, usually in an electric field, onto either nitrocellulose, polyvinylidene difluoride (PVDF) or nylon membrane. After blotting some proteins refold to native conformation in buffer and immunoblotting may detect sequential and even discontinuous epitopes in some cases. Two dimensional electrophoresis using IEF in the first dimension followed by reducing SDS-PAGE in the second dimension, then blotting has become a popular way to efficiently separate individual proteins and test antibody binding. Again some proteins may refold properly following transfer and present discontinuous epitopes, others will not. The membranes are blocked with non-specific protein and non-ionic detergent to inhibit non-specific antibody binding (Goodman and Leach, 2004). Then the membrane is incubated with diluted human serum to allow IgE binding. After washing unbound serum antibodies the membrane is incubated with diluted polyclonal or monoclonal anti-human IgE antibodies that have been labeled in some way to allow detection of bound IgE. The label may be a radioactive nuclide such as ^{125}I , or an enzyme, either alkaline phosphatase or horseradish peroxidase, which is then detected as a visibly stained band if colorimetric substrates are used or by image capture if electrochemiluminescent (ECL) substrates are used. As in all immunoassays, the sensitivity and specificity of each assay must be tested using appropriate controls. Controls should include positive and negative allergen controls for the test subjects, an extract of non-GM and the GM crop, a no-serum control to test secondary detection antibody stickiness and non-allergic serum. The controls are necessary to establish a no-detection threshold. Additionally it is useful to apply the pure GM protein or GM extract at two or more concentrations to demonstrate a dose response. It is also quite important if at all possible to verify GM sample loading and specificity with animal IgG antibodies to the purified GM protein. If apparent IgE binding is demonstrated, the specificity of

binding may be verified by inhibition assay where the sera is pre-incubated with the recombinant protein, or with irrelevant protein prior to incubating with fresh western blots. If specific, the bands should be significantly reduced by the appropriate inhibitor, while if the binding is not specific, the binding should not be significantly reduced or it should be reduced when irrelevant proteins are used as inhibitors (Goodman and Leach, 2004).

7.2. Immunospots

There are advantages and disadvantages to the various methods described below but they are all based on the same principle, binding the target protein (usually native conformation) to a solid substrate, followed by blocking, incubation with serum, removal of unbound IgE, incubation with a highly specific labeled anti-IgE and then detection. The experimental design for the comparison is critical to demonstrate IgE binding to the test protein relative to controls, or between allergic and non-allergic subjects. Regarding assay methods, the primary difference between the techniques is the level of sensitivity, which is primarily related to the amount of bound antigen target. However, there is a complex interaction of abundance and affinity that influences both sensitivity and specificity. In all cases secondary assay testing may be required using various levels of inhibitors to demonstrate specificity of IgE binding. It is also important to demonstrate that the protein target is deposited on the solid phase as expected and if possible measure the quantity by independent means such as antigen-specific IgG measurement. Not all proteins bind equivalently to the various substrates that may be used in these assays.

The ancillary equipment needed to perform the specific assays can make a tremendous difference in cost and complexity, but can also add to the accuracy of measurement. The cost of production or purification of proteins may also be significant. In the simplest format, the dot-blot, the protein or extract is spotted onto nitrocellulose or PVDF membrane, and then air-dried to fix it to the substrate (Chew et al., 1999). After incubating with sera the IgE may be detected by colorimetric reporter system which means the reading can be visual, without instruments. However quantification would require spots of various concentrations and the range of error is likely to be great. Alternatively an image analyzer or densitometer may be used to estimate binding intensity. A more modern version of the immunospot is the microarray system where hundreds or thousands of protein spots are sprayed onto a slide in carefully controlled grid-positions before applying test sera (Jahn-Schmid et al., 2003). The target is usually spotted in duplicate or triplicate to ensure reproducibility. The advantage of microarrays is the small amount of sera that may be used to test binding to many proteins at once. The anti-IgE is usually labeled with a fluorescent marker. If assay conditions are optimized for each allergen the assays may be quite specific and sensitive. The disadvantage is that the antigen application and bound IgE readout require special equipment, increasing costs.

The most commonly used ELISA in allergy testing is the direct IgE binding assay. The protein is bound to a plastic plate, either through hydrophobic interactions between the protein and plastic, or through polar interactions if a charge modified plastic is used. The binding capacity per unit flat surface area is low compared to the binding capacity of an equivalent flat surface on a membrane or paper or synthetic disc. The dynamic range is probably less in ELISA than dot blots or other three dimensional substrates (Wahyuni et al., 2003). However, in many ways the ELISA offers a quantitative measurement and relatively high through-put that is not possible with some other formats. Non-specific binding can be a serious problem that varies between subjects and test proteins and the plate must be blocked with an appropriate blocker, usually gelatin, bovine serum albumin or non-fat dry milk (NFDm). After

