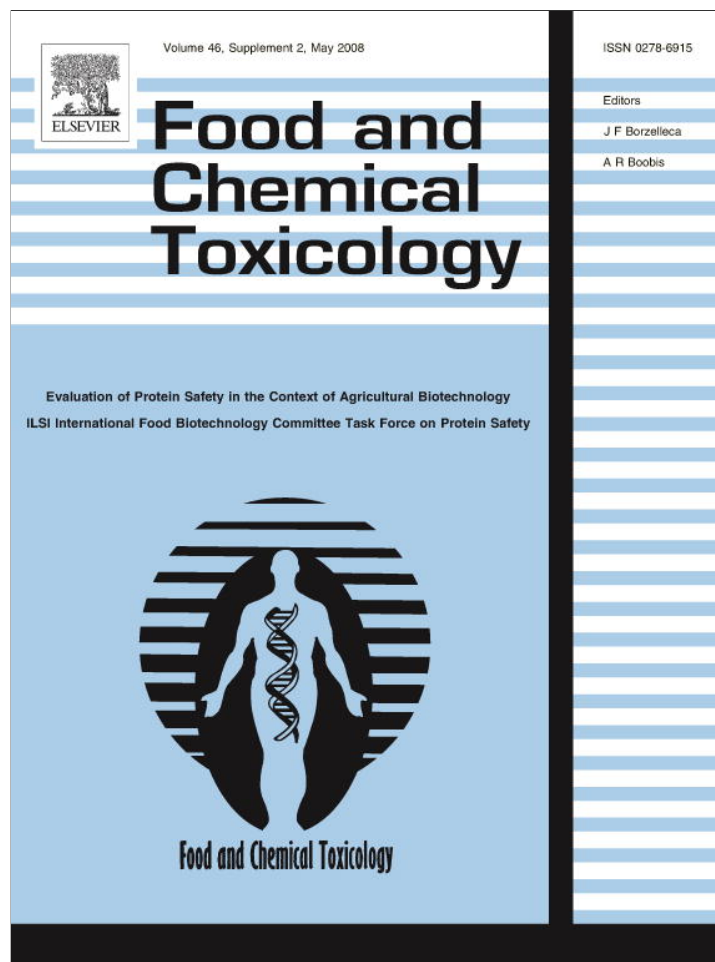


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



## Evaluation of protein safety in the context of agricultural biotechnology

Bryan Delaney<sup>a</sup>, James D. Astwood<sup>b,1</sup>, Helen Cunny<sup>c,2</sup>, Robin Eichen Conn<sup>d</sup>, Corinne Herouet-Guicheney<sup>e</sup>, Susan MacIntosh<sup>f</sup>, Linda S. Meyer<sup>g</sup>, Laura Privalle<sup>h</sup>, Yong Gao<sup>i,3</sup>, Joel Mattsson<sup>i</sup>, Marci Levine<sup>j,\*</sup>, ILSI International Food Biotechnology Committee Task Force on Protein Safety

<sup>a</sup> Pioneer Hi-Bred, A DuPont Company, Johnston, IA, USA

<sup>b</sup> Monsanto Company, St. Louis, MO, USA

<sup>c</sup> Bayer CropScience LP, Research Triangle Park, NC, USA

<sup>d</sup> Cargill Inc., Wayzata, MN, USA

<sup>e</sup> Bayer CropScience, Sophia-Antipolis, France

<sup>f</sup> MacIntosh & Associates, St. Paul, MN, USA

<sup>g</sup> Syngenta Biotechnology, Inc., Research Triangle Park, NC, USA

<sup>h</sup> BASF, Research Triangle Park, NC, USA

<sup>i</sup> Dow AgroSciences, Indianapolis, IN, USA

<sup>j</sup> International Life Sciences Institute, Washington, DC, USA

Received 3 July 2007; accepted 19 January 2008

### Abstract

One component of the safety assessment of agricultural products produced through biotechnology is evaluation of the safety of newly expressed proteins. The ILSI International Food Biotechnology Committee has developed a scientifically based two-tiered, weight-of-evidence strategy to assess the safety of novel proteins used in the context of agricultural biotechnology. Recommendations draw upon knowledge of the biological and chemical characteristics of proteins and testing methods for evaluating potential intrinsic hazards of chemicals. Tier I (potential hazard identification) includes an assessment of the biological function or mode of action and intended application of the protein, history of safe use, comparison of the amino acid sequence of the protein to other proteins, as well as the biochemical and physico-chemical properties of the proteins. Studies outlined in Tier II (hazard characterization) are conducted when the results from Tier I are not sufficient to allow a determination of safety (reasonable certainty of no harm) on a case-by-case basis. These studies may include acute and repeated dose toxicology studies and hypothesis-based testing. The application of these guidelines is presented using examples of transgenic proteins applied for agricultural input and output traits in genetically modified crops along with recommendations for future research considerations related to protein safety assessment.

© 2008 ILSI. Published by Elsevier Ltd. All rights reserved.

**Keywords:** Novel protein; Biotechnology; GM; Safety assessment; Hazard; Case study

\* Corresponding author. Tel.: +1 202 659 0074; fax: +1 202 659 3859.

E-mail address: [mlevine@ilsi.org](mailto:mlevine@ilsi.org) (M. Levine).

<sup>1</sup> Present address: ConAgra Foods, Omaha, NE, USA.

<sup>2</sup> Present address: National Toxicology Program, NIEHS, Research Triangle Park, NC, USA.

<sup>3</sup> Present address: Monsanto Company, St. Louis, MO, USA.

## 1. Summary

An important component of the safety assessment of agricultural products produced through biotechnology by expression of transgenic proteins is evaluation of the safety of newly expressed proteins. Because proteins introduced into crops via genetic modification using recombinant DNA techniques may not have been components of foods or feeds previously consumed, the safety of these proteins for humans or animals may not be known. The ILSI International Food Biotechnology Committee collaborated with experts on protein and food safety to develop a scientifically based approach to assess the safety of candidate novel proteins (CNPs) by drawing upon knowledge of the biological and chemical characteristics of proteins, available testing methods for evaluating intrinsic hazards of chemicals, and illustrative case studies of protein safety assessment. The scope is limited to food and feed safety, therefore assessment of crop performance and crop safety, including whole food animal feeding studies, pleiotropic effects, environmental aspects, or ethical considerations are not discussed. The potential for allergenicity is only briefly discussed since other publications have recently evaluated this area in significant detail (Hileman et al., 2002; Codex, 2003; Goodman et al., 2005; Ladics et al., 2007; Thomas et al., 2007).

Proteins are a necessary component of the diet of humans and other mammals. The mammalian digestive system degrades dietary proteins into constituent amino acids which are efficiently absorbed and reincorporated into new proteins (Day, 1996; US EPA, 2000). Because of the sensitivity of dietary proteins to digestion and the minimal potential for absorption of intact proteins from the GI system, the overwhelming majority of dietary proteins possess no potential for systemic toxicity (Sjoblad et al., 1992). Accordingly, consumption of proteins, as a general class of macronutrients is not normally associated with adverse effects. However, a limited number of proteins have demonstrated toxicity to humans and other mammals. Some animals (e.g., scorpions, snakes) produce venoms containing protein toxins that are only active parenterally (Magalhães et al., 1998; Sidell et al., 1997). Other toxic proteins are produced by pathogenic bacteria and there are some well-known plant-produced protein toxins such as the ribosome inhibiting protein ricin (Franz and Jax, 1997). Certain other proteins including lectins and enzyme inhibitors that are components of plants are considered antinutrients because, while they are not particularly toxic, repeated exposure to them can result in decreased utilization of dietary nutrients.

A two-tiered, weight-of-evidence strategy has been developed to evaluate the safety of transgenic proteins rather than a decision tree approach. This is a more flexible approach and it takes into account the totality of the data in a holistic manner. Ideally, the predictive value of each piece of evidence should be well understood in order to give certain data more 'weight' than others during the assessment, thus leading to more confidence in the overall assessment.

Tier I (potential hazard identification) includes an assessment of the biological function or mode of action and intended application of the protein, assessment of the history of safe use (HOSU) of a particular protein (Constable et al., 2007), a comparison of the amino acid sequence of the protein to other known proteins, particularly, those known to be toxic or allergenic and those considered to be antinutrients, and an evaluation of certain physical properties of the transgenic protein.

Studies outlined in Tier II (hazard characterization) are conducted when the results from Tier I are not sufficient to allow a determination of safety (reasonable certainty of no harm; Codex, 2003) and are conducted on a case-by-case basis. These studies may include acute toxicology studies. Acute toxicology studies are often conducted via oral exposure because that is the most likely route of exposure to the transgenic protein, and mice are used because less test substance is required. However, some studies using other routes of exposure, such as intraperitoneal or intravenous administration, have been conducted. Depending on the results from the Tier I assessment, additional toxicology studies and hypothesis-based testing could be considered on a case-by-case basis.

To demonstrate the application of the safety assessment framework to proteins introduced into GM crops, this document presents case studies on six transgenic proteins: (i) potato virus Y coat protein (viral resistance in potatoes); (ii) CP4 5-enolpyruvylshikimate-3-phosphate synthase (CP4 EPSPS; provides glyphosate tolerance); (iii) phosphinothricin acetyltransferase (PAT; provides glufosinate tolerance); (iv) phosphomannose isomerase (selectable marker that provides tolerance to mannose); (v) Cry1Ab protein, a delta endotoxin from *Bacillus thuringiensis* that provides protection from European corn borer; and (vi) a plant-derived antifungal protein (AFP). Applying the tiered approach, five of the case study proteins demonstrated reasonable certainty of no harm. One protein, the antifungal protein (AFP) was found to require more detailed hypothesis-based testing.

While the results demonstrated that the tiered-approach, as described in this document, is effective in assessing the safety of transgenic proteins used in GM crops, recommendations for future research considerations related to protein safety assessment are included.

## 2. Introduction

A number of field crops, including corn, soybeans, cotton, canola, potato, tomato, and squash have been genetically modified by recombinant DNA technology to express exogenous genes and have been marketed after extensive evaluation of food, feed and environmental safety (ISAAA, 2006). Expression of the introduced gene leads to one or more proteins being produced in the plant (*in planta*) that endow these crops with targeted specific traits. In many cases, the transgenic proteins (protein produced in the genetically modified crop) will be present in the food or

feed obtained from these crops which will be referred to in this documents as genetically modified (GM) crops.

As a macronutrient, protein is an essential component of the human diet and, although individual proteins mediate myriad biological functions, consumption of proteins as a class of dietary substances is not inherently associated with adverse effects (FAO/WHO, 1996). However, a limited number of protein toxins are known to exist. In light of this, and because some proteins introduced into crops through genetic modification using recombinant DNA techniques may not have been components of foods previously consumed by humans (i.e., may lack a presumed history of safe use), it is appropriate to evaluate the safety of newly introduced transgenic proteins as part of the overall safety assessment of GM crops.

Insights into the safety assessment of proteins can be gained by examining the concepts and approaches applied to the risk assessment of food additives or GRAS (generally recognized as safe) food ingredients and chemicals, which are relatively well established (US FDA, 1983).

The first concept to consider in the safety assessment of any transgenic protein is historical information about the safe use or prior exposure to that particular protein or very similar proteins; also known as “history of safe use”. The degree to which history of safe use is documentable relative to the intended new or expanded uses as a food guides subsequent steps for additional testing.

Follow-up testing for food additives or GRAS food ingredients and chemicals typically includes, but is not necessarily limited to, rodent toxicology studies to identify potential health hazards associated with exposure to these substances. However, proteins differ from most food additives or chemicals in a number of important ways. Beginning with the physical properties, food additives and chemicals are generally smaller molecular weight xenobiotics while proteins are relatively large biological macromolecules. The propensity for systemic absorption of any orally consumed substance is typically inversely proportional to the size of the molecule; so small molecules are more likely to be absorbed intact than larger ones (Fricker and Drewe, 1996). Most dietary proteins are sufficiently large that their absorption intact through the gastrointestinal (GI) tract is very limited, without evidence of involvement of channels, active transport, or disruption of tight junctions. This is not the case for all proteins, as certain proteins, such as ovalbumin, are absorbed intact (Tsume et al., 1996). However, the method of absorption and quantitative estimates of the amount of ovalbumin absorbed following oral exposure are unknown. In addition, the biological activity of proteins is dependent on their structural integrity. The acidic conditions and proteolytic enzymes present in mammalian GI systems efficiently denature and degrade most proteins into constituent amino acids and small peptides that primarily serve as a source of nutrients. Accordingly, the protein structure and biological activity of most proteins are usually lost following ingestion (Metcalf et al., 1996).

This document describes a scientifically based, two-tiered weight-of-evidence approach to assess the safety of candidate novel proteins (CNPs) by drawing upon knowledge of the biological and physico-chemical characteristics of proteins, the available testing methods for evaluating intrinsic hazard, and illustrative case studies of protein safety assessment. The scope is limited to food and feed safety, therefore assessment of crop performance and crop safety, including whole food animal feeding studies, pleiotropic effects, environmental aspects, or ethical considerations are not discussed as they have been discussed elsewhere (FAO/WHO, 1991; Atherton, 2002; Chassy, 2002; Küiper et al., 2002; Cellini et al., 2004; Chassy et al., 2004). The potential for allergenicity is only briefly discussed since other publications have recently evaluated this area in detail (Hefle et al., 1996; Hileman et al., 2002; Codex, 2003; Goodman et al., 2005; Ladics et al., 2007; Thomas et al., 2007).

As opposed to a decision tree approach, which provides distinct yes/no decisions leading to a specific next step, the weight-of-evidence approach is more flexible and focuses on the totality of the data in a holistic manner. Ideally, the predictive value of each piece of evidence would be well understood to give certain types of data more ‘weight’ than others during the assessment, thus leading to more confidence in the overall assessment. Currently, the predictive value of individual data may not be clearly defined, but additional experience with the assays and methods utilized will continue to improve their predictive value. This weight-of-evidence approach, in which the individual components of the safety assessment process are prioritized into two tiers of testing, depicted in Figs. 1 and 2, optimizes the efficiency and robustness of the process. Case studies of transgenic protein safety, some derived from peer-reviewed literature and others from regulatory submissions, demonstrate the merits and limitations of the framework. Specifically this framework includes:

#### Tier I: Potential Hazard Identification (Fig. 1)

- History of safe use
- Bioinformatics analysis
- Mode of action
- *In vitro* digestibility and stability
- Expression level and dietary intake

Tier II: Hazard Characterization (when Tier I assessment identifies hazard potential, or does not permit a determination of safety; Fig. 2), one or more of the following studies may be conducted as determined on a case-by-case basis

- Acute toxicology study
- Repeated dose toxicology study
- Hypothesis-based evaluations

The available best practices for use and limitations of various methods for evaluation of CNPs will be described

**Tier I: Potential Hazard Identification**

- **History of Safe Use:** The protein, or a structurally and functionally related one, has a history of safe use/consumption in food and the source of the inserted DNA does not raise any toxicological concerns
- **Bioinformatics Analysis:** The protein does not show significant amino acid sequence similarity to known toxins, anti-nutrients, or allergens
- **Mode of Action and Specificity:** The protein acts as intended with a known spectrum of activity
- **In Vitro Digestibility and Lability:** The protein is readily degraded / denatured by digestive enzymes, pH, and/or temperature
- **Expression Level and Dietary Intake:** Protein expression levels in the food crop or crop by-products are determined such that dietary exposure can be estimated

Fig. 1. Tier I: potential hazard identification.

**Tier II: Hazard Characterization**

Determined on a case-by-case basis and might include one or more of the following:

- Acute toxicology assessment of transgenic protein
- Repeated dose toxicology assessment of transgenic protein
- Hypothesis-based Studies

Fig. 2. Tier II: hazard characterization.

in subsequent sections of this document. Assessment of the safety of CNPs is only one component of the overall safety assessment of GM crops which would typically include analysis of the inserted DNA, environmental aspects, and phenotypic and compositional analysis of the GM crop, however, all of these components are beyond the scope of the current document.

### 3. Biology and chemistry of proteins

#### 3.1. Dietary proteins do not typically represent a hazard

The ability of a living organism to grow and reproduce is dependent on the complex interaction of that organism with a variety of biochemical substances of diverse structure and function that exist in the environment. Many of these substances are complex macromolecules composed of amino acids (proteins), nucleic acids (DNA and RNA), oligo and polysaccharides (carbohydrates), and fatty acids (lipids). Many proteins have been isolated and characterized and, accordingly many different functional classes of proteins have been distinguished. Despite the large number of proteins that have been isolated, only a small number have toxic properties. Even among those proteins that have demonstrated toxicity in humans and other mammals, very few have demonstrated toxicity following oral exposure. This is primarily attributable to the fact that proteins are relatively large and labile in the digestive tract and therefore are typically not absorbed intact following oral exposure.

##### 3.1.1. Protein structure and function

Proteins are polymers made from 20 common amino acids connected through covalent peptide bonds in specific

sequences encoded by DNA in the genome. A few amino acids may occur in a modified form (e.g. hydroxyproline), however, they represent only a small fraction of the total amino acid pool. The charge, polarity and hydrophobicity of the amino acids of any proteins largely direct the three-dimensional structure of the protein. Tertiary structural features of proteins can include post-translational modifications such as inter- and intra-chain disulfide bond formation, phosphorylation, and acetylation. Together, these characteristics determine the structure and thus the function of the protein (Smolin and Grosvenor, 2000; Devlin, 2002).

