White Paper on Tier-Based Testing for the Effects of Proteinaceous Insecticidal Plant-Incorporated Protectants on Non-Target Arthropods for Regulatory Risk Assessments

Biotechnology Regulatory Services

Biopesticides and Pollution Prevention Division

April 6, 2007

TABLE OF CONTENTS

EDITOR

Robyn I. Rose, United States Department of Agriculture, Animal Plant Health Inspection Service, Biotechnology Regulatory Services

Dr. Robyn Rose currently works at the United States Department of Agriculture's (USDA) Animal and Plant Health Inspection Service (APHIS) and is a previous employee of the U.S. Environmental Protection Agency (EPA) Office of Pesticide Programs (OPP). Thus, Dr. Rose is familiar with the evaluation of proteinaceous, insecticidal plant-incorporated protectants (PIP), particularly those derived from *Bacillus thuringiensis* from the perspective of both agencies. This paper is not EPA or USDA guidance or policy.

REVIEWERS

A panel of experts was convened on December 15 and 16, 2004 to review a draft of this White Paper and provide expertise and critical input. Comments were considered and incorporated when appropriate. The following is a list of the people who reviewed this White Paper:

Joel Coats, Iowa State University
Stephen Cox, Texas Tech University
Galen Dively, University of Maryland
Brian Federici, University of California at Riverside
Cecile Girard, Canadian Food Inspection Agency
Graham Head, Monsanto Company
Richard Hellmich, USDA, Agricultural Research Service, Iowa State University
Rodney Herman, Dow AgroSciences
Neil Hoffman, USDA, Animal Plant Health Inspection Service, Biotechnology Regulatory Services
Philip Hutton, EPA, Office of Pesticide Programs, Biopesticides and Pollution Prevention Division
Paul Jepson, Oregon State University
Susan Koehler, USDA, Animal Plant Health Inspection Service, Biotechnology Regulatory Services
Raymond Layton, Pioneer HiBred
Keith Matthews, EPA, Office of General Council
Thomas McClintock, EPA, Office of Science Policy and Coordination
William Moar, Auburn University
Joerg Romeis, Biosafety Group, Agroscope Reckenholz-Tänikon Research Station (ART)
Alan Raybould, Syngenta
Robyn Rose, USDA, Animal Plant Health Inspection Service, Biotechnology Regulatory Services
Richard Roush, University of California at Davis
Anthony Shelton, Cornell University

L. Carmen Soileau, USDA, Animal Plant Health Inspection Service, Biotechnology Regulatory Services
Debra Teixera, Springborn Smithers Laboratories
Gail Tomimatsu, EPA, Office of Pesticide Programs, Biopesticides and Pollution Prevention Division
Zigfridas Vaituzis, EPA, Office of Pesticide Programs, Biopesticides and Pollution Prevention Division
Jefferey Wolt, Iowa State University

INTRODUCTION

The purpose of this paper is to summarize scientific and regulatory information on non-target invertebrate testing and evaluation of new products expressing insecticidal proteins in an ecological risk assessment. The information included provides a general discussion for conducting non-target invertebrate ecological risk assessments for crops expressing insecticidal proteins. This paper is not intended as official policy or guidance; it is unknown at this time if it will lead to future guidance.

This document focuses on plants expressing insecticidal proteins because they are currently available in the commercial market and are evaluated for potential non-target invertebrate risks. The document does not address other plants expressing pesticidal traits, such as viral coat proteins, because they have a different mode of action than insecticidal proteins. Additionally, the document does not address genetically engineered (GE) crops expressing insecticidal traits that are non-proteinaceous because USDA and EPA have not evaluated any of these products. However, the discussion below may be relevant in the future for non-protenaceous insecticidal traits.

Currently, both the United States Department of Agriculture (USDA) and Environmental Protection Agency (EPA) conduct risk assessments of plants expressing insecticidal proteins as part of their regulatory process (see EPA-OPP and USDA-APHIS Regulatory section below). Since both USDA and EPA review similar non-target invertebrate data as part of the ecological risk assessment, these agencies coordinated to develop this White Paper discussing the tiered approach to risk assessment.