incubating serum in blocking solution and then on the plate to allow binding, three or four washing steps are used to reduce non-specific binding. The most direct detection system usually has fewer background binding problems. It is important to use a highly specific anti-IgE that is directly coupled to horseradish peroxidase, or alkaline phosphatase. Some sera will have high titers of antigen-specific IgG as well as IgE. The ratio of binding specificity of anti-IgE to IgE vs. IgG should be at least 500 fold as can be tested by directly binding various concentrations of highly purified IgE and IgG directly to the ELISA plate, then incubating with the detection system. Coating of target protein and detection reagent dilutions must be optimized as well as the time and conditions of binding (temperature, salt concentration, non-ionic detergent). As a rule of thumb, the total amount of a purified protein target that often works well is between 50 and 100 ng of specific protein per individual well (96 well plate). If a whole extract is used, it will depend on the concentration of the target in the extract, but usually loading a total of 1 to 20 micrograms of protein extract per well provides a very workable target density. If the sensitivity of the assay is not sufficient, it may be enhanced by using an amplification step, with biotinylated anti-IgE, then a streptavidin, or Neutravidin-labeled with HRP to enhance measurement. However some protein preparations contain biotin and false positive signals may occur, which must be evaluated with a no-serum control. Additionally, NFDm, while usually providing superior blocking, may inhibit a biotin-avidin interaction as most preparations of NFDm contain high levels of biotin. Individual sera from those with specific allergies, irrelevant allergies and no-known allergies, as well as no serum controls should be used to objectively judge specific binding to the target protein. Serum samples are usually diluted between 1:5 (v:v) and 1:20 using blocking solution for diluent to provide a good positive signal to noise ratio. In general more dilute serum samples provide greater specificity as only high affinity binding is likely. However, the sensitivity is reduced as the serum is diluted. After unbound serum is washed away labeled anti-IgE is added and bound and free anti-IgE is removed, then colorimetric substrate is added. Tetra-methylbenzidine (TMB) is a substrate that is used with HRP. Color development must be monitored and should be stopped with an appropriate reagent (e.g. 6N sulfuric acid for HRP-TMB) at a fixed time, or a standard state of development in a positive well in order to optimize the signal to noise ratio.

While a threshold of IgE detection must be set for positive and specific-IgE binding in the predictive assays for the assessment, the threshold cannot be directly tied to a biological response as could be done if the protein was a demonstrated allergen (e.g. Ara h 2 of peanut, Bet v 1 of birch pollen). Instead a statistical assay limit is typically used to indicate an ELISA result from an allergic subject is scored as a positive response. For instance an optical density (OD) reading between 2 and 5 times the average OD of negative control subjects or an OD more than three standard deviations above the non-allergic controls is often used as a positive threshold (Goodman and Leach, 2004). However, those are statistical definitions of positive analytical results, not a measure of an allergic biological response. Another option is to set the limit based on a measure of specific IgE equivalent to a specific amount of IgE standard fixed on the ELISA plate, calibrated to the World Health Organization standard which is defined as 1 kU/L, equal to 2.4 mg of IgE per liter (Plebani et al., 1998). As an example, the original Pharmacia Diagnostics ImmunoCAP was calibrated to 0.35 kU/L as a minimum level of detection. It is extremely important to understand that there is no absolute level of specific IgE above which an individual will have allergic reactions and below which there is no risk of reaction. The serum level measured for individuals with true allergic responses will vary for each allergen and across individuals. Further, in vitro measurements cannot distinguish between IgE binding to a single epitope on a protein, which cannot lead to

mast cell or basophil degranulation, and binding to two or more epitopes, which may lead to degranulation and an allergic response *in vivo*. Complicating factors to consider are that *in vitro* tests measure blood levels of IgE, but allergic responses are usually caused by the presence of allergen at an epithelial surface (mucosal or epidermal) being bound to IgE which is already attached to high affinity FcεRI receptors on mast cells resident in the mucosa. The serum half-life of IgE is relatively short (< 3 days), where the half-life on mucosal mast cells is at least 14 days and there are B cells and plasma cells in the mucosal tissues that produce IgE which may not circulate but is available to re-load receptors on tissue mast cells (Gould et al., 2003).

It is common knowledge that for even the best antigen-specific *in vitro* IgE assays some individuals with clear allergic reactions to the material will test below the statistical limit of detection for the assay, and that some individuals who have absolutely no allergic response will test markedly above the detection limit. Few studies though include enough non-allergic controls to estimate the positive and negative predictive values for any one allergen. In one study a commercially available egg-specific *in vitro* IgE assay (CAPS) was compared to biological responses of double-blinded, placebo-controlled food challenges (Sampson and Ho, 1997). One-half of the individuals having a positive food challenge to egg had greater than 9.8 kU/L IgE to egg by CAPS. The average egg-specific CAPS value for those who were originally suspected of having egg allergy but were DBPCFC negative was 1.0 kU/L. The 95% negative predictive value for the egg-specific CAPS was determined to be 0.6 kU/L in this study (Sampson and Ho, 1997). Yet the background value determined by the CAPS manufacturer is 0.35 kU/L, demonstrating that individuals who are not symptomatic, i.e., not allergic, can have positive IgE binding even in well validated assays (Sampson and Ho, 1997). Such studies demonstrate the complexity of setting biologically meaningful thresholds in the safety assessment. The implication is that assays must include appropriate positive controls for the test material (Hoff et al., 2007). For instance if a sequence of a GM protein scored a high identity compared to a major egg white protein, the egg white protein should be included as a control. If the serum IgE to the egg white protein was 100 kU/L, but only 0.6 kU/L to the GM protein, the biological response to the GM protein is not likely to be significant even though it may be above the average negative control serum. Alternatively, if the IgE binding to the egg white protein was 10 kU/L and the binding to the GM protein was 2 kU/L it is much more likely that the GM protein could elicit an allergic response in the individual.