The human body can produce 11 of the 20 common amino acids from dietary precursors. The other nine amino acids, which cannot be synthesized by the human body, are therefore considered “essential” because they must be obtained exogenously through the diet. In addition, two amino acids that are derived from essential amino acids are considered to be “conditionally” essential because they can be derived metabolically from other amino acids (i.e., tyrosine, which is derived from phenylalanine, and cysteine, which is derived from methionine; Smolin and Grosvenor, 2000; Devlin, 2002).

The number of theoretically possible combinations of amino acids used to produce proteins, as a general class of substances, is practically limitless. Accordingly, the number and variety of protein sequences that exist in nature is very large. To date, the amino acid sequences of more than 2.8 million proteins of diverse structure and function have been either experimentally determined or predicted based on DNA sequence and are annotated in public databases (PFAM, 2005; Uniprot-Swissprot Consortium, 2007; <http://www.expasy.org/sprot/>). Despite the appar-

ently large number of “possible” proteins anticipated on the basis of probability alone (e.g.  $20^{500}$  if the average protein is 500 amino acids long; additional permutations exist if one includes shorter and longer proteins) approximately 74% of all known proteins can be classified into just under 9000 different families according to their relatedness in structure and function (PFAM, 2005). This is a surprisingly narrow constraint on protein structure and function, which has been imposed by evolution and physical chemistry (i.e., not all amino acid combinations are preferred). Most of these protein families are present in both animal and plant tissues and, therefore, the proteins therein could be consumed in the diet.

To better understand the wide range of functional proteins found in plants and animals, an examination focused on mammals and plants is illustrative. It has been estimated that there are more than 250,000 different functional proteins in mammals (Smolin and Grosvenor, 2000). The size of these proteins varies considerably. For example, smaller soluble proteins such as lysozyme (15 kDa) and albumin (69 kDa), medium sized proteins such as immunoglobulins (150 kDa (IgG) to 950 kDa (IgM)), and large proteins such as blood coagulation factor VIII (1120 kDa) have been characterized (Devlin, 2002; Smolin and Grosvenor, 2000). Soluble proteins are found in blood, tissues and cytosol of cells whereas other proteins are primarily located within cell membranes or solid tissues. Many classes of proteins are comparable between mammals and plants (i.e., transcription factors, transmembrane ATPases, etc.). Examples of functional classes of proteins are listed in Table 1. The large diversity of proteins outlined in Table 1 further supports the need for a case-by-case approach in the assessment of protein safety.

The amino acid sequence of a protein is the principle determinant of the tertiary (three-dimensional) structure; however, physical associations between more than one polypeptide chain, when present, determine the quaternary structure that is required for biological function of some proteins. Other aspects such as post-translational processing, pH, and protein–protein interactions can also influence the structure of a protein. Perhaps the most fundamental concept to understanding the biological activity of proteins is that their activity is dependent on structural integrity. That is, loss of protein structure typically leads to loss of biological function.

### 3.1.2. Protein size and lability

The combination of the physical conditions (pH, temperature, and emulsification) and proteolytic enzymes (i.g., pepsin, trypsin, chymotrypsin and bacterial proteases) in the GI system produce an environment that denatures and degrades the structural integrity and functional activity of most dietary proteins.

The potential for systemic absorption of intact dietary proteins (10,000–1,000,000 Da) differs considerably from that for the absorption of smaller molecular weight xenobiotics (200–600 Da). Absorption of small molecular weight

xenobiotics from the GI system is dependent on the specific physical properties of a substance, including the extent of ionization in the GI system, molecular weight, and hydrophobicity. Small xenobiotics are more amenable to passage into the systemic circulation than large intact proteins. However, there are examples of larger proteins being digested into small peptides that are subsequently absorbed intact and bioactive (Strobel, 1998). Nevertheless, a thorough bioinformatics analysis (as described in Section 4.2.2) should identify known bioactive peptides that are contained within the primary amino acid sequence of the transgenic protein. The source and the activity of an identified bioactive peptide would guide further assessment.

### 3.1.3. Protein processing, absorption and synthesis

As part of normal cellular metabolism, proteins are constantly being synthesized and degraded in body tissues. For example, the human body synthesizes approximately 300 g of new protein each day (Smolin and Grosvenor, 2000; Devlin, 2002). Because of this, a constant supply of new amino acids is required some of which are provided by the diet. For example, in the United States, the average protein consumption is 100 g/day (Gerrior et al., 2004). Other amino acids are obtained from endogenous sources from degradation of proteins already present within the body. Degraded enzymes and structural proteins from mucosal cells sloughed off into the gastrointestinal (GI) system are believed to contribute 35–200 g of amino acids daily (Smolin and Grosvenor, 2000; Devlin, 2002).

Degradation of orally consumed proteins begins in the mammalian stomach where the combined action of acidic pH and digestive enzymes hydrolyze the peptide bonds that connect amino acids. The low pH of the stomach often leads to loss of tertiary structure and denaturation of ingested proteins. In addition, pepsin, an endopeptidase active in the low pH environment of the stomach, contributes to protein degradation by cleaving a relatively broad spectrum of peptide bonds. Degradation of dietary proteins continues in the small intestine where proteins and peptide fragments are subjected to an extensive battery of proteolytic digestive enzymes (described in Section 4.2.4.1).

Decades ago, before the low permeability of the GI system to systemic absorption of protein macromolecules was fully appreciated and before most proteins were found to be digestible, protein hormones including somatotropins and insulin were orally administered to humans in an attempt to achieve therapeutic benefits. These efforts were not successful, and this route of administration was abandoned (Astwood, 1970). We now have a better understanding of how proteins are altered by the GI system (Roberts et al., 1996; Webb, 1990). Essentially, all orally consumed proteins are subjected to the same digestive processes regardless of source or function. The process of protein digestion is efficient as only 6–12 g of the daily protein load (100 g from food and recycled, intestinal enzymes and mucosal cells) entering the GI system each day is lost in feces (Smolin and Grosvenor, 2000). It should be noted

Table 1  
Functional classes of proteins

Category	Function	Animal examples	Plant examples
Structural	Maintains cell and tissue integrity, provides structural support, strength and protection	<ul style="list-style-type: none"> <li>• Actin</li> <li>• Myosin – muscle</li> <li>• Keratin – hair, skin, wool, spider webs</li> <li>• Collagen – connective tissue, bones</li> </ul>	<ul style="list-style-type: none"> <li>• Actin</li> <li>• Vacuolar proteins</li> <li>• Microtubules</li> </ul>
Storage	Seed store nutrient proteins required for the growth of the embryo	<ul style="list-style-type: none"> <li>• Ovalbumin</li> <li>• Ferritin</li> <li>• Casein</li> </ul>	<ul style="list-style-type: none"> <li>• Glutenins</li> <li>• Gliadins</li> <li>• Albumins</li> </ul>
Regulatory	Protein hormones control gene expression Involved in growth and differentiation	<ul style="list-style-type: none"> <li>• Polypeptide hormones – insulin,</li> <li>• Growth factors</li> <li>• Cytokines – interferon</li> <li>• Kinases – protein kinase c, mitogen-activated kinase</li> <li>• Signalling – G protein-coupled receptors</li> <li>• Transcription factors</li> </ul>	<ul style="list-style-type: none"> <li>• Transcription factors</li> <li>• Photoreceptors</li> <li>• Hormone receptors</li> <li>• Signalling – G-protein- protein-coupled receptors</li> <li>• Kinases</li> <li>• Activators, repressors</li> </ul>
Enzymes	Most varied and highly specialized class, catalyze a multitude of reactions, catabolic, anabolic, synthetic	<ul style="list-style-type: none"> <li>• Digestion – pepsin, trypsin, lactase, alcohol dehydrogenase</li> <li>• DNA processing – DNA and RNA polymerases</li> <li>• Krebs cycle and glycolytic enzymes</li> <li>• Amino acid biosynthesis enzymes</li> </ul>	<ul style="list-style-type: none"> <li>• Krebs cycle and glycolytic enzymes</li> <li>• DNA processing – DNA and RNA polymerases</li> <li>• Photosynthesis – ribulobisphosphate carboxyase (represents ~50% of the protein in green tissues)</li> <li>• Lignin biosynthetic enzymes</li> <li>• Cellulose biosynthetic enzymes</li> <li>• Amino acid biosynthetic enzymes</li> </ul>
Transporters	Proteins that transport substances (i.e., lipids, vitamins, micronutrients, oxygen) across cell membranes or between different subcellular compartments	<ul style="list-style-type: none"> <li>• Transmembrane ATPases</li> <li>• Hemoglobin</li> <li>• Myoglobin</li> <li>• Serum albumins</li> </ul>	<ul style="list-style-type: none"> <li>• Aquaporin</li> <li>• Transmembrane ATPases</li> <li>• Nitrate uptake proteins</li> <li>• Lipid transfer proteins</li> <li>• Leghemoglobin</li> </ul>
Defense molecules	Antibodies help neutralize invading viruses, bacteria Coagulation and fibrinolysis proteins stop blood loss Hypersensitive response in plants	<ul style="list-style-type: none"> <li>• Specific IgG antibodies</li> <li>• Factor Va, VIIIa, Xa, and thrombin</li> <li>• Plasmin</li> <li>• Defensins, disrupt microbial membranes</li> </ul>	<ul style="list-style-type: none"> <li>• R (resistance) proteins</li> <li>• Hypersensitive response proteins</li> <li>• Systemic acquired resistance proteins</li> <li>• Peroxidases</li> <li>• Chitinases</li> <li>• Glucanases</li> </ul>
Contractile	Endow cells or cell components with the ability to contract, change shape or move about	<ul style="list-style-type: none"> <li>• Actin</li> <li>• Myosin</li> <li>• Troponin C</li> </ul>	<ul style="list-style-type: none"> <li>• Actin</li> <li>• Extensins</li> </ul>

that there are some examples of proteins that are absorbed intact. For example, ovalbumin (stable egg protein), was detected in both the plasma and the lymph fluid after oral administration (Tsume et al., 1996).

The cell membrane is a significant barrier that protects cells from foreign substances (Alberts et al., 2002). While the lipid bilayer of the plasma membrane allows small molecules to passively diffuse into a cell, specialized membrane spanning proteins can transport larger hydrophobic or hydrophilic molecules, including proteins. Other active transport mechanisms, such as pinocytosis, macrocytosis or receptor-mediated endocytosis, can allow very small quantities (0.001–1 directly absorbed, Strobel, 1998).

Despite the relatively efficient system of protein degradation and/or active transport mechanisms, small quantities of dietary proteins can survive the GI tract intact (Ratner et al., 1952). Likewise, certain food matrices and

composition can have a semi-protective effect, but also will influence critical variables in protein digestion including gastric emptying, biliary and pancreatic excretion and peristalsis. In addition, other factors that impact the digestive and absorption processes in certain subpopulations can include: age (infant, elderly), disease (gastroenteritis), and certain types of medication.

### 3.2. Characteristics of proteins that may represent a hazard

As a class of food ingredients, proteins provide nutrients essential for life. Humans have evolved digestive systems that rapidly digest most proteins and convert them into nutrients for further utilization (FAO/WHO, 1996). While most dietary proteins are innocuous, a relatively small number of proteins existing in nature are known to exert toxic effects when ingested. The amino acid sequences for

some of these proteins have been published and they are present in amino acid databases typically used for homology searches. In some cases, these proteins are resistant to digestive processes that degrade innocuous dietary proteins. Of all known toxic proteins, those produced by bacteria are the best-characterized, however, anti-nutritional proteins (e.g., lectins and protease inhibitors) have also been identified in plant sources (Gill, 1982).

### 3.2.1. Pathogenic bacterial toxins

Acute exposure to certain bacterial toxins can cause adverse effects that range from GI tract discomfort to life-threatening dehydration, paralysis, and death (refer to Table 2, Gill, 1987a). Exposure to some bacterial protein toxins occurs via proteins that are formed in food by such organisms as *Clostridium botulinum* or *Staphylococcus aureus*. Botulinum toxins are unique among bacterial protein toxins in that they are produced as protoxins that are progressively degraded as they pass through the GI system. This process liberates a biologically active binary toxin that binds to receptors in the luminal epithelium which facilitate absorption via transcytosis, thereby delivering the intact toxin to the systemic circulation (Sakaguchi et al., 1988). These toxins are extremely potent (Minimal Lethal Dose = 0.5–1 ng/kg body weight) so that only small quantities are necessary to cause paralysis of the nervous system (Sakaguchi et al., 1988; Table 2).

Other pathogenic bacteria are consumed in food, proliferate in the GI system and produce toxic proteins within the intestine thereby bypassing the stomach (Alouf and Freer, 1999). For example, the virulent pathogenic *Escherichia coli* strain O157:H7, as well as *Vibrio cholerae* and *Clostridium perfringens*, produce enterotoxins that attach to intestinal mucosal cells causing direct cytotoxicity (Popoff, 1998). These toxins form pores in cell membranes that allow leakage of water and electrolytes out of the cell into the intestine causing severe diarrhea (e.g. Alouf and Freer, 1999). *V. cholerae* enterotoxin is an example where the sensitivity of a protein toxin to digestive processes may not be a good indicator of the potential for systemic toxicity because the context in which it is typically encountered allows more limited opportunities for exposure to intestinal proteases.

Table 2 is provided as a list of many of the known bacterial and seed toxins. Structural similarity to any of these compounds would be identified in a bioinformatics screen using the amino acid sequence of the transgenic protein, which could trigger further investigation or could serve as the basis for eliminating from development.

### 3.2.2. Plant and animal protein toxins, and antinutrients

Compared to the number of bacterial protein toxins, the number of known protein toxins and antinutrients produced by plants and animals is considerably smaller. Protein toxins that have been identified in plants include antifungal proteins and lectins (Cheeke and Shull, 1985; Liener, 1994a,b; Gatehouse et al., 1999). *In planta* expression of these substances is believed to serve as a protective

function against insect pests and plant pathogens. In some cases they may also be toxic to mammals (Leiner, 1994b; Gatehouse et al., 1999). The most toxic known lectin is ricin, found in the seeds of Castor beans (Robertus, 1991). Ricin is a toxin that interferes with protein synthesis in cells by inactivating ribosomes.

Consumption of uncooked or unprocessed enzyme inhibitors (e.g., trypsin inhibitor, amylase inhibitors) causes anti-nutritive effects by inhibition of digestive enzymes, thereby interfering with the normal processes that digest dietary proteins. At higher levels these enzymes can be toxic. Digestion-resistant lectins (e.g., PHA [*Phaseolus vulgaris* hemagglutinin] found in kidney beans) bind to and directly damage intestinal epithelial cells (Leiner, 1994b). This results in impaired digestion of nutrients and interference with growth of the organism. Many of these substances are safely consumed in the diet because they do not possess an inherent toxicity, are sensitive to digestive proteolysis, or are readily destroyed by cooking (Leiner, 1994b).

In addition to toxic plant proteins, some animals including scorpions and snakes produce venoms that contain protein toxins. However, in most cases, these toxins only exhibit adverse effects when administered via non-oral routes (Magalhães et al., 1998; Sidell et al., 1997).

### 3.2.3. Protein allergens

Although the subject of protein allergenicity has been considered extensively elsewhere (Codex, 2003; Metcalfe et al., 1996; Thomas et al., 2004, 2005; Gibson, 2006), this assessment of protein safety should include at least a cursory discussion of allergenicity. In brief, very few food proteins are allergens, however, all known food allergens are proteins. It is noteworthy that many of the same tools used in comparison of transgenic proteins to known protein toxins are also applicable to comparison with allergenic proteins (e.g., history of safe use, bioinformatics analysis and digestibility).

Historically, evaluation of the potential allergenicity of proteins intended for use in GM crops has involved the use of a decision tree strategy (Metcalfe et al., 1996). The current assessment strategy, as outlined by Codex (2003), focuses on a weight-of-evidence approach recognizing that no single endpoint can be used to predict human allergenic potential. In that context, the following factors are considered: (i) the source of the gene; (ii) the similarity of the amino acid sequence of the protein of interest to that of known allergens; (iii) the stability of the protein to digestion by pepsin in an *in vitro* digestibility assay; and (iv) when necessary, in *in vitro* human sera testing or clinical testing (Codex, 2003; Goodman et al., 2005).