Plants expressing insecticidal traits may be divided into two categories: proteinaceous and non-proteinaceous. Proteinaceous insecticidal traits include those that are currently available commercially to prevent the damage caused by insects and are based on genes derived from the common bacterium *Bacillus thuringiensis* (*Bt*). *Bt* genes are considered to be "reduced risk" pesticides (http://www.epa.gov/pesticides/health/reducing.htm) and, as such, they constitute alternatives to broad spectrum organophosphate insecticides. Currently, there are no non-proteinaceous insecticidal traits being considered by the USDA or EPA for commercial use for protection against insects.

Ecological risk assessment is the process by which regulatory authorities, such as the USDA Animal and Plant Health Inspection Service (APHIS) and EPA Office of Pesticide Programs (OPP), use scientific data on potential hazards and exposure to assess the likelihood of adverse impacts on populations and communities of organisms in the environment. The ecological hazard associated with transgenic crops is related to the potential toxicity of the insecticidal protein to non-target beneficial organisms. The exposure assessment predicts the likelihood that non-target organisms will have a dietary exposure to the expressed protein at or above a hazard threshold level. The ecological risk is determined by the interaction between toxicity and exposure that arises.

This document is intended to offer support to the non-target arthropod ecological risk assessment scheme for insecticidal proteins, including detailed characterization of exposure and risk as described in the EPA's Ecological Risk Assessment Framework document (EPA, 1998). The paper addresses laboratory testing of select, representative

non-target species and, when warranted, appropriate semi-field and field testing. This testing is conducted to determine if there is a hazard associated with the insecticidal trait that could pose an unacceptable risk to non-target arthropods. The risk assessment process addressed in this paper includes, but is not limited to, the tiered approach to testing, appropriate indicator organisms and appropriate endpoints.

EPA-OPP AND USDA-APHIS REGULATORY AUTHORITY

In 1986, the Office of Science and Technology Policy published the Coordinated Framework for Regulation of Biotechnology (51 Federal Register 23303; June 26, 1986), which provides a regulatory framework intended to ensure the safety of biotechnology products using existing statutory authority. Agricultural products produced by biotechnology are regulated by the United States Department of Agriculture, Animal and Plant Health Inspection Service (USDA-APHIS), the Environmental Protection Agency, Office of Pesticide Programs (EPA-OPP), and the Department of Health and Human Services Food and Drug Administration (FDA). The FDA assesses whether a GE food or feed crop is as safe for consumption by humans and animals as its traditionally-bred counterpart, but does not conduct environmental risk assessments (ERA). The responsibilities of USDA-APHIS and EPA-OPP as they relate to non-target arthropod testing and risk assessment are described below.

USDA-APHIS

USDA-APHIS is responsible for protecting the United States' animal and plant resources from agricultural pests and diseases. Under the authority of the Plant Protection Act (June 20, 2000), APHIS regulations (7 CFR 340) provide procedures for obtaining a permit or for submitting a notification, prior to "introducing" a regulated article in the United States. A genetically engineered organism is considered a regulated article if the donor organism, recipient organism, vector or vector agent used in engineering the organism belongs to one of the taxonomic groups listed in the regulation and is also a plant pest, or if there is a reason to believe it is a plant pest. The act of introduction includes any movement into (import) or through (interstate) the United States, or release into the environment outside an area of physical confinement. The regulations also provide for petitions for the determination of nonregulated status. Once a determination of nonregulated status is granted, the product (and its offspring) no longer requires APHIS review for movement or release in the United States. Transgenic plants that have been genetically engineered to express insecticidal proteins are considered regulated articles by APHIS until they are granted non-regulated status through the petition process.

APHIS regulations part 7 CFR 340.6 (c)(4) describe the types of data and information that a developer must submit in support of a petition for nonregulated status. In part, these specifically include under a description of "*known and potential differences from the unmodified recipient organism that would substantiate that the regulated article is unlikely to pose a greater plant pest risk than the unmodified organism from which it was derived*," including effects of the regulated article on non-target organisms and indirect plant pest effects on other agricultural products and, under 7 CFR 340.6 (c) (5), data reports from field trials conducted under APHIS permit or notification that shall include

"methods of observation, resulting data, and analysis regarding all deleterious effects on plants, nontarget organisms, or the environment."