The principles and practical limitations for other assay formats include RAST, EAST and microarrays are similar to ELISA. The protein targets in a RAST assay is usually covalently bound to either a synthetic polymer or cellulose following cyanogens bromide activation of the polymer. Serum IgE binding, washing and detection antibody reagents are used, but the anti-IgE is labeled with ¹²⁵I. Bound anti-IgE is detected in a gamma counter or on radiographic film. EAST and microarrays are similar to RAST or ELISA except detection is a fluorescent system. Since the coupling methods are different between assays and involve different amino acid reactants, some proteins may be detected more specifically in one format compared to another. However, protein to protein variability is hard to judge as almost all comparative tests have been done with different proteins or extracts as is the case for each of the commercial *in vitro* IgE systems (e.g. Plebani et al., 1998). All methods generally work for proteins, but one cannot directly compare results across different assay formats.

7.3. Inhibition assays

Inhibition of IgE binding to target proteins is commonly used to demonstrate specificity of binding and also cross-reactivity. Inhibition

can be used with nearly all assay formats including immunoblots, dot blots, ELISA, RAST and EAST. Usually the solid-phased protein target quantity is held constant while soluble inhibitors, which may include extracts, pure specific proteins, sequence related proteins and irrelevant controls, are used at various concentrations to develop an inhibition curve in a quantitative assay (e.g. Scheurer et al., 1999). Inhibition of immunoblot binding is often performed with a fixed concentration of both solid-phased target and inhibitors to demonstrate qualitative or semi-quantitative differences because of the high volume of sera needed for gel blots (e.g. Goodman and Leach, 2004; Koppelman et al., 2005).

7.4. Examples of serum testing with GM products

Although a serum IgE study was not warranted based on bioinformatics comparison, sera from house dust-mite (HDM) allergic subjects were tested for IgE binding to the insecticidal GM protein Cry 1 F in maize, and compared to IgE binding to the “sequence matched” Der p 7 HDM allergen in a controlled study (Ladics et al., 2006). The bioinformatics alignment that led to these tests was merely an identical six amino acid match between Cry 1 F and the house dust mite allergen *Dermatophagoides pteronyssinus*, Der p 7 (Fig. 1, Ladics et al., 2006). For characterization, the authors demonstrated Cry 1 F specific IgG mouse monoclonal antibody binding to an appropriate size protein in extract of GM maize, but not non-GM maize, and no binding to a high concentration of Der p 7 (Fig. 2, Ladics et al., 2006). Sera from 20 individuals with demonstrated allergy HDM, including 10 that had clear specific IgE to purified Der p 7 were used in the tests. The results were negative in that none of the sera had IgE that bound to a high concentration extract of GM maize or to non-GM maize, even though half had demonstrable IgE binding to Der p 7 in the assays. These results indicated that there was no evidence of potential cross-reactivity of Cry 1 F to Der p 7 in this population of HDM allergic subjects (Ladics et al., 2006). It should therefore be assumed that the risk that HDM allergic individuals might react to Cry 1 F in the GM corn is extremely small.

As noted earlier, a study on GM allergenicity evaluated IgE binding and skin prick test reactivity to a 2S albumin protein encoded by a Brazil nut gene that was transferred into soybeans was tested in 1996. The pure protein bound IgE as did a protein of the same molecular weight from Brazil nut extract and from an extract of GM soybean, but not in an extract of non-GM soybean (Nordlee et al., 1996). The IgE binding was confirmed by RAST-Inhibition and the biological relevance of the IgE was demonstrated by skin prick test using the Brazil nut allergic subjects (Nordlee et al., 1996). As a consequence, the potential product was cancelled by the developer.

7.5. Illustrative model studies related to determination of potential cross reactivity of non-GM proteins

Recently the potential cross-reactivity of proteins in green (*Actinidia deliciosa*), gold (*Actinidia chinensis*) and baby kiwi (*Actinidia arguta*) fruits and a processed product of baby kiwi were evaluated by immunoblotting and ELISA using sera from donors with clear allergy to green and/or gold kiwi (Chen et al., 2006). There were clear differences in detectable IgE binding using the same serum samples to bind proteins in non-reduced, but denatured immunoblotted samples, reduced immunoblotted samples and native extract samples in ELISA. Two of the allergic subjects only demonstrated clear IgE binding to baby kiwi by ELISA and non-reduced immunoblots, but not by reduced immunoblotting. This result indicates their IgE probably binds to conformational (discontinuous) epitopes that do not re-form during incubation. Some individual sera with remarkable binding to green kiwi proteins by ELISA

had barely detectable IgE binding by Immunoblotting (Chen et al., 2006). These results demonstrate that at least for some allergens and patients two or more IgE binding assay techniques may be necessary to demonstrate important IgE binding. In addition, apparent specific IgE binding was noted to specific protein bands when sera from two non-kiwi-allergic subjects were tested by immunoblotting extracts of the fruit, demonstrating that false-positive IgE binding to some proteins is relatively common.