The factors noted above describe the methodology used to determine whether the transgenic protein is similar to that of known allergens and to evaluate the potential for cross reactivity in those persons sensitized to the known allergen. A comparison of the primary amino acid sequence of the transgenic protein to the amino acid sequences of known allergens is used to identify such similarities

Table 2  
Examples of adverse effect levels for bacterial and seed toxins in different species

Toxin	Source organism	Dose/kg of body weight	Mode of action
<i>Pore former</i>			
Aerolysin	<i>Aeromonas hydrophila</i>	7 µg (MLD, iv, mouse)	Pore former
Alpha-toxin, alpha-lysin	<i>Staphylococcus aureus</i>	1.3 µg (LD <sub>50</sub> , oral, rabbit)	Pore former (lyses membrane of platelets and monocytes)
Delta-lysin	<i>Staphylococcus aureus</i>	40,000 µg (LD <sub>50</sub> , oral, rabbit)	Pore former (lyses membrane of erythrocytes and many other cell types)
Enterotoxin	<i>Clostridium perfringens, Type A</i>	81 µg (LD <sub>50</sub> , iv, mouse) 100 µg (diarrhea, oral, mice) <sup>d</sup>	Pore former
Listeriolysin	<i>Listeria monocytogenes</i>	3–12 µg (MLD, mouse)	Pore former
Pneumolysin	<i>Staphylococcus pneumoniae</i>	1–5 µg (LD <sub>50</sub> , iv, rabbit)	Pore former/sulfhydryl-activated cytolysin (lyses cholesterol containing membranes)
Theta-toxin, perfringolysin O	<i>Clostridium perfringens, Type A</i>	5–8 µg (LD <sub>50</sub> , iv, mouse)	Pore former/sulfhydryl-activated cytolysin (lyses cholesterol containing membranes)
<i>Protease</i>			
Kappa-toxin	<i>Clostridium perfringens, Type A</i>	1500 µg (LD <sub>50</sub> , iv, mouse)	Collagenase
Lethal factor	<i>Bacillus anthracis</i>	<114 µg (LD <sub>50</sub> , iv, rat)	Protease of MAPKs (blocks phosphorylation and subsequent protein–protein interaction)
Neurotoxin	<i>Clostridium botulinum, Type A</i>	0.001 µg (MLD, oral, human) 1.0 µg (LD <sub>50</sub> , oral, mouse) <sup>a,b</sup>	Endopeptidase targeting SNAP-25 (blocks neurotransmitter release)
Neurotoxin	<i>Clostridium botulinum, Type B</i>	0.0005 µg (MLD, ip, mouse)	Endopeptidase targeting synaptobrevin-2 (blocks neurotransmitter release)
Tetanus toxin, tetanospasmin	<i>Clostridium tetani</i>	<0.0025 µg (MLD, human)	Endopeptidase targeting VAMP (synaptobrevin) (blocks transmitter release)
<i>Protein synthesis inhibitor</i>			
Ricin	<i>Castor bean</i>	30,000 µg (MLD, oral, mouse) <sup>c</sup>	Cleaves adenine 4324 of 28S RNA
Shigella toxin	<i>Shigella dysenteriae</i>	<0.009 µg (LD <sub>50</sub> , ip, rabbit)	Inactivates 60S ribosomal subunit
<i>Adenylate cyclase regulator</i>			
Cholera toxin	<i>Vibrio cholerae</i>	250 µg (LD <sub>50</sub> , oral, mouse)	Activates intracellular adenylate cyclase
Heat-labile enterotoxins	<i>Escherichia coli</i>	250 µg (LD <sub>50</sub> , iv, mouse) 100 µg (diarrhea, oral, mouse) <sup>f</sup>	Activates intracellular adenylate cyclase
<i>Glycosylation</i>			
Cytotoxin	<i>Clostridium difficile</i>	220 µg (LD <sub>50</sub> , ip, mouse)	Glycosylation of small GTP-binding proteins on many cell types
Enterotoxin, toxin A	<i>Clostridium difficile</i>	0.5 µg (LD <sub>50</sub> , ip, mouse) 2000 µg (diarrhea, oral, mouse) <sup>c</sup> 16,000 µg (MLD, oral, mouse) <sup>e</sup>	Glycosylation of small GTP-binding proteins on intestinal epithelial cells

Ref: Adapted from Gill (1987a,b, 1982).

MLD: minimum lethal dose; iv – intravenously, ip – intraperitoneally.

LD<sub>50</sub>: dose that causes 50% lethality.

<sup>a</sup> Maksymowych et al. (1999).

<sup>b</sup> Hauschild (1989).

<sup>c</sup> Lyerly et al. (1985).

<sup>d</sup> Skjelkvale and Uemura (1977).

<sup>e</sup> Ishiguro et al. (1983).

<sup>f</sup> Finklestein (1973).

(Hileman et al., 2002; Codex, 2003; Thomas et al., 2004, 2005; Goodman et al., 2005; Ladics et al., 2007). The sequence comparison can be performed early in the process of producing a transgenic crop for commercialization. Similarities identified by this methodology can be used to determine whether the transgenic crop may possibly introduce a cross-reactive protein even before any transgenic protein has been isolated. When significant similarities are identified, it may be necessary for that particular transgenic protein to either undergo further evaluation (e.g. serological testing or other evaluation of the biological relevance of the similarity) or that it not be taken to the next level of commercial development.

Serological testing to evaluate the cross reactivity of novel proteins with IgE antibodies from individuals with allergies to known food allergens may also be undertaken if the source of the gene is from a known allergenic food or if the sequence homology search identified homology with known allergenic proteins. In certain circumstances, further clinical testing may also be required. If tests indicate the protein binds IgE from those with allergies, it is unlikely that the novel protein would be developed and registered as a commercial product (Codex, 2003; Metcalfe et al., 1996; Thomas et al., 2005; Gibson, 2006).

An evaluation of the stability of the transgenic protein to digestion by pepsin is also undertaken early in the safety assessment process. The digestibility assay was first conceived as a means to determine the relative stability of a protein to low pH and pepsin protease encountered in the mammalian GI environment. It was originally developed and used as a method to assess the nutritional value of protein sources by predicting amino acid bioavailability (Astwood, 1970). It was then applied in a systematic fashion to testing allergenic food proteins (Astwood et al., 1996) and a standardized procedure has recently been evaluated in an international multi-laboratory ring study (Thomas et al., 2004). While not all stable proteins are allergens (Fu et al., 2002; Herman et al., 2006), for the purpose of allergenicity evaluation, digestible proteins are believed to have lower potential for systemic exposure of the intact protein, and this observation is relevant to other aspects of hazard characterization, i.e., the potential toxicity of proteins.

#### 4. Testing strategy

##### 4.1. A tiered, weight-of-evidence approach to protein hazard assessment

The principles proposed in the current document present a systematic method to evaluate the safety of transgenic proteins, and where necessary, evaluate the lack of anticipated adverse effects in the context of new or expanded consumption patterns. Scientifically, it is not possible to prove a negative (i.e., the complete lack of hazard) with certainty. However, it is possible to provide a scientific basis to conclude that any particular protein is as safe as

another within defined limits. Therefore, a two tiered, weight-of-evidence approach for the safety assessment of transgenic proteins using the types of data or information described below was developed. This chapter also describes the application of an integrated testing strategy, the components of which are applied to the overall weight-of-evidence approach, and provides examples to illustrate appropriate strategies to evaluate the safety of various proteins.

Knowledge about the biochemical properties, amino acid sequence, function, source, and mode of action of a protein can provide insight into the history of safe use and the potential of a protein to cause adverse effects in mammals and other organisms. These properties are known for many toxic proteins as well as the vastly greater number of non-toxic proteins. The tiered application of the tools described within this section can be used to determine, with a high degree of confidence, if the sequence and biochemical properties of CNPs are similar to or different from those of known proteins, including toxic proteins.

The potential hazard assessment Tier 1 includes components to be applied early in the development process. Components of this tier require information about the CNP but only small amounts of the transgenic protein or none at all (Fig. 1).

When data from the basic hazard assessment (i.e., Tier 1) is equivocal, incomplete or identifies a potential hazard, a second tier supplemental assessment (Tier 2) can supply further information as determined on a case-by-case basis and quantitative risk assessment principles may be applied. Risk is a function of both hazard and exposure; therefore, in the absence of a hazard, risk cannot be defined. In most cases a quantitative risk assessment for dietary proteins is thus unnecessary. In contrast to Tier 1 components, those in the second tier will likely require production, isolation, and characterization of larger quantities of the CNP.

In most cases, bacterial or other expression systems are employed to obtain the CNP to conduct these studies for practical reasons because the expression level of most transgenic proteins in plant tissues is so low that it is seldom physically possible to purify the quantities of CNPs from GM crops to conduct the analyses. One of the limitations of conducting these studies with proteins obtained from heterologous systems (e.g., bacteria) is that they may not be folded or otherwise post-translationally modified as they would be when expressed *in planta*. Ideally, studies conducted to assess the individual components of either tier should be conducted with recombinant CNPs that are biochemically and functionally equivalent to those expressed *in planta*. While there is no standard battery of analytical techniques to demonstrate equivalence, functional and physico-chemical equivalence studies including a number of different techniques to determine protein size, N-terminal amino acid sequence comparison, immunoequivalence, post-translational modification, and functional activity of the gene product are typically conducted

with CNPs isolated from heterologous systems prior to conducting any of these studies (Codex, 2003).

As described in the examples in Sections 5.2–5.5, numerous GM crops on the market today express transgenic proteins that demonstrate a reasonable certainty of no harm (no hazard identified) within the first tier testing and would not require supplemental studies (no hazard characterization) to support the same conclusion. In some cases, developers of GM crops decide to conduct additional testing to confirm the protein safety profile or to address questions that may be raised by regulatory agencies. The subsequent sections of this document present a more detailed description of the recommended testing approach and possible outcomes.

#### 4.2. Tier I: potential hazard identification

The individual components of Tier I (potential hazard identification) aid in the identification of potential hazards associated with the CNP. The components include an assessment of the biological function or mode of action and intended application of the CNP, assessment of the history of safe use, and bioinformatics comparison of the amino acid sequence to other proteins to search for similarity to those known to be toxic, allergenic, or anti-nutritive. A recommendation to evaluate certain physical properties of the transgenic protein (e.g., resistance or sensitivity to digestive enzymes) was also included in Tier I. The collective components in Tier I provide a comprehensive assessment that in many cases allows for a determination of whether the particular CNP presents a potential concern for hazard, even though evidence of how to weigh each individual component is still evolving. Concerns raised by one or more of the components of Tier I may drive the safety assessment to consider Tier II.

##### 4.2.1. History of safe use

The history of safe use for any chemical or protein consists of documented evidence of animal and/or human dietary consumption of sufficient duration within a defined population such that it can be concluded that the existing dietary exposure has demonstrated a reasonable certainty of no harm for the majority of consumers.

The history of safe use (HOSU) concept is widely used in a regulatory context to provide guidance on the level of familiarity with respect to probable safety of chemicals or proteins. To establish HOSU, documented evidence should include evidence of exposure and dietary intake estimates in humans and animals (using consumption modeling techniques) and epidemiological or experimental evidence of no harm. Mode of action and specificity of individual proteins can also be considered. While complete absence of HOSU for specific proteins does not indicate that the protein presents a hazard, it could indicate that the weight of evidence from Tier I analysis is incomplete or inconclusive. Such a finding could lead to recommending further analysis of other Tier I components or possible

additional toxicology (Tier II) testing as determined on a case-by-case basis.

The US Food and Drug Administration (FDA) has applied the HOSU concept to GM crops as an extension of the generally recognized as safe (GRAS) process that was created for the safety assessment of food ingredients in 1986 (21 CFR 170.30(f)). In 1992, the US FDA elaborated on this view by specifically including the concept of HOSU in a testing decision tree for proteins in biotechnology (57 FR 22984). In this decision tree, the FDA indicates that proteins that are the same or similar to proteins already in the food supply should be considered GRAS and therefore are exempt from further hazard characterization.

The HOSU concept has also been articulated in recent guidelines developed by the European Food Safety Authority (EFSA) (EFSA, 2006a). In these guidelines, EFSA states that a key consideration in designing a testing strategy for an introduced protein is the analysis of HOSU. As stated in the guidance, “The studies required to investigate the toxicity of a newly expressed protein should be selected on a case-by-case basis, depending on the knowledge available with respect to the protein’s source, function/activity and history of human/animal consumption. In the case of proteins expressed in the GM plant where both the plant and the new proteins have a history of safe consumption by humans and animals, specific toxicity testing might not be required.” (EFSA, 2006a)

In addition to guidance from FDA and EFSA, numerous authoritative consensus documents on food safety have advocated the case-by-case approach and the use of the HOSU concept. For example, Codex has also described the HOSU concept (Codex, 2003).

The primary limitation to the concept of HOSU is the lack of scientific precision associated with the definition (Constable et al., 2007). It is reasonable to conclude that CNPs that are identical in sequence to proteins already found in foods that have been consumed to a significant extent without evidence of adverse effects pose no safety concern assuming similar consumption uses or patterns. However, when the similarity between transgenic proteins and common food ingredients becomes less absolute, the degree of similarity may be defined, but limits for what constitute “similar” are determined on a case-by-case basis. Notwithstanding the subjective application of this concept when applied to proteins that are structurally and functionally similar (but not identical) to those found in the diet with a documentable history of safe use, this is a useful concept in the hazard assessment process of proteins.

Knowledge about the properties of the organism from which the gene was derived can provide further information regarding the history of safe use. It is important to know whether the source organism from which the gene of interest is derived is known to produce toxic, pathogenic, allergenic, or anti-nutritional effects in humans. Genes obtained from sources known to produce proteins that cause adverse effects may be more likely to encode toxic or allergenic proteins if for no other reason than that

they were obtained from a source capable of producing such proteins. Information about the source will assist in identifying tools and relevant data to be considered in the toxicity and allergenicity assessments. Typically the CNP will have been selected on the basis of a known and desired mode of action, which in turn will provide a preliminary indication of hazard. For example, a protein obtained from a common food source may have a more robust history of safe use than a protein obtained from organisms known to be toxic (e.g., *Clostridium botulinum*, the source of botulism toxin) or allergenic (e.g., *Arachis hypogaea* – the source of peanut allergens). For allergy assessment, because 90% of all food allergic reactions are ascribed to eight major food groups, a gene sourced from one of these foods would trigger the need to consider specific serological testing (Thomas et al., 2005; Gibson, 2006). Concerns about the pathogenicity of the source organism are likely to lead to questioning the safety of any protein, even with a well-defined mode-of-action, derived from that organism and might therefore require a more detailed evaluation. However, unless the transgene of interest is known to mediate or participate in the pathogenicity of the source organism, it is unlikely to impact the safety of other proteins obtained from the same organism. The underlying mode of action of a pathogenic organism will direct the evaluation using the testing strategies provided in this hazard assessment.

#### 4.2.2. Bioinformatics analysis

Bioinformatics, in the context of protein analysis, is a term used to describe the application of a scientific discipline, one component of which includes relating biological information to specific amino acid sequences in known proteins. The primary purpose of bioinformatics assessments of transgenic proteins is to assess the degree of amino acid sequence similarity, phylogenetic relationship, or the orthology between different proteins. In this case, the comparison is essentially conducted between a CNP and the amino acid sequences of all known proteins. Information from this type of comparison can be particularly useful where easily recognized amino acid sequence similarity is identified between the transgenic protein and proteins that have a history of safe use in food and/or feed (see Y-coat protein example). Likewise, similarity to those proteins known to be allergenic, toxic, or pharmacologically active can lead to new hypotheses for additional safety evaluations, depending on the nature and degree of similarity. The general concept in bioinformatics testing is that the more similar the amino acid sequence of the CNP is to protein toxins (for example) the more likely it will require hypothesis-based toxicity testing.

Bioinformatics tools include computer algorithms that are used to evaluate the phylogenetic relationships between genes and gene families and to determine the degree of sequence similarity between two or more sequences (DNA or protein). Bioinformatics tools can also be used to predict the tertiary structure of proteins; however, the number of proteins with known three-dimensional struc-

tures against which to compare is considerably smaller than the number of those with known amino acid sequences.

The most commonly used bioinformatics computer tools; such as Fast protein comparison or Fast All (FASTA; Pearson and Lipman, 1988), and Basic Local Alignment Search Tool (BLAST; Altschul et al., 1990) were designed to assess overall similarity between protein sequences. Various scoring matrices assign more “weight” or bias to amino acids that are responsible for tertiary structures, thereby anchoring potentially similar sequences to regions most likely to have a common tertiary structure (e.g., BLOSUM62 scoring matrix algorithm, <http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/Scoring2.html>). A single search with one query sequence (presumably that of the CNP) identifies the closest matches and orders them according to the level of similarity out of all sequences available in a large sequence library. A number of searchable protein or DNA databases such as Uniprot–Swissprot (<http://www.expasy.org/sprot/>) or the National Center for Biotechnology Information at the US National Institute of Health ([www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) have been used for this purpose.

Information obtained from bioinformatics analyses can be used to define or refine the mode of action of specific proteins. Proteins that are highly similar at the level of amino acid sequence might have related, but not necessarily identical, modes of action, for example hydrolase activity with differing specificity (see Section 4.2.3). Another example is the insecticidal Cry proteins from *B. thuringiensis* (*Bt*) subspecies that share a high level of sequence homology and the ability to disrupt insect midgut membranes in sensitive species by causing pore formation and ultimately the death of a limited spectrum of sensitive species. However, the receptor or binding site specificity of individual *Bt* proteins can be extremely narrow, with some affecting only certain Lepidopteran pests and others affecting only Coleopteran pests (Höfte and Whitely, 1989).

Another use of bioinformatics is the development of phylogenetic trees for proteins that are structurally aligned or with similar sequences to putatively characterize the similarity between proteins, or a group of proteins, to other proteins already present in food or feed. These analyses may be used in combination with HOSU analysis to establish that an entire class of proteins may already be ubiquitous in food and that the mode of action(s) relevant to that class is unlikely to cause adverse effects in humans, animals, or the environment.