In 1996, APHIS published a *Guide for Preparing and Submitting Petitions for Genetically-Engineered Plants* (available at: http://www.aphis.usda.gov/brs/pdf/usergen8.pdf) that addresses the types of data to be submitted in support of petitions, including guidance regarding plants engineered with new pesticidal phenotypes (Section VII, Adverse Consequences of New Cultivar Introduction). To improve consistency of reviews and data submitted for unconfined releases, representatives of the USDA-APHIS, the EPA-OPP and the Canadian Food Inspection Agency (CFIA) Plant Biosafety Office met in September 2000 and produced a bilateral document that describes the commonalities and differences in environmental data considered for the environmental assessment of unconfined releases of transgenic plants, including data to assess effects on non-target organisms (http://www.aphis.usda.gov/brs/canadian/appenannex2e.pdf). The concepts in this document provided the basis for the development of *Unconfined Release into the Environment* by the North American Plant Protection Organization (NAPPO) in Standard RSPM No. 14 on the Importation and Release (into the environment) of Transgenic Plants in NAPPO Member Countries (U.S., Canada, and Mexico) (http://www.nappo.org/Standards/NEW/RSPM14-UpdatedOct03-e.pdf).

Consistent among these documents is the importance of field test data and observations on pest susceptibility, data and information characterizing the levels of toxicants and anti-nutrients normally expressed in the plant, and data characterizing new pesticidal component and its mode of action. Also of importance for regulated articles that contain insecticidal proteins are data concerning the specificity of the toxin, its routes of exposure, and the effects from the introduction of the transgenic plants relative to their non-transgenic counterparts on representative non-target organisms. The information required by EPA is often submitted to APHIS to assess non-target effects, including any that might arise with federally listed threatened and endangered species, in Environmental Assessments (EA) prepared by APHIS in response to petitions for deregulation in compliance with the National Environmental Policy Act (NEPA) of 1969 (42 U.S.C. 4321 *et seq.)* and the pursuant implementing regulations (40 CFR 1500-1508, 7 CFR Part lb; 7 CFR Part 372). These EAs can be viewed on the APHIS website (http://www.aphis.usda.gov/brs/not_reg.html).

EPA

EPA is responsible for regulating the sale, distribution, and use of pesticides in order to protect human health and the environment. If a GE plant produces a substance that is intended to be used for "preventing, destroying, repelling or mitigating any pest," the substance and the genetic material necessary to produce the substance are designated as pesticides under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA). The term plant-incorporated protectant (PIP) was designated by the EPA to describe substances that plants produce to protect them from damage caused by arthropod pests and diseases, although EPA only registers PIPs produced through genetic engineering. A PIP is defined as "*the pesticidal substance that is produced in a plant and the genetic material necessary for the production of that substance*" (40 C.F.R. § 174.3). Since PIPs

are pesticidal substances that plants produce from genetic material that is native to the plant or has been intentionally introduced into the plant's genome, EPA regulates the protein, in the case of proteinaceous PIPS, and the genetic material necessary to produce it, but not the plant itself (http://www.epa.gov/pesticides/biopesticides/whatarebiopesticides.htm).

FIFRA section 3(a) requires, with some exceptions, that a pesticide be registered under the Act prior to distribution or sale in the United States. To register a pesticide, EPA evaluates the proposed pesticide to ensure that its use will not pose an unreasonable risk to human health or the environment. Under FIFRA section 5, EPA issues Experimental Use Permits (EUP) for research conducted on more than ten acres to allow prospective registrants to generate information or data necessary to register a pesticide. In addition, the Federal Food, Drug, and Cosmetic Act (FFDCA) authorizes EPA to establish tolerances (maximum limits) or exemptions from the requirement of a tolerance for residues of pesticides in food. Under FFDCA, food that contains pesticide residues is considered adulterated and subject to seizure by FDA unless EPA has issued a tolerance exemption or a tolerance (and the residue is within the limits). Regulatory requirements, criteria, and procedures applicable to PIPs are outlined in 40 CFR 174 and 40 CFR 152 (66 Federal Register 37772: Regulations under the Federal Insecticide, Fungicide, and Rodenticide Act for Plant-Incorporated Protectants (Formerly Plant-Pesticides); July 19, 2001).