Other examples may be found in studies of allergenic legumes and tree nuts. Vicilins are common seed storage proteins in legumes and a number of tree nuts. A small number of peanut allergic individuals are reported to have clinically important allergic reactions when consuming specific other legumes and in some cases it appears to be cross-reactivity due to structural similarity of major seed storage proteins across species (Beardslee et al., 2000; Moneret-Vautrin et al., 1999). In rare instances cross-reactivity due to IgE that specifically binds to peanut allergens may even occur when consuming tree nuts (de Leon et al., 2005). Wensing et al. (2003) studied three patients who experience anaphylaxis after eating green pea and also report having some allergic symptoms when eating peanut. Western blotting of reduced, denatured proteins from peanut (extracts and Ara h 1), and extract and partially purified vicilin of pea, as well as reciprocal inhibition assays demonstrated cross-reactive binding of patient IgE with higher binding avidity to pea. Inhibition ELISA assays demonstrated more than 100 fold higher IgE binding to pea vicilin compared to Ara h 1 for all three patients. The Wensing et al. (2003) in vitro IgE binding data corresponds to the more mild symptoms these individuals experience when consuming peanut, compared to pea and provide a good example of how carefully controlled studies can help identify potentially important cross-reactivity.

In the case of some legume proteins we know the sequence identity of proteins that are cross-reactive and can begin to evaluate thresholds of sequence matching that might be used to guide regulatory decisions about when to test for potential cross-reactivity. For example, the pea vicilin tested by Wensing et al. (2003) is 92% identical to vicilin in lentil (Len c 1) by FASTA, 55% identical to a soybean glycinin and 53% identical to Ara h 1 of peanut over the 420 amino acid alignment. Based on an 80 amino acid alignment performed with the search routine on <http://www.allergenonline.com/>, the highest identity for any 80 aa segment is 96% to Len c 1, 76% to glycinin and 66% identical to a segment of Ara h 1. Interestingly, serum tests by Wensing et al. (2003) demonstrated IgE binding of pea allergic subjects was cross-reactive to Ara h 1, but only about 1/100th as potent as the pea vicilin based on inhibition tests. The clinical history data for the three patients in that study indicated that at least one of them was also allergic to lentil, and a number of other legumes and one was allergic to at least three beans. There was no data on other sources of allergic reactions for the third subject. Since the symptoms reported for the three were less severe to peanut than to pea, and the IgE binding efficiency was less than 1/100th, one might conclude that an overall identity of ~ 50%, or an 80 aa identity match of ~ 66% is likely to indicate low level cross-reactivity.

Another case where cross-reactivity may be responsible at least for initiation of allergic responses to a legume protein is the soybean pathogenesis related (PR) protein that was originally called SAM22 and is now known as Gly m 4 (Mittag et al., 2004). The protein is a homologue of the Bet v 1 family of proteins which are highly conserved across plant families. It is also generally susceptible to heat denaturation and indeed, many of the individuals with significant IgE binding to the food related Bet v 1 homologues have reduced or no allergic reactions when the food is heated, demonstrating the conformational nature of IgE binding. Searching the allergen database of the Food Allergy Research and Resource Program (<http://www.allergenonline.com/>) by FASTA using the de-

fault parameters of the search page demonstrates that overall best full-length sequence identity of Gly m 4 is to bean PR protein at approximately 80% identity and the highest identity score for an 80 amino acid alignment to that protein is 89% (unpublished). Matches of nearly the same identity were found to lupine and peanut PR proteins. The highest match to a non-legume PR protein was to the hazelnut allergen Cor a 1.04 at 58% overall identity and a maximum match of 61% to the best 80 amino acid alignment. The highest Bet v 1 match was slightly lower. The vegetable Bet v 1 homologue, Dau c 1 matched with 39% identity overall and 41% as the best 80 aa segment score. Since most people are exposed to homologous PR proteins from many different airway and food sources, identifying the exact sensitizing source for the individual is impossible. Further, merely performing in vitro IgE binding tests and even inhibition assays with these homologous proteins does not prove that the specific proteins cause allergic reactions for the individual subject. It is essential to take a very careful allergic history of the subject and to perform careful quantitative IgE cross-inhibition studies with the suspected cross-reactive proteins and less identical related proteins if one is to fully understand cross-reactivity. An example of such a study is the work reporting cross-reactivity of cherry and celery Bet v 1 homologues, in which homologues of roughly 40% shared identity appear to have some level of clinical cross-reactivity from proteins that by in vitro assays are roughly 100 fold different in binding avidities (Scheurer et al., 1999). However, understanding the relationships when individuals are exposed to homologues of many sources is more complex than simple sequence identities or in vitro results might suggest as demonstrated by more complex in vitro assays (Mittag et al., 2006). The complex sequence relationships of this highly ubiquitous family of proteins discussed above may contribute to the observed diverse patterns of clinical and laboratory cross-reactivities. Based on the additional information from studies like those of Scheurer et al. (1999) and Mittag et al. (2006), when evaluated with the sequence comparison data suggest that the sequence identity criteria used to flag proteins that may represent a risk of clinical cross-reactivity could be raised above the 35% identity match over any segment of 80 amino acids that has been recommended by FAO/WHO (2001) and Codex (2003). Additionally taking into account the small number of structural protein families that are thought to cause most allergic reactions (Jenkins et al., 2005), low level identity matches to proteins that do not fall into those families are probably less likely to represent an important risk of clinical cross-reactivity based on known incidences of allergy.