Although specific guidelines have been developed for allergenicity (Hileman et al., 2002; Codex, 2003; Goodman et al., 2005; Ladics et al., 2007), guidelines for similarity of homologous proteins in the areas of toxicology, pharmacology, mode of action, and history of safe use are determined on a case-by-case basis. The general concept that is evaluated in this type of analysis is that homologous proteins tend to have the same or similar function. In the context of comparing the amino acid sequence of a CNP to

known proteins toxins, ideally, bioinformatics analysis can be used to demonstrate “lack of similarity” between CNPs and protein toxins. While no formal guidelines have been established for what constitutes a significant sequence similarity between a CNP and a protein toxin, one general recommendation is that proteins sharing less than 20% identity over 100 or more amino acids should not be considered homologues (Doolittle, 1990).

The power and robustness of bioinformatics analyses has increased exponentially since the early 1980s, when only limited numbers of gene and/or protein sequences had been determined empirically. Today there are millions of sequences in these databases with more added each day. This implies that there is a great deal of representative data upon which to base hypotheses and conclusions with respect to functionality, orthology or phylogeny. Currently, bioinformatics analysis is mostly limited to linear sequence analysis with higher order suggested by certain residues. As structural analyses progress, bioinformatics should become an even more powerful tool. However, more biologically relevant test data are needed in order to increase the predictability of bioinformatics search results and databases that may be used to support safety assessments.

#### 4.2.3. Mode of action and specificity

Knowledge about both the mode of action and the functional specificity of a CNP are important factors in the hazard assessment of transgenic proteins as they are the basis of the testing strategy and study design. For purposes of this discussion, mode of action is not used in a typical toxicological sense, but is used from a functional perspective. The mode of action is the mechanism, by which a protein acts *in vivo*. For most currently available GM crops, the mode of action applies to the *in vivo* interaction of the protein (e.g., Bt-proteins) with receptors of susceptible insect species or the interaction between the transgenic protein and a particular class of herbicides *in planta*. Uncertainty about the safe use of any particular protein is reduced if the mode of action can be shown to have low relevance for humans. The better the specificity of a mode of action is understood for any protein, the greater the probability to predict whether it will have adverse effects. Before a CNP is considered for transfer into a crop, the physico-chemical properties, structural characteristics, function, and the mode of action of the protein should be assessed.

Specificity is one of the most significant features of protein molecules in biological systems. As described in the previous section on bioinformatics, the relatedness between proteins can be determined from amino acid sequence similarity to better understand the functional role or classification of any particular protein. Similar proteins can be identified in many different organisms that serve the same purpose. For example, an enzyme is specific for a reaction because its active site is constrained to react only with its own substrate(s). Similarly, the specificity of peptide hormones is based primarily on the interaction with specific receptor proteins.

Information on the role of a particular protein in the donor organism and the systems in which it interacts can assist in evaluating the potential risk of human and non-target exposure to the transgenic protein in other organisms, such as plants. This insight is based on whether the conditions for the protein to perform its function (including the availability of specific molecules that the protein interacts with [e.g., substrate, cofactors, receptors]) exist in the bodies of humans, other mammals, and/or non-target organisms.

For instance, the mechanism of glyphosate tolerance in GM plants is based on the fact that the only known physiological target of glyphosate is the plant endogenous EPSPS – a key enzyme involved in the shikimic acid pathway of the aromatic amino acid biosynthesis. Different from plant endogenous EPSPS, the transgenic EPSPS expressed in the GM plants is not inactivated by glyphosate, thereby conferring the tolerance to glyphosate (Nida et al., 1996; Padgett et al., 1995; see details in case study 5.3).

The protein that a transgene encodes should be characterized to determine the physico-chemical properties and structural characteristics, including protein size; isoelectric point; post-translational modification; amino acid sequence; stability to pH, temperature, and chemical or biochemical agents; and secondary and tertiary structure. Secondary and tertiary structural information may be relevant to understand the mode of action for a protein at the molecular level. Protein characterization provides fundamental information about a protein, which is useful for evaluating bioinformatics analysis and essential to elucidation of its mode of action.

After characterization of a protein, the function and mode of action of the protein at the cellular and molecular level should be investigated. Because of the broad diversity of protein functions and properties, there is no single approach or study design that one can follow to investigate the mode of action for different proteins and this must be determined on a case-by-case basis. As an example, several insecticidal crystalline (Cry) proteins discovered from the bacterium *Bt* have been used in transgenic crops to confer resistance to certain insect pests. Investigation of the mechanism of toxicity of Cry proteins in targeted insects is typically conducted by evaluating the function of the specific activity of Cry proteins in nature. Based on the solubility of the *Bt* proteins as well as specific insect midgut receptors that are not found on mammalian cells in the digestive tract, these Cry insect toxins have no effects on humans, other mammals, or non-susceptible insects even if they could survive passage in the gut (Höfte and Whitely, 1989; Van Rie et al., 1989; see details in case study 5.4).

If there is any safety concern for the transgenic protein after protein characterization, elucidation of the mode of action and specificity, further *in vitro* or *in vivo* testing may be conducted. Any data and information about the function and mode of action of a protein should be reviewed together with data generated by other techniques. There is always the possibility that the function and mode

of action of a protein in one biological system may not be applicable to another. Mode of action and specificity studies provide information that is directly relevant to safety, and these data also guide collateral studies to enhance the overall assessment of protein safety. Therefore, studies conducted in a relevant model for the target species are important.

#### 4.2.4. *In vitro* evaluation of stability

Proteins that are unstable in the GI system are more likely to be safe following oral consumption than those that resist digestion if for no other reason than they are unlikely to retain biological activity following degradation. The potential instability of proteins to pH and digestive enzymes are part of the hazard identification of CNPs. Specifically, the pH and digestive enzymes examined represent the conditions to which the CNP would likely be exposed during consumption. In some cases, information about the stability of a CNP under standard conditions of heat and other forms of processing sometimes applied to obtain edible fractions from grains may also provide further information about the potential exposure to CNPs. However, it should be noted that this level of information is typically not required. The endpoint to be examined for thermal and pH effects is the functionality of protein (e.g., enzymatic activity, bioactivity, etc.), while the presence or absence of protein sequence integrity is the focus for digestibility. Knowing the stability of the protein expands the general knowledge base of the characteristics contributing to a more complete understanding of the mode of action of the protein and provides information about the condition of the protein during human, animal and non-target organism exposure.

**4.2.4.1. *In vitro* digestibility.** Many proteins are sensitive to proteolysis in the mammalian GI system because of the acidic conditions and the abundant quantity and variety of proteases and peptidases in the stomach and the mechanical processing. However, little is known about the stability of most dietary proteins when considered in the context of the total number of dietary proteins that exist in nature, or in the context of complex food matrices as would be found in a typical meal.

Some dietary proteins are degraded within a matter of seconds during exposure to digestive enzymes, whereas others are resistant to *in vitro* digestive proteolysis even following extended digestion. Some proteins appear to be digested completely, albeit slowly when evaluated in *in vitro* digestibility assays. The digestion of many proteins lies in the middle of the continuum, making the results more difficult to interpret. Loss of integrity of some intact proteins has also been reported, in some cases, to lead to the generation of stable lower molecular weight peptides in SDS-polyacrylamide gel electrophoresis (SDS-PAGE) analysis. Evidence of slow or limited protein digestibility does not indicate that the protein is necessarily a hazard, but proteins that are resistant to digestive proteolysis might

be more likely to be absorbed in a biologically active form, if for no other reason than that more of the active form reaches the absorptive part of the GI system intact.

Simulated mammalian gastric fluid (SGF) has been used to assess the *in vitro* digestibility of CNPs. For *in vitro* digestion studies, SGF is typically prepared according to specifications of the US Pharmacopoeia (USP, 2000). Numerous variations in SGF assay parameters have been reported. Recently, Thomas et al. (2004) established a standardized protocol for evaluating the *in vitro* digestibility of proteins in SGF in the context of an inter-laboratory ring study. Although not as widely utilized as that of SGF, some studies have evaluated the *in vitro* digestibility of proteins in simulated mammalian intestinal fluid (SIF). The SIF differs from SGF in that the pH is neutral (7.5) and the digestion is mediated by pancreatin. Pancreatin is a mixture of different enzymes including amylase, proteases (such as trypsin and chymotrypsin), lipases, and ribonucleases, all of which are commonly found in the mammalian intestine. To gain a more realistic assessment, proteins may also be subjected to sequential digestion. That is, exposure to SGF followed by neutralization and exposure to SIF prior to analyzing protein size by SDS-PAGE. If a protein is digested rapidly during an exposure to SGF alone, or during exposure to SIF following digestion in SGF, the probability of being absorbed by epithelial cells of the small intestine in a biologically active form would be extremely low.

**4.2.4.2. Functional stability to temperature and pH.** Information about the stability of a protein can lead to a more thorough understanding of the mode of action as well as the effects of exposure to various conditions on the transgenic protein. The temperatures and pHs involved in processing of the crop should first be identified to set proper parameters for the testing protocol. The test substance may be the purified protein (recombinant, homologous or heterologous, or native) or specific plant tissue containing the protein, such as the grain. An alternative approach is to evaluate the activity of the transgenic protein in food components after isolation from the GM crop.

The temperatures typically evaluated in such production process analyses range from 0 °C to 100 °C and vary in duration for time intervals of up to 60 min. For example, corn is generally wet or dry milled when being processed for various food applications (Eckhoff et al., 1993). The milling process subjects the grain to temperatures of 50–60 °C for up to an hour, so testing at 50 or 60 °C might be appropriate. Stability to temperature and pH extremes may increase the level of concern about a protein, but stability by itself is not an indication that the protein is toxic or allergenic. Where there are concerns about the harmful or toxic properties of a protein due to protein function *per se*, the demonstration of lack-of-function resulting from cooking, milling or other processing, or otherwise will contribute to the hazard identification of that protein in food or feed (Thomas et al., 2007).

#### 4.2.5. Protein expression level and dietary intake

Expression level(s), tissue expression patterns, and dietary intake of transgenic proteins in GM crops are components of a comprehensive safety assessment. Levels of expression of transgenic proteins in various plant tissues can be determined analytically to: (i) define the amount of protein present in seeds and other plant parts; (ii) calculate expected exposure levels to humans and non-target organisms; (iii) support the effective dose level of the protein needed for the phenotype; (iv) demonstrate stability of the encoded transgenic protein during breeding; and (v) evaluate variations across different environmental conditions.

Typically, intake of a protein is estimated by considering actual expression levels in consumed tissues (i.e., fruit or grain vs. leaves) and by considering a comprehensive evaluation of food consumption practices of the population. Mathematical models for combining food consumption and food composition data have been used for estimating dietary intake, nutritional status and exposure to contaminants or pesticide residues for many years (Exponent, Inc.'s FARE™ and DEEM™ [Washington DC]; USDA WWEIA, NHANES – <http://www.ars.usda.gov/Services/docs.htm?docid=14018>). The most appropriate model will depend upon the specific analysis and on the available data. To characterize the intake of foods, it is necessary to know the amount of food consumed and the frequency with which the food is consumed. It is also important to include the impact of protein stability and processing effects that are evaluated in the stability Section 4.2.4. The intake of proteins can then be estimated by multiplying the intake estimates by the concentration of the transgenic protein in the food. The potential impact on nutritional status can be estimated by calculating intakes using existing databases of nutrient composition for traditional counterparts of the new food and comparing the results to analyses using the nutrient composition for the transgenic food. Sensitive subpopulations can also be included in this analysis when warranted.

Some proteins are found in GM crops at low levels. One strategy for dealing with very low levels of exposure that has not been applied to proteins is that of the threshold of toxicological concern (TTC; Cheeseman et al., 1999). TTC refers to the establishment of a human intake threshold value for all chemicals, below which there is a very low probability of an appreciable risk to human health (60 FR 36581–36596). The TTC is not a threshold at which no toxicity is observed, as the Average Daily Intake (ADI) is in traditional risk assessment paradigms, but rather a predicted exposure threshold at which toxicological data from structurally similar compounds indicates large margins of safety (100×–1000×).

This concept was originally proposed by Frawley (1967) and is the basis for the US FDA's "threshold of regulation" (TR) process for dealing with components of food-contact materials that pose a negligible risk (US FDA, 1983). FDA's threshold of regulation is based on an evaluation of the dietary levels at which toxic effects (typically carcino-

genesis) are observed in long-term animal feeding studies conducted with chemicals which could potentially become components of food (including indirect additives). A threshold was set at a concentration of 0.5 parts per billion (ppb). This concept of food safety has been elaborated by subsequent reports (Munro et al., 1999; Kroes et al., 2000, 2004; Barlow et al., 2001). Kroes et al. (2004) indicated that TTC was not normally applied to proteins because the data were insufficient to define a TTC for allergenicity. In the case where hazards are identified with CNPs in Tier I analysis, it might be possible to use this concept to evaluate existing data to determine if TTC values can be determined for proteins or structural classes of proteins.

#### 4.3. Tier II: hazard characterization

A second tier of testing has been developed to characterize hazard for CNPs when the information from the first tier does not provide clear evidence to make a determination of safety (reasonable certainty of no harm). Specifically, Tier II: hazard characterization components could be conducted to evaluate potential hazards associated with dietary exposure to CNPs on a case-by-case basis (Fig. 2).

When necessary to investigate the safety of CNPs beyond the components in Tier I as determined on a case-by-case basis, studies in the second tier can include an evaluation of the acute toxicity of the purified protein. Studies to assess the acute toxicity of CNPs are conducted because it has been stated that proteins that are toxic most likely act through acute mechanisms of action (Sjoblad et al., 1992). Additional testing might be undertaken depending on results from the acute study or if data suggest that evaluation of repeated administration is necessary. Acute toxicology studies are typically conducted in mice via oral exposure because that is the most relevant route of exposure. However, in some cases, other routes of exposure such as intraperitoneal or intravenous administration have been conducted (Hérouet et al., 2005).

Another option in hazard characterization is the conduct of repeated dose studies (up to 28 days) with purified transgenic proteins. However, studies of this type have not been done historically with transgenic proteins. Hypothesis-based testing might be considered if information related to the mode of action raises concern or bioinformatics analysis identifies sequence similarity to known toxins. One or more testable hypotheses should be developed for advanced *in vitro* or *in vivo* testing (see more details in case study 5.5). If the protein is stable to digestive enzymes, determination of the biological fate of the protein may need to be considered on a case-by-case basis.

##### 4.3.1. Acute toxicology assessment of transgenic protein

Acute toxicity testing of transgenic proteins is performed in rodents (usually mice) to assess potential mammalian toxicity following a single exposure to high concentrations of the protein. Acute toxicology studies are required by the EPA to assess the hazards of protein-based pesticides,

which are the active components of microbial pesticides or that are Plant-Incorporated Protectants (PIPs; US EPA, 2000). While acute toxicity studies are not necessarily required for transgenic proteins that are not PIPs, they have sometimes been conducted as well. To date, no tested transgenic proteins used in GM crops have demonstrated adverse effects in acute toxicity studies even at extremely high doses (Table 3). Because the US EPA considers acute toxicity studies “confirmatory” for proteins that have already been used safely within previously approved biopesticides, such as microbial sprays (Betz et al., 2000), additional safety studies beyond acute oral toxicity studies are generally not conducted. The results from these studies with CNPs confirmed the lack of acute toxicity and suggest that there is little actual information to be obtained from conducting them that would not otherwise have been obtained from the first Tier of the safety assessment.

Not all transgenic proteins have pesticidal properties; rather, some impart traits such as herbicide tolerance, drought tolerance, delayed ripening, and alterations in

nutrient content. These traits are imparted by altering the expression of enzymes that catalyze a specific or a narrow spectrum of reactions. Knowledge about these transgenic proteins obtained from Tier 1 typically provides enough information to determine whether they would be likely to cause adverse effects. Consequently, acute dosing studies for non-pesticidal proteins have not been viewed to be necessary to confirm their safety by some regulatory authorities. For example, the European Food Safety Authority does not require acute or repeated dose toxicity testing for proteins with a history of safe use (EFSA, 2006a). Guidelines for testing the safety of “novel” proteins or for GM crops do not specify that acute toxicology testing is necessary to assess potential toxicity. Rather, EFSA advocates a case-by-case basis to determine whether acute and/or repeated dose toxicity studies are necessary to demonstrate a reasonable certainty of no harm that could be attributable to exposure to the transgenic protein (EFSA, 2006a). Codex guidelines also provide latitude for assessing the safety of transgenic proteins, which can be directly or indirectly established as having a history of safe use in the context of the factors described in the Hazard Identification testing.

For PIPs, US EPA and OECD guidelines recommend a high dose (limit dose) assessment in acute oral toxicology studies (US EPA, 1998; OECD, 2002). In these studies, laboratory animals are dosed one time via oral gavage. The limit dose is defined as either 2000 or 5000 mg/kg of body weight by OECD and US EPA, respectively. However, the actual dose used will depend on a variety of factors including the purity and solubility of the test material in an appropriate dosing vehicle. After dosing, test animals are observed daily for 14 days for body weight changes and clinical signs of adverse effects. At the end of the in-life phase, a gross necropsy is also conducted to determine whether obvious pathological changes to major organs are evident. Control groups such as vehicle controls and/or protein controls (administered the same dose of a known non-toxic protein that was given to test animals) are sometimes included. Control groups can provide background information on spontaneous changes and help to interpret the biological relevance of spontaneous findings in test animals. Additional parameters, such as clinical pathology (hematology, blood chemistry) are typically not included in acute toxicity studies, but could be added to the design of acute toxicity study on a case-by-case basis. However, if these additional response variables are included in acute toxicity studies, it is important to know that the laboratory study where they are conducted has sufficient corresponding data from historical control animals against which to compare the results.