According to FIFRA Section 3, guidelines should be developed that specify the kinds of information required to support the registration of a pesticide. Standards should be commensurate with the anticipated extent of use, pattern of use, and the level and degree of potential beneficial or adverse effects on humans and the environment. If additional data are required to maintain an existing registration of a pesticide, all existing registrants of the pesticide should be notified. PIPs that are sold and distributed for planting by farmers have conditions on the registration that must be followed such as areas where Bt cotton cannot be planted.

In the United States, the EPA has provided the basis for developing ecological risk assessments for safety determinations for PIPs. An EPA risk assessment is based on a risk *vs.* benefit analysis and considers the potential for unreasonable adverse effects to occur. Data requirements for the commercial registration of a PIP by EPA are similar to those for microbially-derived pesticides because specific requirements for GE crops have not yet been developed. Regulatory guidance for ecological risk assessment of products containing PIPs is governed principally by the 1982 Subdivision M guidelines for microbial pesticides, as amended in 1989. Additional regulation of PIPs was implemented through rulemaking (FR, 2001); however, there were no changes proposed in the guidance for the types of tests to be conducted or in the way in which the tests were to be conducted. In 1996, the EPA's Office of Prevention, Pesticides and Toxic Substances (OPPTS) published the 885 Series *Microbial Pesticide Test Guidelines* consistent with the original Subdivision M guidance (http://www.epa.gov/pesticides/biopesticides/regtools/guidelines/microbial_gdlns.htm). EPA addresses environmental risk assessments by a tiered testing scheme in the EPA's Microbial Testing Guidelines Overview (OPPTS 885.0001) (EPA 1996).

The microbial pesticide test guidelines are based on toxic and pathogenic endpoints and do not consider sublethal effects.

NON-TARGET INVERTEBRATE TESTING WITHIN THE CONTEXT OF REGULATORY REVIEW OF PIPS

Ecological Risk Assessment for Invertebrate Non-target Organisms

An ecological risk assessment in a regulatory context is conducted to facilitate decision-making with regard to identifying potential undesirable impacts and options for mitigating them. A risk assessment is a synthesis of sufficient information to determine whether the risks of a proposed course of action are acceptable. To minimize data requirements and avoid unnecessary tests, risk assessments are structured such that risk is determined first from estimates of hazard and exposure under "worst-case" conditions. A lack of adverse effects under these conditions may provide enough confidence that there is no risk and no further data would be needed. Hence, tests conducted early in an investigation tend to be broad in scope but relatively simple in design, and can be used to demonstrate acceptable risk under most conceivable conditions. Later tests are often designed to assess risk under more realistic conditions particularly when early studies suggest potentially unacceptable risk. These later tests are more complex than earlier studies, and conclusions may be limited to the particular conditions of the test. Use of this "tiered" framework saves valuable time and resources by organizing the studies in a cohesive and coherent manner and eliminating unnecessary lines of investigation. A pre-defined and agreed upon framework also provides a common basis for evaluating and comparing risks, and serves to develop a common language among regulators, the regulated and the scientific community.

Assessing the potential for ecological or plant pest risks from plants containing insecticidal proteins involves estimating the likelihood that the presence of the gene will have adverse effects on the environment. A risk exists if the presence of the inserted gene has hazardous effects on non-target organisms and these organisms are exposed to the insectidal protein. Estimates of hazard and exposure allow regulatory agencies to determine the likelihood that an insecticidal protein may cause a problem and also to determine the scale of that problem. Acquisition of relevant hazard and exposure data, therefore, is a fundamental objective of any testing program. How to acquire the data will depend on the problem formulation. The problem formulation is a preliminary assessment in which an analysis plan is developed that considers such things as the risk management goals, relevant and available risk information, and the technical tools that are available. For crops genetically engineered to express insectidal proteins, there are several components of the preliminary assessment that are particularly important:

1. Overall management goals and assessment endpoints
2. Hazard identification
3. Exposure identification
4. Test endpoints
5. Iterative or tiered approach
6. Scientific and public concerns

Overall Management Goals and Assessment Endpoints

The overall purpose of an environmental risk assessment is to determine the likelihood and magnitude of adverse effects on the environment stemming from a particular action, and hence to help make decisions and manage risk. Environmental legislation defines broad management goals. For example, the purpose of FIFRA is to protect the public health and the environment from the misuse of pesticides. For risk assessment purposes, it is necessary to derive explicit expressions of the environmental value to be protected (U.S. EPA 1998) and also to define these expressions in a scientifically operational way. These expressions are called assessment endpoints, and comprise an entity and some property of that entity (Newman, 1998), such as the population size of a certain species or the concentration of a particular chemical in groundwater.