8. Summary

Evaluating the potential allergenicity of proteins encoded by genes transferred into GM organisms may require protein-specific IgE testing with serum from individuals allergic to the source of the gene, if the source is a significant allergen, or when the sequence of the protein has a high identity score compared to a known allergenic protein (e.g., greater than 50% identical over the full-length, or conservatively, greater than 35% identical over any 80 aa segment). Under those circumstances the serum donors must be carefully chosen to include individuals who are clinically proven to be allergic to the source or the source of the potentially cross-reactive allergen. Further, tests materials must be well characterized and the assay demonstrated to be sensitive and specific enough to demonstrate specific binding. Under most circumstances the assay design must be capable of identifying sequential and conformational epitopes. If the protein is glycosylated in the GM organism, the tests should evaluate IgE binding to determine whether the carbohydrate may serve as a target for binding. If it is, further

evaluation is warranted to evaluate the biological or clinical relevance of binding (e.g. basophil histamine release or skin prick tests). Absolute thresholds of IgE binding have not been, and will likely never be established which can identify allergenic risks with 100% certainty. Tests of potential cross-reactivity should therefore always include the relevant primary allergen. Although there is little direct evidence, proteins sharing only 0.1 to 1% avidity based on inhibition studies are unlikely to induce significant basophil histamine release, and probably unlikely to induce a significant clinical reactions compared to the primary allergen (e.g. Sheurer et al., 1999; de Leon et al., 2007). It must be recognized that many proteins appear to have exhibit some low level of in vitro IgE binding using sera of those without clinical allergy. As assays are rarely calibrated to indicate specific levels of antigen-specific IgE, appropriate positive and negative allergen and serum controls must be included in each assay to demonstrate the relative level of IgE binding.

Conflict of interest statement

The author is an occasional consultant to food, commodity and biotechnology companies to help evaluate potential allergenicity of novel proteins and products of biotechnology.

The author owns stock in Monsanto Company as a result of previous employment-retirement program.

The author has funding from the U.S. EPA Star Grant program for improving serum IgE test methods for the evaluation of allergenicity of genetically engineered crops.

The author manages the AllergenOnline database in the Food Allergy Research and Resource Program, Department of Food Science and Technology, University of Nebraska, Lincoln. The database is funded under contract by six agricultural biotechnology companies, and provides a peer review of sequences deemed to be allergens or putative allergens.