The intravenous (IV) route of exposure has also been used to assess the acute toxicity of transgenic proteins (Hérouet et al., 2005). Parenteral routes of exposure may be considered to ensure high systemic exposure based on a considerable experience reported in the scientific literature with IV injections of toxic proteins (see Table 2). However, routes of exposure that bypass the GI system should

Table 3  
Results from acute mouse toxicology studies with proteins in GM crops<sup>a</sup>

Protein	Source	Trait	Dose (mg/kg body weight) and route of exposure <sup>b</sup>
Cry1Ab	<i>Bacillus</i>	Insect	>4000 (oral)
Cry1Ac	<i>thuringiensis</i>	resistance	>4200 (oral)
Cry2A	subsp. <i>kurstaki</i>		>3000 (oral)
Cry3A	<i>Bacillus</i>	Insect	>5200 (oral)
	<i>thuringiensis</i>	resistance	
	subsp. <i>tenebrionis</i>		
Cry3Bb1	<i>Bacillus</i>	Insect	>3850 (oral)
	<i>thuringiensis</i>	resistance	
	subsp. <i>kumamotoensis</i>		
Cry1F	<i>Bacillus</i>	Insect	>600 (oral)
	<i>thuringiensis</i>	resistance	
	subsp. <i>aizawi</i>		
Cry9C	<i>Bacillus</i>	Insect	>3760 (oral)
	<i>thuringiensis</i>	resistance	>0.3 (iv)
	subsp. <i>tolworthi</i>		
VIP3A	<i>Bacillus</i>	Insect	>3675 (oral)
	<i>thuringiensis</i>	resistance	
Phosphinothricin acetyltransferase (PAT)	<i>Streptomyces hygroscopicus</i>	Glufosinate tolerance	>10.0 (iv)
	<i>Streptomyces viridochromogenes</i>		>4200 (oral)
5'-Enolpyruvyl shikimate 3-phosphate synthase (CP4 EPSPS)	<i>Agrobacterium</i> sp.	Glyphosate tolerance	>1000 (oral)
Phosphomannose isomerase (PMI) NPTII	<i>E. coli</i>	Selectable marker	>5000 (oral)
β-glucuronidase (GUS)		Selectable marker	>100 (oral)

<sup>a</sup> Personal communication – Bayer CropScience, BASF, Dow AgroSciences, Pioneer Hi-Bred, Monsanto Company, and Syngenta Biotechnology.

<sup>b</sup> No adverse effects observed at the indicated dose.

be conducted with caution for a number of reasons. First, in most cases the test substances being evaluated are obtained from bacterial expression systems, accordingly even good preparations may contain small amounts of contaminants including bacterial lipopolysaccharides (e.g., endotoxin). Parenteral administration of endotoxin can cause effects including mortality that could lead to the erroneous conclusion that a transgenic protein is acutely toxic. For example, the LD<sub>50</sub> of endotoxin in a 20 g mouse has been estimated to be 100–500 µg (Beutler et al., 1985; Suzuki et al., 2000). Thus, these types of studies carry an inherent risk of overestimating the toxicity of a CNP that could be present in food or feed. Secondly, proteolytic degradation that normally occurs in the GI system following oral exposure does not occur when a CNP preparation is administered parenterally. Thus any potential intrinsic toxicity of CNPs could be exacerbated when, in fact, they present minimal (if any) potential for systemic absorption intact as, in general, IV LD<sub>50</sub> values are considerably lower than oral LD<sub>50</sub> values (see Tables 2 and 3). Another concern with conducting acute toxicity studies with CNPs parenterally is that a number of toxic proteins including PHA-E, have a well-characterized ability to cause adverse effects in the GI system itself. Accordingly, parenteral administration could, in some cases, bypass a potential target organ where adverse changes may otherwise have been detected following oral exposure to the CNP. Finally, there is no accepted limit dose for parenteral acute toxicity studies as there is with oral testing guidelines.

Nevertheless, there are some advantages to conducting acute toxicity studies parenterally. To begin, results from acute toxicity studies conducted with CNPs administered parenterally can be compared to the acute toxicity values of known protein toxins as an indicator of absolute intrinsic toxicity. Intravenous exposure can be considered to mimic worst case scenario by considering that 100% of the entire transgenic protein is absorbed into the body. In addition, the IV route of exposure allows an analysis of potential lethal and non-lethal effects with substantially small quantities of CNP preparation than are required to conduct them orally. Furthermore, in some cases, assessment of the acute toxicity of proteins by IV route of exposure has been recommended by regulatory authorities. In particular, for proteins expressed in non-food plants, such as cotton, where products obtained from these plants will be widely used in surgery and production of hygienic substances, parenteral exposure is feasible.

For CNPs that reveal similarity to known antinutrient proteins from Tier 1 analysis, it may not be prudent to rely solely on the outcome from Hazard Identification and acute toxicity studies for safety assessment. For example, certain lectins and protease inhibitors are toxic to certain insect pests, but also cause toxicity when fed to mammals. Evidence of toxicity with these types of substances may not be observed during acute toxicity testing. In these cases, repeated dose toxicity studies may be warranted (US EPA, 2000; Van Haver et al., 2003).

#### 4.3.2. Repeated dose toxicology assessment with transgenic proteins

As indicated in previous sections of this document, in most cases, repeated dose toxicology studies have not been necessary to evaluate the safety of CNPs. Should it be necessary to conduct repeated dose toxicity studies with CNPs, it is anticipated that the study design would be consistent with those conducted with food ingredients as described in OECD Guideline 407 (OECD, 1995; EFSA, 2006a,b). These guidelines were developed to screen synthetic chemicals for a broad range of possible adverse outcomes. The exposure levels are based on the concentrations of the transgenic protein likely to be encountered in the human diet. However, OECD Guideline 407, section 30, states: "Overall, there is a need for a flexible approach, depending on the species and the observed and/or expected effect with a given compound." It is recommended that rodents would be dosed for 28 consecutive days either by gavage or by adding the test protein to the diet. The guideline defines the limit (e.g., highest) dose in repeated dose 28 day exposure studies as 1000 mg/kg body weight per day. A major impediment to this testing in rodents is the difficulty of producing large amounts of the test protein required to conduct studies of this design. For these types of studies, mice would be preferred to rats in order to reduce the amounts of purified protein required, unless information indicates that rats would be a more appropriate species. However, even with mice, the amount of protein that would be required to conduct a 28-day study would likely exceed 25 g of purified transgenic protein.

As there is no evidence to suggest that protein digestion is altered as a result of repeated exposure, if a CNP is susceptible to digestion, is not homologous to known protein toxins and is not acutely toxic, repeated dose administration of the protein would be unlikely to contribute additional valuable information to the overall assessment. On a case-by-case basis, the contribution of repeated dose oral or dietary toxicity studies with CNPs should be weighed carefully against the potential confounding factors and the difficulty of producing large quantities of protein.

In toxicology, it is well accepted that dose influences toxicological outcome (Conolly and Lutz, 2004). Excessive doses and inappropriate routes of exposure may provide evidence of mechanisms of action that would not occur with relevant routes and exposure levels. In this context, a guideline 'limit dose' of 1000 mg/kg/day dose level in repeated dose toxicity studies may be very high compared to possible human exposures from plant-incorporated transgenic proteins. An alternative approach to setting the high-exposure level would be to start with predicted human exposure, and multiply this value by at least 100, but probably not more than 1000. In most instances, a risk assessment approach will result in a dose that is considerably less than the limit dose of 1000 mg/kg/day.

#### 4.3.3. Hypothesis-based studies

If one of the components of the hazard characterization proposed in Tier 2 suggests a possible hazard, hypothesis-

based testing might be considered on a case-by-case basis. An example is provided for antifungal protein (AFP) where mode of action and bioinformatics sequence similarity data raised possible concerns (see more details in case study 5.5). If a protein or protein fragment is relatively stable to digestive enzymes, determination of the rate of disappearance or the biological fate of the protein might be considered (Thomas et al., 2004; Herman et al., 2003). If exposure assessment indicates a high level of exposure to the protein, additional testing might be considered if dietary intakes significantly exceed a reasonable extrapolation of history of safe use. For example, safe food use at 1 mg/kg/person/day may be acceptable, but food use at 100 mg/kg/person/day may not represent a reasonable margin of safety relative to known dietary intakes or available toxicology data.

## 5. Examples

To demonstrate the effectiveness of the proposed approach in this document to assess the hazard potential of individual proteins, the application of the principles is discussed in relation to six transgenic proteins as examples. Each protein was evaluated using this framework with a discussion about the results and approach that were followed. They are also summarized in Table 4 at the end of this section.

### 5.1. Potato virus Y coat protein

Transgenic expression of the potato virus Y coat protein protects potatoes from infection by Potato Virus Y (PVY). PVY is endemic in potatoes and causes significant damage to the potato crop. Roger Beachy's laboratory characterized coat protein mediated resistance to virus infection in the 1980's and found that expression of coat protein in plant cells prevented reinfection by additional virus (Powell-Abel et al., 1980). The expression level of Y coat protein in transgenic lines is 0.01% of that found in "naturally" infected tubers. Potato virus Y coat protein is identical to the potato virus Y coat protein found in the diet when infected potatoes are consumed (Powell-Abel et al., 1980). Based on data from surveys of tuber infection in the United States and Europe, calculations showed that the consumption of potato virus Y coat protein from GM potatoes would be less than that currently in the diet, even if these GM potatoes were to reach 100% market share. Since there is a strong history of safe use and there is no new dietary exposure, no further safety assessment of the potato virus Y coat protein is warranted.

### 5.2. Phosphinothricin acetyltransferase protein

The phosphinothricin acetyltransferase (PAT) proteins, which are encoded by the *bar* coding sequence from *Streptomyces hygroscopicus* or the *pat* coding sequence from *Streptomyces viridochromogenes*, are present in glufosinate-ammonium tolerant plant varieties of various crops such as corn, cotton, rice, canola (oilseed rape), and soy-

bean. The PAT proteins encoded by *bar* and *pat* are similar enough to be considered homologues (Wehrmann et al., 1996). The PAT enzymes acetylate L-phosphinothricin, the active isomer of the glufosinate-ammonium herbicide, resulting in tolerance of transgenic plants to post-emergent application of the non-selective herbicide. The determination of the safety and toxicological profile of the PAT protein is independent of the crop variety and applies to all plants that express this protein. Because the expression of the PAT protein in transgenic plants, regardless of crop, is extremely low (<0.1% of total protein), safety studies were conducted with PAT proteins produced in *E. coli*. The safety assessment of PAT proteins has been described (Hérouet et al., 2005).

#### 5.2.1. Potential hazard identification evaluation – Tier 1

5.2.1.1. *History of safe use.* The safety assessment begins with the consideration of the safety of the bacterial *Streptomyces* donor organisms. These *Streptomyces* species are common soil microbes, widespread in nature, and found all over the world. They have not been reported as being toxic or allergenic to humans or animals (Kutzner, 1981). Humans are probably frequently exposed to *Streptomyces* through the consumption of roots and vegetables and there is no evidence to indicate harm from such exposure.

5.2.1.2. *Bioinformatics analysis.* Bioinformatics investigations of the PAT proteins did not reveal any indications of similarity to known toxins or allergens (Hérouet et al., 2005). The complete amino acid sequence of the PAT proteins was compared with all protein sequences present in seven large public databases. The algorithm used was BLAST and the scoring matrix BLOSUM62. The epitope search using eight continuous amino acids did not match with any known allergens. The results showed that the PAT proteins have sequence similarity only with other acetyltransferase proteins, for which no adverse effects have been reported following consumption. In addition, there were no potential N-glycosylations identified.

5.2.1.3. *Mode of action.* The PAT proteins share similar structural and functional properties with the class of acetyltransferase proteins, which are widely distributed in nature. The PAT enzyme adds an acetyl group to the substrate L-phosphinothricin, the active isomer of the herbicide glufosinate-ammonium. The addition of the acetyl group produces N-acetyl glufosinate, which is not herbicidally active. Therefore, plants, which contain the PAT enzyme, are resistant to the herbicidal effects of glufosinate-ammonium. The PAT enzymes are highly substrate specific for L-phosphinothricin and studies show that they do not act on even the closest substrate analogs (Thompson et al., 1987).

5.2.1.4. *In vitro digestibility.* PAT proteins were rapidly degraded in human simulated gastric and intestinal fluids (Hérouet et al., 2005). The rapid degradation (within

Table 4  
Summary of examples

Examples	Tier I Potential hazard identification					Tier II Hazard characterization		
	History of safe use	Bioinformatics	Gene source	Mechanism of action/ function	Digestibility <i>in vitro</i>	Acute toxicity	Sub-chronic toxicity	Hypothesis-based testing
Y coat protein (PVY)	Potato virus Y commonly infects as many as 20% of all potato tubers consumed (need exact data for Y coat protein)	Potato virus Y coat protein is not similar to allergens or toxins, and is similar to other plant virus coat proteins commonly consumed	Plant viruses are consumed incidentally	Plant viral coat proteins have no enzyme activity, and the mechanism of 'immunity' is to prevent further infection by Potato virus Y	No data required	No data required	No data required	N/A
Phosphomannose isomerase (PMI)	Found in various plant and animal foods, and is found in many bacteria – also endogenous to humans	PMI shares similarity to other sugar isomerases and no similarity to toxins	<i>E. coli</i>	Well described enzymatic activity with known substrate specificity that poses no concern	Digestible	Non-toxic (NOEL 3030 mg/kg)	No data required	N/A
Phosphinothricin acetyltransferase (PAT)	Found ubiquitously in the environment and on foods	PAT shares similarity to other acetyltransferases and no similarity to toxins	<i>Streptomyces</i>	Well described enzymatic activity with known substrate specificity	Digestible	Practically non-toxic (NOEL 10 mg/kg iv; MOS > 1000)	No data required	N/A
5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS)	EPSPS are ubiquitous in all plants, microbes and many fungi	CP4 EPSPS shares similarity to other EPSPS and no similarity to toxins	<i>Agrobacterium</i>	Well described enzymatic activity with known substrate specificity that poses no concern	Digestible	Practically non-toxic (MOS > 1000)	No data required	N/A
Delta-endotoxin Cry1Ab	Found in many common microbial pesticides used for 40 years (with prior toxicological evaluations)	Cry1Ab shares similarity to other Bacillus crystal proteins and no similarity to mammalian toxins	<i>Bacillus thuringiensis</i>	As an insecticidal protein, the specificity and potency has been characterized and shown to be restricted to Lepidopteran insect pests with no practical toxicity to organisms outside of insects	Digestible	Practically non-toxic (MOS > 1000)	No data required	N/A
Antifungal protein (plant defensins) (AFP)	Found in various plant and animal foods	AFP shares similarity to other plant defensins as well as insect defensins (such as scorpion toxin). Some classes of insect defensins are toxic to animals and humans	Plants	AFP's act by interfering with membrane functionality and growth of fungal hyphae. Insect defensins act directly as neurotoxins via membrane integral channel proteins on the basis of protein–protein interactions	Stable	Not assessed	No data required	Several <i>in vitro</i> studies were completed to show lack of AFP specificity for mammalian neurons – lack of toxicity to mammalian neurons

30 s) shows that these proteins do not survive in the digestive environment.

Since no hazard has been identified during the Tier I assessment process, described herein, there is no reason to perform additional analysis. However, the developers choose to conduct acute toxicology studies via IV exposure to strengthen the evaluation of safety.

### 5.2.2. Hazard characterization evaluation – Tier 2

**5.2.2.1. Acute toxicity in the mouse.** To have a direct assessment of the protein toxicity, an acute study was conducted via the IV route of administration. Animals (mice) treated IV with PAT proteins had no visible signs of systemic toxicity, in contrast to the positive control which induced 100% mortality within 10 min at the same dose. The animals were observed for 14 days followed by a post-mortem examination in which no adverse effects were observed (IV dose of 10 mg/kg body weight). Additionally, no adverse effects were observed in mice following an acute oral toxicity study with PAT at a dose of 2500 mg/kg.

While these additional studies went beyond the recommendation, they served to further confirm that the PAT proteins pose no risk to human or mammalian consumption (Hérouet et al., 2005).

### 5.3. CP4 enolpyruvylshikimate-3-phosphate synthase protein

Glyphosate-tolerant GM plants express the transgenic protein 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS), which is an enzyme derived from *Agrobacterium* sp. strain CP4 (this enzyme is identified as CP4 EPSPS). Glyphosate effectively inhibits plant endogenous EPSPS, thereby disrupting the aromatic amino acid synthesis pathway and leading to plant death. Different from plant endogenous EPSPS, the CP4 EPSPS expressed in the GM plants is not inactivated by glyphosate thereby conferring the plant tolerance to glyphosate (Nida et al., 1996; Padgett et al., 1995; Stallings et al., 1991).

The safety assessment for the CP4 EPSPS enzyme introduced into soybean to produce an herbicide tolerant crop has been described (Harrison et al., 1996; Nair et al., 2002).

#### 5.3.1. Potential hazard identification evaluation – Tier 1

**5.3.1.1. History of safe use.** The CP4 EPSPS protein, derived from a common soil bacterium, is a member of the class of EPSPS proteins found ubiquitously in plants and microorganisms (Padgett et al., 1996).

**5.3.1.2. Bioinformatics analysis.** No similarity was found to known protein allergens or toxins based on amino acid sequence homology searches (Harrison et al., 1996).

**5.3.1.3. Mode of action.** All plants, bacteria, and fungi contain EPSPS enzymes, but they are not present in humans and other mammals because mammals do not have the metabolic machinery to synthesize aromatic amino acids. Therefore, they do not have specific substrates or receptors

with which plant or CP4 EPSPS protein can interact. Based on knowledge of the mode of action of this protein, it is not likely that it will cause harm to humans or animals (Padgett et al., 1991).

**5.3.1.4. In vitro digestibility.** The CP4 EPSPS protein showed rapid degradation in simulated digestive fluid assays (Harrison et al., 1996).

Because no hazard has been identified during the Tier I assessment process described herein, there is no reason to perform additional analysis. However, the developers chose to conduct an acute toxicology study to strengthen the safety evaluation.

#### 5.3.2. Hazard characterization evaluation – Tier 2

**5.3.2.1. Acute toxicity.** The acute oral toxicity assessment of CP4 EPSPS in mice showed no adverse effects at >572 mg/kg body weight (Harrison et al., 1996).

### 5.4. *Bacillus thuringiensis* Cry1Ab protein

The *cry1Ab* gene used to produce insect-protected corn lines is a modification of the *cry1Ab* gene isolated from the DNA of *B. thuringiensis* subsp. *kurstaki* strain HD-1 and was originally designated *cryIA(b)* according to the nomenclature of Höfte and Whitely (1989), which has since been revised to *cry1Ab* (Crickmore et al., 1998). *Bt* corn was developed through a specific genetic modification to be resistant to attack by European corn borer (ECB; *Ostrinia nubilalis*), a major insect pest of corn in agriculture. *Bt* corn produces a truncated version of the insecticidal protein, Cry1Ab, derived from *B. thuringiensis*.

#### 5.4.1. Potential hazard identification evaluation – Tier 1

**5.4.1.1. History of safe use.** Microbial *Bt*-based products have been used commercially more than 50 years by growers, including organic growers. As a demonstration of the safety of *Bt*-based products, formulated microbial *Bts* have no re-entry lag time and no pre-harvest interval, thus microbial *Bt*-treated crops can be harvested and consumed immediately after application (<http://www.epa.gov/oppfead1/cb/ppdc/2002/may02transcript.htm>). Only the safest products lack a re-entry lag time and pre-harvest interval. The US. EPA regulates re-entry and pre-harvest intervals to protect both workers and consumers.

**5.4.1.2. Bioinformatics analysis.** The Cry1Ab protein shows no amino acid sequence similarity to known protein toxins or allergens, other than other *Bt* proteins (Sanders et al., 1998).

**5.4.1.3. Mode of action.** The current consensus on Cry protein mode of action is that the crystals of *Bt* protoxins are first solubilized in the midgut of susceptible insects where the pH is high (e.g. typically 9–11 for lepidopteran larvae).

Solubilization is followed by activation of protoxins by specific midgut protease cleavage to form active toxins, which are roughly half of the original molecular size.  $\delta$ -Endotoxins, such as the Cry1Ab protein, act by selectively binding to specific sites localized on the brush border midgut epithelium of susceptible insect species (Höfte and Whitely, 1989; Zhuang and Gill, 2002). Following binding, cation-specific pores are formed that disrupt midgut ion flow, interfere with nutrient uptake and thereby cause paralysis and death. Cry1Ab is insecticidal only to lepidopteran insects, and its specificity of action is directly attributable to the presence of specific binding receptor proteins, such as aminopeptidases and cadherins, in the target insects (Zhuang and Gill, 2002). There are no binding sites for  $\delta$ -endotoxins of *B. thuringiensis* on the surface of mammalian intestinal cells, therefore, livestock animals and humans are not susceptible to these proteins (Van Rie et al., 1989).

**5.4.1.4. *In vitro* digestibility.** The Cry1Ab protein is rapidly degraded and its insecticidal activity is lost under conditions that simulate mammalian digestion (Sanders et al., 1998).

Although the Tier I assessment did not identify a hazard, the Cry1Ab protein is an insect toxin (a PIP), and it is recommended that an acute toxicity study be conducted.

#### 5.4.2. Hazard characterization evaluation – Tier 2

**5.4.2.1. Acute and repeated dose toxicity.** There were no indications of toxicity as measured by treatment related adverse effects in mice administered Cry1Ab protein by oral gavage at levels greater than 4000 mg/kg body weight (US EPA, 1996, 2000; Table 3).

While additional evaluations are not recommended according to the new assessment described here, a large body of historical studies has further supported a long history of safe use. Acute, repeated dose, and chronic toxicology studies conducted over the past 50 years establish the safety of the microbial *Bt* products, including those that contain Cry1Ab (Betz et al., 2000). The US EPA has determined that *Bt* microbial products show no adverse effects in numerous toxicological studies conducted and has concluded that these products are not toxic or pathogenic to humans. These studies support the safety of Cry1Ab protein and are fully consistent with the history of safe use for the Cry1Ab protein, which has been demonstrated as highly selective for insects, with no activity against other types of living organisms such as mammals, fish, birds, or invertebrates (Betz et al., 2000; US EPA, 1996). The acute oral toxicity test performed, by the product developer, confirms the safety of the Cry1Ab protein.

#### 5.4.3. Phosphomannose isomerase protein

Phosphomannose isomerase (PMI) is an enzyme that catalyzes the reversible interconversion of mannose 6-phosphate and fructose 6-phosphate. PMI is common in nature and is found across animal kingdoms, but is less ubiquitous in the plant kingdom, being absent in many plants (Golds-

worthy and Street, 1965; Lee and Matheson, 1984). Plant cells lacking the enzyme are incapable of surviving on synthetic medium containing mannose as sole carbon source. But expression in plants of the *E. coli manA* gene encoding PMI allows cells to utilize mannose as a carbon source and survive on media containing mannose. PMI, therefore, has utility as a selectable marker for transformation of many plant species which normally do not express this enzyme. The determination of the safety and hazard profile of the PMI protein is independent of the crop variety and applies to all plants that express this protein.

#### 5.4.4. Potential hazard identification evaluation – Tier 1

**5.4.4.1. History of safe use.** PMI is ubiquitous in nature. The gene encoding PMI activity has been cloned from several species of bacteria and yeast, as well as from mammals, including humans (Miles and Guest, 1984; Darzins et al., 1986; Shinabarger et al., 1991; Collins and Hackett, 1991; Schmidt et al., 1992; Smith et al., 1992; Proudfoot et al., 1994a,b). Functionally equivalent PMI enzymes with significant amino acid homology to the *E. coli* PMI have been identified in many organisms (including the gut microflora in humans). Broad expression and exposure in humans and many food items have been shown.

Because a homologue of *E. coli* PMI is expressed by humans, and the PMI protein is found in human intestinal flora as well as some but not all plants, there has always been a background of human exposure and a low quantity of PMI found in the human diet. However, quantitative expression levels in foods are not currently available so a quantitative exposure assessment could not be conducted.

**5.4.4.2. Bioinformatics analysis.** The amino acid sequence of the PMI protein from *E. coli* was compared to the latest posting of the National Center for Biotechnology Information (NCBI) The PMI protein sequence showed no significant homology to any known toxins or allergens (Reed et al., 2001).

**5.4.4.3. Mode of action.** Mannose and mannose derivatives are common constituents of living cells and are major elements of intermediary metabolism. Mannose is phosphorylated by hexokinase to mannose 6-phosphate and in the presence of PMI enters the glycolytic pathway after isomerization to fructose 6-phosphate. The only known function of PMI is to bridge glucose metabolism with mannose 6-phosphate production. The effect of mannose on plants was first described over 40 years ago (Stelid, 1954; Morgan and Street, 1959). Plant cells genetically transformed to express PMI can utilize mannose as a carbon source, thus giving them a growth advantage on mannose-containing media. In the absence of PMI, mannose 6-phosphate accumulates, inhibiting phosphoglucose thereby blocking glycolysis and cell growth. PMI is also present in humans. Lack of the active enzyme in humans is associated with carbohydrate deficient glycoprotein syndrome, which can be fatal, but is treatable (Niehues et al., 1998).

5.4.4.4. *In vitro* digestibility. PMI was rapidly degraded by *in vitro* simulated mammalian gastric fluid (SGF). No intact PMI was detected upon immediate sampling of the reaction mixture (Reed et al., 2001).

Since no hazard was been identified during the Tier I assessment process described herein, there was no reason to perform additional analysis. However, the developers chose to conduct an acute toxicology study to strengthen the evaluation of safety.

#### 5.4.5. Hazard characterization evaluation – Tier 2

5.4.5.1. *Acute toxicity study.* An acute oral toxicity study in the mouse was conducted with PMI protein purified from an *E. coli* over-expression system. PMI was administered by oral gavage to male and female mice at a dose of 3030 mg PMI/kg body weight. There was no evidence of toxicity after 14 days and the LD<sub>50</sub> for PMI protein was greater than 3030 mg/kg (Reed et al., 2001).

The additional toxicity testing performed with PMI protein, confirms the conclusion that the PMI protein is safe (reasonable certainty of no harm) for humans or animals.

#### 5.5. Antifungal proteins

Antifungal plant defensins are highly basic, cysteine-rich proteins found in a variety of food plants that exhibit fungistatic/fungicidal activity towards filamentous fungi by altering hyphal growth and morphology (Terras et al., 1992; Broekaert et al., 1997). Antifungal proteins (AFP) are a ubiquitous class of plant proteins that inhibit fungal hyphae growth. The AFP functions through receptor-mediated mechanisms as sodium channel blockers and potassium channel stimulators. Cell damage from ion leakage (K<sup>+</sup>, Ca<sup>2+</sup>, Na<sup>+</sup>) causes hyperbranching during fungal hyphae tip growth resulting in cell necrosis and ultimately cell death.

##### 5.5.1. Potential hazard identification evaluation – Tier 1

5.5.1.1. *History of safe use.* Antifungal proteins are found in foods in many crops that become part of the diet including potato, pepper, flour, barley and corn.

5.5.1.2. *Bioinformatics analysis.* Antifungal proteins (AFPs) are members of a large superfamily of small, cysteine-rich proteins, termed defensins or magainins, which includes scorpion toxins.

5.5.1.3. *Mode of action.* The mode of action of the antifungal proteins (AFP's) has not been established, however, they have been reported to stimulate the influx of Ca<sup>2+</sup> and the efflux of K<sup>+</sup> ions in filamentous fungi, *in vitro* (Thevisen et al., 1996; De Samblanx et al., 1997), and mutational analysis studies have shown that enhanced antifungal activity is associated with the increased influx of Ca<sup>2+</sup> (De Samblanx et al., 1997). Antifungal proteins show an apparent high degree of specificity toward filamentous fungi; absence

of toxicity towards a variety of human cells *in vitro* has been demonstrated (Terras et al., 1992).

5.5.1.4. *In vitro* digestibility. AFPs are relatively resistant to denaturation by pH, proteolysis and heat. AFPs are also relatively resistant to proteolysis in simulations of the enzymatic environment of the stomach/small intestine and share structural similarity to defensins and magainins, thus identifying a potential hazard under the Tier I criteria and hazard characterization; both acute toxicity and hypothesis driven testing strategies would be recommended. The developer has conducted a set of hypothesis driven testing to better understand the potential biological activity of AFP on mammalian cells. It is important to note that in this particular example, none of the data from the Tier II assessment has been published and was provided as a personal communication (Dr. James Astwood).

##### 5.5.2. Hazard characterization – Tier 2

5.5.2.1. *Acute and repeated dose toxicity.* No acute or repeated dose toxicology testing has yet to be conducted, but would be eventually required if commercialization is sought, consistent with this assessment framework.

5.5.2.2. *Hypothesis-based testing.* Mouse neocortical neuronal cell cultures respond to injurious insults with representative alterations in ion channel signaling systems. These cell cultures were chosen as the testing system because they are derived from nerve tissue, the hypothesized site of likely AFP activity and are routinely used in cell biological, biochemical, electrophysiological, and cytotoxicological investigations under physiological and pathological conditions. Cultured cortical neurons retain most of the intrinsic cellular properties of their *in vivo* counterparts (and, more broadly, of central neurons and peripheral neurons in general) and share high similarity to human ion channels.

The antifungal protein AFP (10 μM) and scorpion toxin CsE-v3 (10 μM) showed no cytotoxic effect on cell viability of cultured cortical neurons. Cytotoxicity was only observed when the concentration of CsE-v3 was increased to 100 μM. This high concentration of AFP was not tested, partly because large quantities of AFP were difficult to obtain and because human exposure to such high concentration is extremely unlikely to occur and represents a 1000× excess over expected daily consumption. AFP at 10 μM did not affect the primary electrophysiological parameters of the neuronal cell membrane such as membrane potential and generation of action potentials, and did not affect voltage-gated Ca<sup>2+</sup> and K<sup>+</sup> currents.

Voltage-gated Na<sup>+</sup> channels are integral plasma membrane proteins responsible for the rapidly rising phase of action potentials in most excitable tissues. Membrane depolarization activates Na<sup>+</sup> channels and produces a rapid increase in Na<sup>+</sup> permeability, followed by a reduction in Na<sup>+</sup> permeability due to a fast channel inactivation process that terminates the action potential. Scorpion toxins typically produce many-fold increases in the inactivation time

constant of Na<sup>+</sup> currents, resulting in massive prolongation of the current (Wang and Strichartz, 1982; Duval et al., 1989; Kaneda et al., 1989; Lee et al., 2000). Such large effect on Na<sup>+</sup> current inactivation typically results in large increases in action potential duration (Duval et al., 1989). In comparison to these prominent known effects of scorpion toxins, the effect of 10 M AFP on Na<sup>+</sup> current inactivation was quite small. Importantly, no change in Na<sup>+</sup> current amplitude or action potential firing rate was seen.

Hypothesis-based testing of the AFP did not affect electrophysiological parameters of mammalian neuron cells as might have been predicted from the bioinformatics analysis that found similarity to known toxins. However, AFP should not be further developed for commercialization until toxicology testing can be conducted and the stability of the protein can be better evaluated.

## 6. Conclusions of testing strategy and recommendations

A robust safety assessment framework for proteins has been outlined in this document. A weight-of-evidence approach is described with a two-step hazard identification and hazard characterization. The Tier I potential hazard identification includes both information about the protein (history of safe use and bioinformatics) and laboratory analysis (expression analysis, digestibility, mode of action) requiring only small quantities of purified protein. While many of the analyses described in this document have been optimized or standardized (e.g. *in vitro* digestibility), bioinformatics analysis for protein toxicity is not as fully defined as bioinformatics analysis for protein allergenicity. Additionally, unlike amino acid sequence comparisons between CNPs and allergenic proteins, there are no guidelines to define the level of similarity between CNPs, known safe proteins and protein toxins, nor are there established databases containing the sequences of proteins that are considered toxic. The utility of the bioinformatics search tools and criteria used to identify similarities to protein toxins has not been experimentally determined. Further work in these areas would standardize the approach by setting criteria that ultimately improve the predictive capabilities of the bioinformatics process. Finally, the entire weight-of-evidence concept would be further strengthened by more standardized methods that have established positive and negative predictive values. This would allow an improved ability to give some data more 'weight' than others during the assessment, thus leading to a greater degree of confidence in the overall safety assessment.

Another concept outlined in this document that could be further developed is the threshold of toxicological concern (TTC). TTC refers to the establishment of a human intake threshold value for all chemicals, below which there is a very low probability of an appreciable risk to human health. It might be difficult to apply this same approach to evaluate the allergenicity of transgenic proteins due to the diversity of human allergic reactions. However, research is ongoing to further define the minimal elicitation dose of an aller-

genic food that would cause an adverse reaction in allergic individuals. It might be possible to evaluate existing data for proteins to determine if TTC values can be determined for proteins or structural classes of proteins.

The tools discussed in the Tier I potential hazard identification evaluate the hazard potential from the physical and functional perspectives. If no potential for hazard is identified and history of safe use is demonstrated, this meets the standard of reasonable certainty of no harm. It is also important to note that this assessment process does not address allergenicity, which is evaluated in other publications (Codex, 2003; Thomas et al., 2004, 2005; Gibson, 2006). As summarized in the examples above, extensive characterization and toxicological evaluations have been conducted for proteins introduced into crops via biotechnology. In light of the proposed tiered, weight-of-evidence approach, studies to date have exceeded what is proposed in this document. It is also evident from these examples that the history of safe use and mode of action/function are useful in determining a hazard identification strategy. By working with proteins that are: (1) already safely consumed from the food supply or are similar to proteins in the food supply, and (2) well characterized with respect to biological activity, it is unlikely that transgenic proteins subjected to this rigorous hazard assessment battery represent a significant risk. Indeed, while experience has shown that many of these proteins have been subject to direct toxicity testing, the weight-of-evidence approach would conclude that direct toxicity testing in the Tier II hazard characterization is not essential to the conclusions of safety, and that these tests should be understood as confirmatory for these cases.