References

- Aalberse, R.C., 2000. Structural biology of allergens. *J. Allergy Clin. Immunol.* 106, 228–238.
- Ballmer-Weber, B.K., Wangorsch, A., Bohle, B., Kaul, S., Kundig, T., Fotisch, K., van Ree, R., Vieths, S., 2005. Component-resolved in vitro diagnosis in carrot allergy: does the use of recombinant carrot allergens improve the reliability of the diagnostic procedure? *Clin. Exp. Allergy* 35, 970–978.
- Ballmer-Weber, B.K., Wuthrich, B., Wangorsch, A., Fotisch, K., Altmann, F., Vieths, S., 2001. Carrot allergy: double-blinded, placebo-controlled food challenge and identification of allergens. *J. Allergy Clin. Immunol.* 108 (2), 301–307.
- Beardslee, T.A., Zeece, M.G., Sarath, G., Markwell, J.P., 2000. Soybean glycinin G1 acidic chain shares IgE epitopes with peanut allergen Ara h 3. *Int. Arch. Allergy Immunol.* 123 (4), 299–307.
- Burks, A.W., Cockrell, G., Connaughton, C., Guin, J., Allen, W., Helm, R., 1994. Identification of peanut agglutinin and soybean trypsin inhibitor as minor legume allergens. *Int. Arch. Allergy Immunol.* 105, 143–149.
- Chakraborty, S., Chakraborty, N., Datta, A., 2000. Increased nutritive value of transgenic potato by expressing a nonallergenic albumin gene from *Amaranthus hypochondriacus*. *Proc. Natl. Acad. Sci. USA* 97 (7), 3724–3729.
- Chen, L., Lucas, J.S., Hourihane, J.O., Lindemann, J., Taylor, S.L., Goodman, R.E., 2006. Evaluation of IgE binding to proteins of hardy (*Actinidia arguta*), gold (*Actinidia chinensis*) and green (*Actinidia deliciosa*) kiwifruits and processed hardy kiwifruit concentrate, using sera of individuals with food allergies to green kiwifruit. *Food Chem. Toxicol.* 44, 1100–1107.
- Chew, F.T., Yi, F.C., Chua, K.Y., Fernandez-Caldas, E., Arruda, L.K., Chapman, M.D., Lee, B.W., 1999. Allergenic differences between the domestic mites *Blomia tropicalis* and *Dermatophagoides pteronyssinus*. *Clin. Exp. Allergy* 29 (7), 982–988.
- Codex Alimentarius Commission, 2003. Proposed draft annex on the assessment of possible allergenicity of the draft guideline for the conduct of food safety assessment of foods derived from recombinant DNA plants. Joint FAO/WHO Food Standard Program, Appendix IV; pp. 57–60.
- de Leon, M.P., Drew, A.C., Glaspole, I.N., Suphioglu, C., Rolland, J.M., O'Hehir, R.E., 2005. Functional analysis of cross-reactive immunoglobulin E antibodies: peanut-specific immunoglobulin E sensitizes basophils to tree nut allergens. *Clin. Exp. Allergy* 35 (8), 1056–1064.
- de Leon, M.P., Drew, A.C., Glaspole, I.N., Suphioglu, C., O'Hehir, R.E., Rolland, J.M., 2007. IgE cross-reactivity between the major peanut allergen Ara h 2 and tree nut allergens. *Mol. Immunol.* 44, 463–471.
- Dolan, W.K., 2003. IgE antibody in the serum-detection and diagnostic significance. *Allergy* 58, 717–723.
- Elkayam, O., Tamir, R., Pick, A.I., Wysenbeek, A., 1995. Serum IgE concentrations, disease activity, and atopic disorders in systemic lupus erythematosus. *Allergy* 50 (1), 94–96.
- 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. Rome, Italy, pp. 22–25.
- Goodman, R.E., 2006. Practical and predictive bioinformatics methods for the identification of potentially cross-reactive protein matches. *Mol. Nutr. Food Res.* 50, 655–660.
- Goodman, R.E., Hefle, S.L., 2005. Gaining perspective on the allergenicity assessment of genetically modified food crops. *Expert Rev. Clin. Immunol.* 1 (4), 561–578.
- Goodman, R.E., Hefle, S.L., Taylor, S.L., van Ree, R., 2005. Assessing genetically modified crops to minimize the risk of increased food allergy: a review. *Int. Arch. Allergy Immunol.* 137, 153–166.
- Goodman, R.E., Leach, J.N., 2004. Assessing the allergenicity of proteins introduced into genetically modified crops using specific human IgE assays. *J. A.O.A.C. Intl.* 87 (6), 1423–1432.
- Gould, H.J., Sutton, B.J., Bevil, A.J., Bevil, R.L., McCloskey, N., Coker, H.A., Fear, D., Smurthwaite, L., 2003. The biology of IgE and the basis of allergic disease. *Annu. Rev. Immunol.* 21, 579–628.