Where uncertainty relating to components of the Tier I potential hazard identification is established from the above testing scheme (Fig. 1), higher tier hazard characterization testing is recommended. The Tier II hazard characterization evaluation focuses on toxicology testing (e.g. acute single dose, repeated dose) and hypothesis-based testing, which is formulated on a case-by-case basis from the results of the Tier I potential hazard identification evaluation. Protein toxins are generally believed to act through acute mechanisms of action (Sjoblad et al., 1992). Therefore, when a protein demonstrates no acute toxicity using a standard laboratory mammalian test species, the findings will support the determination that the protein will be non-toxic to humans and other mammals, and will not present a hazard under realistic exposure scenarios. One difficulty in conducting acute oral toxicity tests is the need for large quantities of highly purified transgenic protein. While an intravenous route of exposure would significantly reduce the protein quantities needed, there is a lack of data comparing the oral and intravenous route of exposure for a range of toxic and non-toxic proteins. This is another area where additional work would be valuable.

A repeated dose oral toxicity study with purified protein would be considered only if there are safety concerns about a dietary protein that were not adequately addressed by the

weight-of-evidence from the wide array of other studies, including the acute toxicity test.

Finally, a comprehensive description of the mode of action of a transgenic protein can also reveal areas of potential toxicity that will require further examination in hypothesis-based testing, as was highlighted by the AFP example. As additional proteins are considered as candidates for transformation into crops, the importance and weight given to history of safe use and mode or mechanism of action cannot be understated. Further, it is anticipated that the Tier II hazard characterization evaluation could become more relevant for future proteins such as those that lack a history of safe use or for which the mode of action is not well defined.

## 7. Support

This manuscript was made possible by support from the members of a task force of the ILSI International Food Biotechnology Committee. The International Life Sciences Institute (ILSI) is a non-profit, worldwide foundation established in 1978 to advance the understanding of scientific issues relating to nutrition, food safety, toxicology, risk assessment, and the environment. ILSI also works to provide the science base for global harmonization in these areas. By bringing together scientists from academia, government, industry, and the public sector, ILSI seeks a balanced approach to solving problems of common concern for the well-being of the general public. ILSI receives financial support from industry, government, and foundations.

Task force member	Representative
BASF Bayer CropScience, LP Cargill, Inc.	Laura Privalle Junguo Zhou Robin Eichen-Conn, <i>Chair</i>
Dow AgroSciences Monsanto Company Pioneer Hi-Bred, A DuPont Company Syngenta Biotechnology, Inc.	Daland Juberg Gary Bannon Bryan Delaney Linda Meyer

## Conflict of interest statement

The authors declare that there are no conflicts of interest.

## Acknowledgements

Task force members wish to express special thanks to Susan MacIntosh for her energy and diligence throughout this project. Past task force representatives James Astwood (Monsanto Company, until October, 2006), Helen Cunny (Bayer CropScience, until February, 2007), Joe Dybowski (Dow AgroSciences, until March 2006), Doug Hard (Renn-

essen, LLC; member through 2005), Trish Malarkey (Syngenta, until November 2004), Joel Mattsson (retired, Dow Agrosciences); and science advisors Joe Borzelleca and John Doull, have each provided thoughtful comments and written text throughout the project. We also would like to acknowledge the contributions of Barbara Petersen, Marcia Van Gemert, and Nga Tran at Exponent during the initial stages of this project.

The Task Force also wishes to thank the ILSI staff members, Ms. Lucyna Kurtyka, Senior Science Program Manager until July 14, 2006, and Dr. Marci Levine, Staff Scientist, beginning July 25, 2006, for their dedicated efforts in seeing this project to completion.

This document has been reviewed by individuals internationally recognized for their diverse perspectives and technical expertise. It must be emphasized, however, that the content of this document is the responsibility of the authors, and not the responsibility of the reviewers, nor does it represent an endorsement by the institutions the reviewers are associated with. The authors and Task Force members would like to thank the following individuals for their participation in the review process and for providing many constructive comments and suggestions:

Juan Carlos Batista, SENASA, Buenos Aires, Argentina.  
Paul Brent, Food Standards Australia New Zealand, Product Standards Program, Canberra, Australia.  
Moises Burachik, Argentine Secretariat of Agriculture, Buenos Aires, Argentina.  
Gary Cromwell, University of Kentucky, Department of Animal Sciences, Lexington, USA.  
Neuza Maria Brunoro Costa, Universidade Federal de Viçosa, Viçosa, Brazil.  
Howard Davies, Scottish Crop Research Institute, Mylnfield, Invergowrie, UK.  
Johanna Dwyer, Tufts-New England Medical Center, Boston, MA, USA.  
Karl-Heinz Engel, Technical University of Munich, Freising-Weihenstephan, Germany.  
Flavio Finardi Filho, University of Sao Paulo, Sao Paulo, Brazil.  
Suzanne S. Harris, International Life Sciences Institute (ILSI), Human Nutrition Institute, Washington, DC, USA.  
HO-il Kim, National Institute of Agricultural Sciences, Suwon, Korea.  
Shirong Jia, Chinese Academy of Agricultural Sciences, Biotechnology Research Institute, Beijing, China.  
David Jonas, Industry Council for Development of the Food & Allied Industries, Ty Glyn Farm, UK.  
Lisa Kelly, Food Standards Australia New Zealand, Product Standards Program, Canberra, Australia.  
Franco Lajolo, University of Sao Paulo, Faculdade de Ciências Farmacêuticas, Sao Paulo, Brazil.  
Sun Hee Park, Korean Food and Drug Administration, Seoul, Korea.  
Wayne Parrott, The University of Georgia, Athens, Georgia, USA.

William Price, US Food and Drug Administration, Center for Veterinary Medicine Rockville, Maryland, USA.

Delia Rodriguez-Amaya, Universidade Estadual del Campinas, Brazil

Sylvia Rowe, SR Strategy, Washington, DC, USA.

Tee E Siong, Cardiovascular, Diabetes and Nutrition Research Center, Institute for Medical Research, Kuala Lumpur, Malaysia.

## References

- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. *Biology of the Cell*. Garland Science, New York, NY.
- Alouf, J., Freer, J., 1999. *The Comprehensive Sourcebook of Bacterial Protein Toxins*, second ed. Academic Press, New York, NY.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Astwood, E.G., 1970. Anterior pituitary hormones and related substances. In: Goodman, L.S., Gilman, A. (Eds.), *The Pharmacological Basis of Therapeutics*, fourth ed. MacMillan Publishing Company, New York, pp. 1538–1565.
- Astwood, J.D., Leach, J.N., Fuchs, R.L., 1996. Stability of food allergens to digestion *in vitro*. *Nature Biotech.* 14, 1269–1273.
- Atherton, K., 2002. Safety assessment of genetically modified crops. *Toxicology*, 421–426.
- Barlow, S.M., Kozianski, G., Wurtzen, G., Schlatter, J., 2001. Threshold of toxicological concern for chemical substances in the diet. A report of a workshop held on 5–6 of October, 1999, in Paris, France organized by ILSI Europe Threshold of Toxicological Concern Task Force. *Food Chem. Toxicol.* 39, 893–905.
- Betz, F., Hammond, B.G., Fuchs, R.L., 2000. Safety and advantages of *Bacillus thuringiensis*-protected plants to control insect pests. *Regul. Toxicol. Pharmacol.* 32, 156–173.
- Beutler, B., Milsark, I.W., Cerami, A.C., 1985. Passive immunization against cachectin/tumor necrosis factor protects mice from lethal effects of endotoxin. *Science* 229, 869–871.
- Broekaert, W.F., Cammue, B.P.A., De Bolle, M.F.C., Thevissen, K., De Samblanx, G.W., Osborn, R.W., 1997. Antimicrobial peptides from plants. *Crit. Rev. Plant Sci.* 16, 297–323.
- Cellini, F., Chesson, A., Colquhoun, I., Constable, A., Davies, H.V., Engel, K.H., Gatehouse, A.M.R., Kärenlampi, S., Kok, E.J., Leguay, J.-J., Lehesranta, S., Noteborn, H.P.J.M., Pedersen, J., Smith, M., 2004. Unintended effects and their detection in genetically modified crops. *Food Chem. Toxicol.* 42, 1089–1125.
- Chassy, B., 2002. Food safety evaluation of crops produced through biotechnology. *J. Am. College Nutr.* 21, 1665–1735.
- Chassy, B., Hlywka, J.J., Kleter, G.A., Kok, E.J., Kuiper, H.A., McGloughlin, M., Munro, I.C., Phipps, R.H., Reid, J.E., 2004. Nutritional and safety assessments of foods and feeds nutritionally improved through biotechnology. *Food Nutr. Bull.* 26, 436–442.
- Cheeke, P.R., Shull, L.R., 1985. *Proteins and amino acids. Natural Toxicants in Feed and Poisonous Plants*. AVI Publishing Co., Westport CT, pp. 235–291.
- Cheeseman, M.A., Machuga, E.J., Bailey, A.B., 1999. A tiered approach to threshold of regulation. *Food Chem. Toxicol.* 37, 387–412.
- 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.
- Collins, V.L., Hackett, J., 1991. Sequence of the phosphomannose isomerase encoding gene of *Salmonella typhimurium*. *Gene* 103, 135–136.
- Conolly, R.B., Lutz, W.K., 2004. Nonmonotonic dose–response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicol. Sci.* 77, 151–157.
- Constable, A., Jonas, D., Cockburn, A., Davi, A., Edwards, G., Hepburn, P., Herouet-Guicheney, C., Knowles, B., Moseley, B., Oberdörfer, R., 2007. ‘History of safe use’ as applied to the safety assessment of novel foods and foods derived from genetically modified organisms. *Food Chem. Toxicol.* 45, 2513–2525.
- Crickmore, N., Zeigler, D.R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Dean, D.H., 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* 62, 807–813.
- Darzins, A., Franz, B., Vanags, R.I., Chakrabarty, A.M., 1986. Nucleotide sequence analysis of the phosphomannose isomerase gene (*pmi*) of *Pseudomonas aeruginosa* and comparison with the corresponding *Escherichia coli* gene *manA*. *Gene* 42, 293–302.
- Day, P.R., 1996. The biology of plants. *CRC Crit. Rev. Food Sci. Nutr.* 36, 39–47.
- De Samblanx, G.W., Goderis, I.J., Thevissen, K., Raemaekers, R., Fant, F., Borremans, F., Acland, D.P., Osborn, R.W., Patel, S., Broekaert, W.F., 1997. Mutational analysis of a plant defensin from radish (*Raphanus sativus* L.) reveals two adjacent sites important for antifungal activity. *J. Biol. Chem.* 271, 1171–1179.
- Devlin, T.M., 2002. *Textbook of Biochemistry with Clinical Correlations*, fifth ed. John Wiley & Sons, New York, NY.
- Doolittle, R.F., 1990. Searching through sequence databases. *Methods Enzymol.* 183, 99–110.
- Duval, A., Malecot, C.O., Pelhate, M., Rochat, H., 1989. Changes in Na channel properties of frog and rat skeletal muscles induced by the AaH II toxin from the scorpion *Androctonus australis*. *Pflugers Arch.* 415, 361–371.
- Eckhoff, S.R., Rausch, K.D., Fox, E.J., Tso, C.C., Wu, X., Pan, Z., Buriak, P., 1993. A laboratory wet milling procedure to increase reproducibility and accuracy of product yields. *Cereal Chem.* 70, 727–732.
- EFSA, 2006a. Guidance document of the Scientific Panel on Genetically Modified Organisms for the risk assessment of genetically modified plants and derived food and feed. *EFSA J.* 9, 1–100.
- EFSA, 2006b. Safety and Nutritional Assessment of GM Plant derived Foods/Feed: The role of animal feeding trials. Draft for public consultation. December 15, 2006. <[http://www.efsa.europa.eu/en/science/gmo/gmo\\_consultations/gmo\\_AnimalFeedingTrials.html](http://www.efsa.europa.eu/en/science/gmo/gmo_consultations/gmo_AnimalFeedingTrials.html)>.
- FAO/WHO, 1991. Strategies for assessing the safety of foods produced by biotechnology. Report of a joint FAO/WHO Expert Consultation. WHO, Geneva.
- FAO/WHO, 1996. *Biotechnology and Food Safety Report*. Report of a joint FAO/WHO Expert Consultation, Rome, Italy, 30 September–4 October, 1996. FAO Food and Nutrition Paper 61. Food and Agriculture Organisation of the United Nations, Rome. <<ftp://ftp.fao.org/es/esn/food/biotechnology.pdf>>.
- Franz, D.R., Jax, N.K., 1997. Ricin toxin. In: Zajtcuk, R., Bellamy, R.F. (Eds.), *Textbook of Military Medicine: Medical aspects of chemical and biological warfare*. Office of The Surgeon General Department of the Army, United States of America, TMM Publications, Washington, DC.
- Frawley, J.P., 1967. Scientific evidence and common sense as a basis for food-packaging regulations. *Food Cosmet. Toxicol.* 5, 293–308.
- Finklestein, R.A., 1973. Cholera. *CRC Crit. Rev. Microbiol.* 2, 553–623.
- Fricker, G., Drewe, J., 1996. Current concepts in intestinal peptide absorption. *J. Pept. Sci.* 2, 195–211.
- Fu, T.-J., Abbott, U.R., Hatzos, C., 2002. Digestibility of food allergens and nonallergenic proteins in simulated gastric fluid and simulated intestinal fluid – a comparative study. *J. Agric. Food Chem.* 50, 7154–7160.
- Gatehouse, A.M., Norton, E., Davison, G.M., Babbe, S.M., Newell, C.A., Gatehouse, J.A., 1999. Digestive proteolytic activity in larvae of tomato moth, *Lacanobia oleracea*; effects of plant protease inhibitors *in vitro* and *in vivo*. *J. Insect Physiol.* 45, 545–5558.