- Hamada, M., Furusyo, N., Urabe, K., Morita, K., Nakahara, T., Kinukawa, N., Nose, Y., Hayashi, J., Furue, M., 2005. Prevalence of atopic dermatitis and serum IgE values in nursery school children in Ishigaki Island, Okinawa, Japan. *J. Dermatol.* 32 (4), 248–255.
- 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.
- Hoa, T.T.C., Al-Babili, S., Schaub, P., Potrykus, I., Beyer, P., 2003. Golden Indica and Japonica rice lines amenable to deregulation. *Plant Physiol.* 133, 161–169.
- Hoff, M., Son, D.-Y., Gubesch, M., Ahn, K., Lee, S.-L., Vieths, S., Goodman, R.E., Ballmer-Weber, B., Bannion, G.A., 2007. Serum testing of genetically modified soybeans with special emphasis on potential allergenicity of the heterologous protein CP4 EPSPS. *Mol. Nutr. Food Res.* 51, 946–965.
- Ikematsu, K., Tachimoto, H., Sugisaki, C., Syukuya, A., Ebisawa, M., 2006. Feature of food allergy developed during infancy (1) – relationship between atopic dermatitis and food allergy (in Japanese). *Arerugi* 55 (2), 140–150.
- Ivanciu, O., Schein, C.H., Braun, W., 2003. Database and computational tools for allergenic proteins. *Nucl. Acids Res.* 31, 359–362.
- James, C., 2003. Preview: Global Status of Commercialized Transgenic Crops: 2003. ISAAA Briefs, No. 30, Ithaca, ISAAA.
- Jahn-Schmidt, B., Harwanegg, C., Hiller, R., Bohle, B., Ebner, C., Scheiner, O., Mueller, M.W., 2003. Allergen microarray: comparison of microarray using recombinant allergens with conventional diagnostic methods to detect allergen-specific serum immunoglobulin E. *Clin. Exp. Allergy* 33 (10), 1443–1449.
- Jenkins, J.A., Griffiths-Jones, S., Shewry, P.R., Breiteneder, H., Mills, E.N.C., 2005. Structural relatedness of plant food allergens with specific reference to cross-reactive allergens. *J. Allergy Clin. Immunol.* 115, 163–170.
- König, A., Cockburn, A., Crevel, R.W.R., Debruyne, E., Graftstroem, R., Hammerling, U., Kimber, I., Knudsen, I., Kuiper, H.A., Peijnenburg, A.A.C.M., Penninks, A.H., Poulsen, M., Schaub, M., Wal, J.M., 2004. Assessment of the safety of foods derived from genetically modified (GM) crops. *Food Chem. Toxicol.* 42, 1047–1088.
- Koppelman, S.J., de Jong, A.H., Laaper-Ertmann, M., Peeters, K.A.B.M., Knulst, A.C., Hefle, S.L., Knol, E.F., 2005. Purification and immunoglobulin E-binding properties of peanut allergen Ara h 6: evidence for cross-reactivity with Ara h 2. *Clin. Exp. Allergy* 35, 490–497.
- 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 p7 protein with human sera positive for Der p7-IgE. *Regul. Toxicol. Pharmacol.* 44, 136–143.
- Laske, N., Niggemann, B., 2004. Does the severity of atopic dermatitis correlate with serum IgE levels? *Pediatr. Allergy Immunol.* 15, 86–88.
- Lauer, I., Foetisch, K., Kolarich, D., Ballmer-Weber, B.K., Conti, A., Altmann, F., Vieths, S., Scheurer, S., 2004. Hazelnut (*Corylus avellana*) vicilin Cor a 11: molecular characterization of a glycoprotein and its allergenic activity. *Biochem. J.* 383, 327–334.
- Linneberg, A., Husemoen, L.L.N., Nielsen, N.H., Madsen, F., Frolund, L., Johansen, N., 2006. Screening for allergic respiratory disease in the general population with the ADVIA Centaur Allergy Screen Assay. *Allergy* 61, 344–348.
- 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 derived from genetically engineered crop plants. *Crit. Rev. Food Sci. Nutr.* 36, S165–S186.
- Mittag, D., Vieths, S., Vogel, L., Becker, W.M., Rihs, H.P., Helbling, A., Wuthrich, B., Ballmer-Weber, B.K., 2004. Soybean allergy in patients allergic to birch pollen: clinical investigation and molecular characterization of allergens. *J. Allergy Clin. Immunol.* 113 (1), 148–154.
- Mittag, D., Batori, V., Neudecker, P., Wiche, R., Friis, E.P., Ballmer-Weber, B.K., Vieths, S., Roggen, E.L., 2006. A novel approach for investigation of specific and cross-reactive IgE epitopes on Bet v 1 and homologous food allergens in individual patients. *Mol. Immunol.* 43, 268–278.
- Moneret-Vautrin, D.A., Guerin, L., Kanny, G., Flabbee, J., Fremont, S., Morisset, M., 1999. Cross-allergenicity of peanut and lupine: the risk of lupine allergy in patients allergic to peanuts. *J. Allergy Clin. Immunol.* 104 (4 Pt. 1), 883–888.

- Morton, R.L., Schroeder, H.E., Bateman, K.S., Chrispeels, M.J., Armstrong, E., Higgins, T.J.V., 2000. Bean α -amylase inhibitor 1 in transgenic peas (*Pisum sativum*) provides complete protection from pea weevil (*Bruchus pisorum*) under field conditions. *Proc. Natl. Acad. Sci. USA* 97, 3820–3825.
- National Academy of Sciences, 2004. Safety of Genetically Engineered Foods: Approaches to Ab-testing Unintended Health Effects. National Academies Press, Washington.
- Nordlee, J.A., Taylor, S.L., Townsend, J.A., Thomas, L.A., Bush, R.K., 1996. Identification of a Brazil-nut allergen in transgenic soybeans. *New. Engl. J. Med.* 334, 688–692.
- Paine, J.A., Shipton, C.A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstone, A.L., Drake, R., 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nat. Biotechnol.* 23 (4), 482–487.
- Pasini, G., Simonato, B., Curioni, A., Vincenzi, S., Cristaudo, A., Santucci, B., Peruffo, A.D., Giannattasio, M., 2002. IgE-mediated allergy to corn: a 50 kDa protein, belonging to the reduced soluble proteins, is a major allergen. *Allergy* 57, 106–898.
- Pastorello, E.A., Pompei, C., Pravettoni, V., Farioli, L., Calamari, A.M., Scibilia, J., Robino, A.M., Conti, A., Iametti, S., Fortunato, D., Bonorni, S., Ortolani, C., 2003. Lipid-transfer protein is the major maize allergen maintaining IgE-binding activity after cooking at 100°C, as demonstrated in anaphylactic patients and patients with positive double-blind, placebo-controlled food challenge results. *J. Allergy Clin. Immunol.* 112, 775–783.
- Plebani, M., Bernardi, D., Basso, D., Borghesan, F., Faggian, D., 1998. Measurement of specific immunoglobulin E: intermethod comparison and standardization. *Clin. Chem.* 44 (9), 1974–1979.
- Prescott, V.E., Campbell, P.M., Moore, A., Mattes, J., Rothenberg, M.E., Foster, P.S., Higgins, T.J.V., Hogan, S.P., 2005. Transgenic expression of bean α -amylase inhibitor in peas results in altered structure and immunogenicity. *J. Agric. Food. Chem.* 53, 9023–9030.
- Purohit, A., Laffer, S., Metz-Favre, C., Verot, A., Kricsek, F., Valenta, R., Pauli, G., 2005. Poor association between allergen-specific serum immunoglobulin E levels, skin sensitivity and basophil degranulation: a study with recombinant birch pollen allergen Bet v 1 and an immunoglobulin E detection system measuring immunoglobulin E capable of binding to FC ϵ RI. *Clin. Exp. Allergy* 35, 186–192.
- Rascon-Cruz, Q., Sinagawa-Garcia, S., Osuna-Castro, J.A., Bohorova, N., Parades-Lopez, O., 2004. Accumulation, assembly, and digestibility of amarantin expressed in transgenic tropical maize. *Theor. Appl. Genet.* 108, 335–342.
- Sampson, H.A., 2004. Update on food allergy. *J. Allergy Clin. Immunol.* 113, 805–819.
- Sampson, H.A., Ho, D.G., 1997. Relationship between food-specific IgE concentrations and the risk of positive food challenges in children and adolescents. *J. Allergy Clin. Immunol.* 100, 444–451.
- Scheurer, S., Son, D.Y., Boehm, M., Karamloo, F., Franke, S., Hoffman, A., Haustein, D., Vieths, S., 1999. Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. *Mol. Immunol.* 36, 155–167.
- Schroeder, H.E., Gollasch, S., Moore, A.E., Tabe, L.M., Craig, S., Hardie, D.C., Chrispeels, M.J., Spencer, D., Higgins, T.J.V., 1995. Bean α -amylase inhibitor confers resistance to the pea weevil (*Bruchus pisorum*) in transgenic peas (*Pisum sativum* L.). *Plant Physiol.* 107, 1233–1239.
- Scurlock, A.M., Burks, A.W., 2004. Peanut allergenicity. *Ann. Allergy Asthma Immunol.* 93 (Suppl. 3), S12–S18.
- Shibasaki, M., Sumazaki, R., Isoyama, S., Takita, H., 1992. Interaction of lectins with human IgE: IgE-binding property and histamine releasing activity of 12 plant lectins. *Int. Arch. Allergy Immunol.* 98, 18–25.
- 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, 252–258.
- Sinagawa-Garcia, S.R., Rascon-Cruz, Q., Valdez-Ortiz, A., Medina-Godoy, S., Escobar-Guierrez, A., Paredes-Lopez, O., 2004. Safety assessment by in vitro digestibility and allergenicity of genetically modified maize with an Amaranth 11S globulin. *J. Agric. Food Chem.* 52, 2709–2714.
- Stadler, M.B., Stadler, B.M., 2003. Allergenicity prediction by protein sequence. *FASEB J.* 17, 1141–1143.
- van Ree, R., 2002. Carbohydrate epitopes and their relevance for the diagnosis and treatment of allergic diseases. *Int. Arch. Allergy Immunol.* 129 (3), 189–197.
- van Ree, R., Cabanes-Macheteau, M., Akkerdaas, J., Milazzo, J.-P., Loutelier-Bourhis, C., Rayon, C., Villalba, M., Koppelman, S., Aalberse, R., Rodriguez, R., Faye, L., Lerouge, P., 2000. β (1,2)-xylose and α (1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J. Biol. Chem.* 275 (15), 11451–11458.
- Wahyuni, S., van Ree, R., Mangali, A., Supali, T., Yazdanbakhsh, M., Sartono, E., 2003. Comparison of an enzyme linked immunosorbent assay (ELISA) and a radioallergosorbent test (RAST) for detection of IgE antibodies to *Brugia malayi*. *Parasite Immunol.* 25, 609–614.
- Wensing, M., Knulst, A.C., Piersma, S., O'Kane, F., Knol, E.F., Koppelman, S.J., 2003. Patients with anaphylaxis to pea can have allergy caused by cross-reactive IgE to vicilin (Ara h 1). *J. Allergy Clin. Immunol.* 111, 420–424.
- Young, N.M., Thibault, P., Watson, D.C., Chrispeels, M.J., 1999. Post-translational processing of two α -amylase inhibitors and an arcelin from the common bean, *Phaseolus vulgaris*. *FEBS Lett.* 446, 203–206.
- Zavazal, V., Krauz, V., 1985. Lectin-binding ability of immunoglobulin E and its participation in triggering of mast cells. *Folia Microbiol. (Praha)* 30 (3), 237–246.