- Gerrior, D., Bente, L., Hiza, H., 2004. Nutrient content of the US food supply, 1909–2000 Home Economics Research Report No. 56. US Department of Agriculture, Center for Nutrition Policy and Promotion.
- Gibson, J., 2006. Bioinformatics of protein allergenicity. *Mol. Nutr. Food Res.* 50 (7), 591–670.
- Gill, M.D., 1982. Bacterial toxins: a table of lethal amounts. *Microbiol. Rev.* 46, 86–94.
- Gill, M.D., 1987a. Bacterial toxins: description. In: Laskin, A.I., Lechevalier, H.A. (Eds.), *CRC Handbook of Microbiology*, second ed., . In: *Toxins and Enzymes*, vol. VIII CRC Press, Boca Raton, FL, pp. 3–18.
- Gill, M.D., 1987b. Bacterial toxins: lethal amounts. In: Laskin, A.I., Lechevalier, H.A. (Eds.), *CRC Handbook of Microbiology*, second ed., . In: *Toxins and Enzymes*, vol. VIII CRC Press, Boca Raton, FL, pp. 127–135.
- Goldsworthy, A., Street, H.E., 1965. The carbohydrate nutrition of tomato roots VIII. The mechanism of the inhibition of D-mannose of the respiration of excised roots. *Ann. Bot.* 29, 45–58.
- 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.
- Harrison, L.A., Bailey, M.R., Naylor, M., Ream, J., Hammond, B., Nida, D.L., Burnette, B., Nickson, T.E., Mitsky, T., Taylor, M.L., Fuchs, R.L., Padgett, S.R., 1996. The expressed protein in glyphosate-tolerance soybean, 5-enolpyruvyl-shikimate-3-phosphate synthase from *Agrobacterium* sp. strain CP4, is rapidly digested *in vitro* and is not toxic to acutely gavaged mice. *J. Nutr.* 126, 728–740.
- Hauschild, A., 1989. *Clostridium botulinum*. In: Doyle, M.P. (Ed.), *Foodborne Bacterial Pathogens*. Marcel Dekker, New York, NY, pp. 111–190.
- Hefle, S.L., Nordlee, J.A., Taylor, S.L., 1996. Allergenic foods. *CRC Crit. Rev. Food Sci. Nutr.* 36, 69–89.
- Herman, R.A., Schafer, B.W., Korjagin, V.A., Ernest, A.D., 2003. Rapid digestion of Cry34Ab1 and Cry35Ab1 in simulated gastric fluid. *J. Agric. Food Chem.* 51, 6823–6827.
- Herman, R.A., Storer, N.P., Gao, Y., 2006. Digestion assays in allergenicity assessment of transgenic proteins. *Environ. Health Persp.* 114, 1154–1157.
- Höfte, H., Whitely, H.R., 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* 53, 242–255.
- Hérouet, C., Esdaile, D., Mallyon, B.A., Debruyne, E., Schulz, A., Hendrickx, K., van der Klis, R.-J., Rouan, D., 2005. Safety evaluation of the phosphinothricin acetyltransferase proteins encoded by the pat and bar sequences that confer tolerance to glufosinate-ammonium herbicide in transgenic plants. *Regul. Toxicol. Pharmacol.* 41, 134–149.
- 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.
- ISAAA (International Service for the acquisition of agri-biotech applications), 2006. <<http://www.isaaa.org/>>.
- Ishiguro, M., Mitarai, M., Harada, H., Sekine, I., Kikutani, M., 1983. Biochemical studies on oral toxicity of ricin. I. Ricin administered orally can impair sugar absorption by the rat small intestine. *Chem. Pharm. Bull.* 31, 3222–3227.
- Kaneda, M., Oyama, Y., Ikemoto, Y., Akaike, N., 1989. Scorpion toxin prolongs an inactivation phase of the voltage-dependent sodium current in rat isolated single hippocampal neurons. *Brain Res.* 487, 192–195.
- Kroes, R., Galli, C., Munro, I., Schilter, B., Tran, L.-A., Walker, R., Wurtzen, G., 2000. Threshold of toxicological concern for chemical substances in the diet: a practical tool for assessing the need for toxicity testing. *Food Chem. Toxicol.* 38, 255–312.
- Kroes, R., Renwick, A.G., Cheeseman, M., Kleiner, J., Mangelsdorff, I., Piersma, A., Schilter, B., Schlatter, J., van Schthorst, F., Vos, J.G., Wurtzen, G., 2004. Structure-based thresholds of toxicological concern (TCC): guidance for applications to substances present at low levels in the diet. *Food Chem. Toxicol.* 42, 65–83.
- Küiper, H.A., Kleter, G.A., Noteborn, H.P.J.M., Kok, E.J., 2002. Substantial equivalence – an appropriate paradigm for the safety assessment of genetically modified foods. *Toxicology*, 427–431.
- Kutzner, H.J., 1981. The family *Streptomycetaceae*. In: Starr, M.P., Stolp, H., Trüper, H.G., Ballows, A., Schlegel, H.G. (Eds.), *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria*. Springer-Verlag, Berlin, pp. 2028–2072.
- 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, 985–998.
- Lee, B.T., Matheson, N.K., 1984. Phosphomannose isomerase and phosphoglucoisomerase in seeds of *Cassia coluteoides* and some other legumes that synthesize galactomannan. *Phytochemistry* 23, 983–987.
- Lee, D., Gurevitz, M., Adams, M.E., 2000. Modification of synaptic transmission and sodium channel inactivation by the insect-selective scorpion toxin Lqh $\alpha$ IT. *J. Neurophysiol.* 83, 1181–1187.
- Liener, I.E., 1994a. Antinutritional factors related to proteins and amino acids. In: Hui, Y.H., Gorham, J.R., Murell, K.D., Cliver, D.O. (Eds.), *Foodborne Disease Handbook. Disease Caused by Hazardous Substances*. Marcel Dekker, New York, NY, pp. 261–309.
- Leiner, I.E., 1994b. Implications of antinutritional components in soybean foods. *CRC Crit. Rev. Food Sci. Nutr.* 34, 31–67.
- Lyerly, D.M., Saum, K.E., MacDonald, D.K., Wilkins, T.D., 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. *Infect Immun.* 47, 349–352.
- Magalhães, M.D.M., Vina, G., Arantes, R.M.E., Santos, T.M., Cunha-Melo, J.R., 1998. The mouse as an experimental model for *Tityus serrulatus* scorpion envenoming. *Acta Cir. Bras.* 13, 260–263.
- Maksymowych, A.B., Reinhard, M., Malizio, C.J., Goodnough, M.C., Johnson, E.A., Simpson, L.L., 1999. Pure botulinum neurotoxin is absorbed from the stomach and small intestine and produces peripheral neuromuscular block. *Infect Immun.* 67, 4708–4712.
- 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. *CRC Crit. Rev. Food Sci. Nutr.* 36, 165–186.
- Miles, S.J., Guest, J.R., 1984. Nucleotide sequence and transcriptional start point of the phosphomannose isomerase gene (*manA*) of *Escherichia coli*. *Gene* 32, 41–48.
- Morgan, D.R., Street, H.E., 1959. The carbohydrate nutrition of tomato roots. VII. Sugars, sugar phosphates and sugar alcohols as respiratory substrates for excised roots. *Ann. Bot.* 23, 89–105.
- Munro, I.C., Kennepohl, E., Kroes, R., 1999. A procedure for the safety evaluation of flavouring substances. *Food Chem. Toxicol.* 37, 207–232.
- Nair, R.S., Fuchs, R.L., Schuette, S.A., 2002. Current methods for assessing safety of genetically modified crops as exemplified by data on Roundup Ready soybeans. *Toxicol. Pathol.* 30, 117–125.
- Nida, D.L., Kolacz, K.H., Buehler, R.E., Deaton, W.R., Schuler, W.R., Armstrong, T.A., Taylor, M.L., Ebert, C.C., Rogan, G.J., 1996. Glyphosate-tolerant cotton: genetic characterization and protein expression. *J. Agric. Food Chem.* 44, 1960–1966.
- Niehues, R., Hasilik, M., Alton, G., Korner, C., Schiebe-Sukumar, M., Koch, H.G., Zimmer, K., Wu, R., Harms, E., Reiter, K., von Figura, K., Freeze, H.H., Harms, H.K., Marquardt, T., 1998. Carbohydrate-deficient glycoprotein syndrome type 1b. *J. Clin. Invest.* 101, 1414–1420.
- OECD, 2002. Guideline for the testing of chemicals: acute oral toxicity. OECD Test Guideline 401. Organization for Economic Co-operation and Development, Paris, France.
- OECD, 1995. Guideline for the testing of chemicals: repeated dose 28-day oral toxicity study in rodents. OECD Test Guideline 407. Organization for Economic Co-operation and Development, Paris, France.
- PFAM, 2005. Protein families database of alignments and HMMs. <[www.sanger.ac.uk/Software/Pfam](http://www.sanger.ac.uk/Software/Pfam)>.
- Padgett, S.R., Kolack, K.H., Delannay, X., Re, D.B., La Vallee, B.J., Tinius, C.N., Rhodes, W.K., Otero, Y.I., Barry, G.F., Eichholtz, D.A.,

- Peschke, Nida, D.L., Taylor, N.B., Kishore, G.M., 1995. Development, identification, and characterization of a glyphosate-tolerant soybean line. *Crop Sci.* 35, 1451–1461.
- Padgett, S.R., Re, D.B., Gasser, C.S., Eichholtz, D.A., Frazier, R.B., Hironaka, C.M., Levine, E.B., Shah, D.M., Fraley, R.T., Kishore, G.M., 1991. Site-directed mutagenesis of a conserved region of the 5-enolpyruvylshikimate-3-phosphate synthase active site. *J. Biol. Chem.* 266, 22364–22369.
- Padgett, S.R., Re, D.B., Barry, G.F., Eichholtz, D.E., Delannay, F.X., Fuchs, R.L., Kishore, G.M., Fraley, R.T., 1996. New weed control opportunities: development of soybeans with a roundup ready gene. In: Duke, S.O. (Ed.), *Herbicide-resistant Crops: Agricultural, Environmental, Economic, Regulatory, and Technical Aspects*. CRC Press, New York, NY, pp. 53–84.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85, 2440–2448.
- Popoff, M.R., 1998. Interactions between bacterial toxins and intestinal cells. *Toxicology* 36, 665–685.
- Proudfoot, A.E.I., Turcatti, G., Wells, T.N.C., Payton, M.A., Smith, D.J., 1994a. Purification, cDNA cloning and heterologous expression of human phosphomannose isomerase. *Eur. J. Biochem.* 219, 415–423.
- Proudfoot, A.E., Payton, M.A., Wells, T.N., 1994b. Purification and characterization of fungal and mammalian phosphomannose isomerases. *J. Protein. Chem.* 13, 619–627.
- Powell-Abel, P., Nelson, R.S., De, B., Hoffmann, N., Rogers, S.G., Fraley, R.T., Beachy, R.N., 1980. Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* 232, 738–743.
- Ratner, B., Untracht, S., Collins-Williams, C., 1952. Allergenicity of modified and processed foodstuffs. I. The use of a dual ingestion passive transfer test to determine the allergenicity of foodstuffs in man. *Ann. Allergy* 10, 675–681.
- Reed, J., Privalle, L., Powell, M.L., Meghji, M., Dawson, J., Dunder, E., Suttie, J., Wenk, A., Launis, K., Kramer, C., Chang, Y., Hansen, G., Wright, M., 2001. Phosphomannose isomerase: an efficient selectable marker for plant transformation. *In vitro cell. Dev. Biol. – Plant* 37, 127–132.
- Roberts, P.R., Burney, J.D., Black, K.W., Zaloga, G.P., 1996. Effect of chain length on absorption of biologically active peptides from the gastrointestinal tract. *Digestion* 60, 332–337.
- Robertus, J.D., 1991. The structure and action of ricin, a cytotoxic N-glycosidase. *Semin. Cell Biol.* 2, 23–30.
- Sakaguchi, G., Ohishi, T., Kozaki, S., 1988. Botulism – structure and chemistry of botulinum toxins. In: Herdegree, M.C., Tu, A.T. (Eds.), *Handbook of Natural Toxins*. Marcel Dekker, New York, NY, pp. 191–216.
- Sanders, P.R., Lee, T.C., Groth, M.E., Astwood, J.D., Fuchs, R.L., 1998. Safety assessment of insect-protected corn. In: Thomas, J.A. (Ed.), *Biotechnology and Safety Assessment*, second ed. Taylor and Francis, Oxford, UK, pp. 241–256.
- Schmidt, M., Arnold, W., Niemann, A., Kleickmann, A., Puhler, A., 1992. The *Rhizobium meliloti pmi* gene encodes a new type of phosphomannose isomerase. *Gene* 122, 35–43.
- Shinabarger, D., Berry, A., May, T.B., Rothmel, R., Fialho, A., Chakrabarty, A.M., 1991. Purification and characterization of phosphomannose isomerase-guanosine diphospho-D-mannose pyrophosphorylase. A bifunctional enzyme in the alginate biosynthetic pathway of *Pseudomonas aeruginosa*. *J. Biol. Chem.* 266, 2080–2088.
- Sidell, F.R., Takafuji, E.T., Franz, D.R., 1997. Part I: Warfare, weaponry, and the casualty. In: Sidell, F.R., Takafuji, E.T., Franz, D.R. (Eds.), *Textbook of Military Medicine*, In: *Medical Aspects of Chemical and Biological Warfare*, vol. 3. Borden Institute, Walter Reed Medical Center, Washington, DC.
- Smith, D.J., Proudfoot, A.E.I., Freidli, L., Klig, L.S., Paravicini, G., Payton, M.A., 1992. PMI40, an intron-containing gene required for early steps in yeast mannosylation. *Mol. Cell. Biol.* 12, 2924–2930.
- Skjelkvale, R., Uemura, T., 1977. Experimental diarrhoea in human volunteers following oral administration of *Clostridium perfringens* enterotoxin. *J. Appl. Bacteriol.* 43, 281–286.
- Sjoblad, R.D., McClintock, J.T., Engler, R., 1992. Toxicological considerations for protein components of biological pesticide products. *Regul. Toxicol. Pharmacol.* 15, 3–9.
- Smolin, L.A., Grosvenor, M.B., 2000. *Nutrition Science & Applications*, third ed. Saunders College Publishing, New York, NY.
- Stallings, W.C., Abdel-Meguid, S.S., Lim, L.W., Shieh, H.S., Dayringer, H.E., Leimgruber, N.K., Stegeman, R.A., Anderson, K.S., Sikorski, J.A., Padgett, S.R., Kishore, G.M., 1991. Structure and topological symmetry of the glyphosate target 5-enolpyruvylshikimate-3-phosphate synthase: a distinctive protein fold. *Proc. Natl. Acad. Sci. USA* 88, 5046–5050.
- Stelid, G., 1954. Toxic effects of D-mannose, 2-desoxy-D-glucose, and D-glucosamine upon respiration and ion absorption in wheat roots. *Physiol. Plant* 7, 173–181.
- Strobel, S., 1998. Oral tolerance: immune responses to food antigens. In: Metcalf, D.D., Sampson, H., Simon, R. (Eds.), *Food Allergy: Adverse Reactions to Foods and Food Additives*. Blackwell Science Publishing, Massachusetts, pp. 107–136.
- Suzuki, K., Tateda, K., Matsumoto, T., Gondaira, F., Tsujimoto, S., Yamaguchi, K., 2000. Effects of interaction between *Escherichia coli* verotoxin and lipopolysaccharide on cytokine induction and lethality in mice. *J. Med. Microbiol.* 49, 905–910.
- Terras, F.R., Schoofs, H.M., De Bolle, M.F.C., Van Leuven, F., Rees, S.B., Vanderleyden, J., Cammue, B.P.A., Broekaert, W.F., 1992. Analysis of two novel classes of plant antifungal proteins from radish (*Raphanus sativus* L.) seeds. *J. Biol. Chem.* 267, 15301–15309.
- Thevissen, K., Ghazi, A., De Samblanx, G.W., Brownlee, C., Osborn, R.W., Broekaert, W.F., 1996. Fungal membrane responses induced by plant defensins and thionins. *J. Biol. Chem.* 271, 15018–15025.
- Thomas, K., Aalbers, M., Bannon, G.A., Bartels, M., Dearman, R.J., Esdaile, D.J., Fu, T.J., Glatt, C.M., Hadfield, N., Hatzos, C., Hefle, S.L., Heylings, J.R., Goodman, R.E., Henry, B., Herouet, C., Holsapple, M., Ladics, G.S., Landry, T.D., MacIntosh, S.C., Rice, E.A., Privalle, L.S., Steiner, H.Y., Teshima, R., van Ree, R., Woolhiser, M., Zawodny, J., 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.
- Thomas, K., Bannon, G., Hefle, S., Herouet, C., Holsapple, M., Ladics, G., MacIntosh, S., Privalle, L., 2005. In silico methods for evaluating human allergenicity to novel proteins: Bioinformatics Workshop Meeting Report, 23–24 February 2005. *Toxicol. Sci.* 88, 307–310.
- Thomas, K., Herouet-Guicheney, C., Ladics, G., Bannon, G., Cockburn, A., Crevel, R., Fitzpatrick, J., Mills, C., Privalle, L., Vieths, S., 2007. Evaluating the effects of food processing on the potential human allergenicity of novel proteins: international workshop report. *Food Chem. Toxicol.* 45, 1116–1122.
- Thompson, C.J., Movva, N.R., Tizard, R., Crameri, R., Davies, J.E., Lauwereys, M., Botterman, J., 1987. Characterization of the herbicide-resistance gene bar from *Streptomyces hygroscopicus*. *EMBO J.* 6, 2519–2523.
- Tsume, Y., Taki, Y., Sakane, T., Nadai, T., Sezaki, H., Watabe, K., Kohno, T., Yamashita, S., 1996. Quantitative evaluation of the gastrointestinal absorption of protein into the blood and lymph circulation. *Biol. Pharm. Bull.* 19, 1332–1337.
- Uniprot-Swissprot Consortium, 2007. The Universal Protein Resource (UniProt). *Nucleic Acids Research*, vol. 35, Database issue D193–D197. <[http://nar.oxfordjournals.org/cgi/reprint/35/suppl\\_1/D193.pdf](http://nar.oxfordjournals.org/cgi/reprint/35/suppl_1/D193.pdf)>. doi:10.1093/nar/gkl1929, database online at <http://www.ebi.ac.uk/swissprot/>.
- US FDA, 1983. *Toxicological principles for the safety assessment of direct food additives and color additives used in food*. Redbook. US FDA, Bureau of Foods, Washington DC.
- US EPA, 1998. *Health effects test guidelines. OPPTS 870.1100 Acute oral toxicity*. <[http://www.epa.gov/opptsfrs/publications/OPPTS\\_Harmonized/870\\_Health\\_Effects\\_Test\\_Guidelines/Series/870-1100.pdf](http://www.epa.gov/opptsfrs/publications/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Series/870-1100.pdf)>.

- US EPA, 1996. EPA Fact Sheet for *Bacillus thuringiensis* subspecies *kurstaki* Cry 1A (b) Delta Endotoxin and Its Controlling Sequences as Expressed in Corn. <<http://www.epa.gov/pesticides/biopesticides/ingredients>>.
- US EPA, 2000. Background document for the FIFRA scientific advisory panel on mammalian toxicity assessment guidance for protein plant pesticides. <<http://www.epa.gov/scipoly/sap/2000/june/mammaltox.pdf>>.
- USP, 2000. The United States Pharmacopeia 24. Simulated gastric fluid, TS. In: Board of Trustees (Eds.), The National Formulary 19. United States Pharmacopeial Convention, Inc., Rockville, MD, pp. 2235.
- Van Haver, E., DeSchrijver, A., Devos, Y., Lievens, S., Renckens, S., Moens, W., 2003. Guidance notes for the safety assessment of genetically modified crops for food and feed use. Report of the Belgium Biosafety Advisory Council, Scientific Institute of Public Health. Royal Library of Belgium Deposit No. D/2003/2505/16.
- Van Rie, J., Jansens, S., Höfte, H., Degheele, D., Van Mellaert, H., 1989. Specificity of *Bacillus thuringiensis*  $\delta$ -endotoxins. Eur. J. Biochem. 186, 239–247.
- Wang, G.K., Strichartz, G., 1982. Simultaneous modifications of sodium channel gating by two scorpion toxins. Biophys. J. 40, 175–179.
- Webb, K.E., 1990. Intestinal absorption of protein hydrolysis products: a review. J. Anim. Sci. 68, 3011–3022.
- Wehrmann, A., Van Vliet, A., Opsomer, C., Botterman, J., Schul, A., 1996. The similarities of *bar* and *pat* gene products make them equally applicable for plant engineers. Nature Biotechnol. 14, 1274–1278.
- Zhuang, M., Gill, S.S., 2002. Mode of action of *Bacillus thuringiensis* toxins. In: Voss, G., Ramos, G. (Eds.), Chemistry of Crop Protection, . In: Progress and Prospects in Science and Regulation. Wiley & Sons, New York, NY, pp. 213–236.