The assessment endpoint of relevance to plants expressing insecticidal proteins is the abundance and diversity of non-target organisms in and near fields of these crops. The risk assessment seeks to estimate the likelihood of an impact on non-target invertebrate abundance (e.g., number of organisms) through measurements of the hazard (toxicity) of the protein and estimates of the environmental exposure to the protein. An acceptable risk is typically set as a measure of hazard (such as the NOEC or LC_{50}) occurring at or above a certain concentration of the insecticidal protein. The acceptability of the risk is not defined by the risk assessment, but an operational definition of acceptable risk is essential to determine when sufficient data have been collected. For example, if a protein were fed to a test organism at a concentration at least ten times the estimated environmental concentration and no effects were observed, the risk might be deemed acceptable.

Hazard Identification for Insecticidal Proteins

Hazard identification involves the evaluation of available information on the insecticidal protein to predict the mode of action, potential toxic effects (e.g., mortality and/or delayed development), and the identity of potentially sensitive species. This process helps with the design of initial screening tests, including the choice of test species, the life history stage, the duration of exposure to the protein, and test endpoints. For example, much information is available on the mode of action for *Bt* proteins. The specific mechanism can vary considerably among *Bt* proteins, but the organism-level effects in susceptible species have been predominantly identified as gut toxicity resulting in reduced growth and development, increased mortality, or both. Sometimes particular life-stages are known to be less susceptible than others. Overall, the high specificity of *Bt* proteins can result in fewer sensitive species compared to a broad-spectrum

insecticide. However, more species may need to be tested if broad-spectrum insecticidal traits are developed and suitable surrogate species are not available. All of this information should be integrated to determine which groups of organisms are likely to be exposed and susceptible to an insecticidal protein.

Exposure Identification

Exposure identification is a comprehensive evaluation of data and information regarding the source of the introduced materials (e.g., taxonomy, physiology, ecology and genetics) and predicted exposure to non-target organisms. Important information includes the source, frequency, intensity, and duration of exposure. For insecticidal proteins, the source of exposure will mainly be ingestion of plant tissues that express the protein (e.g., green tissue, pollen, roots, or seeds) or through movement of protein into the soil, and to a generally lesser degree, from consumption of exposed prey. Because different types of promoters, as well as other factors, can yield different concentrations of the insecticidal protein in various tissues over the course of the plant's life cycle, information on expression is important in evaluating potential exposure. Exposure identification for a Bt protein, for example, would require data on expression in different parts of the plant to which non-target arthropods may be exposed, and at different phases of the growth (e.g., young *versus* senescing leaves). If no expression occurred in pollen, then no exposure to pollen feeders would be expected by this route. The information from the exposure identification phase can be used to develop a list of potentially exposed organisms.

Test Endpoints

Test endpoints are measures of response in a test and are related to the assessment endpoint by a conceptual or empirical model. Relative to regulatory practices, several key factors are considered in deriving test endpoints: the surrogate concept, exposure duration, and standardized endpoints. The surrogate concept in ecological risk assessment involves the selection of indicator organisms to represent a group of taxonomically or functionally related species. The concept of selecting indicator organisms evolved because it is not practical or possible to test all components of an ecosystem. For example, Collembola and earthworms have been commonly used as indicators for soil-inhabiting organisms. An uncertainty factor is often used in extrapolating from the indicator organism response to the level used for environmental regulation because the indicator may not be the most sensitive member of the group. The uncertainty factor ensures that the indicator organism test is conservatively representative of the relevant group of non-target organisms.

In assessing the appropriate duration of toxicity tests, it is customary to separately consider acute and chronic tests that describe effects arising from short- and long-term exposures, respectively (see 'duration of exposure' section below). Initial phases of the risk assessment tend to focus on acute or sub-acute studies, conducted at very high exposure concentrations. If no effects are observed in these studies at levels with a specified margin of safety, then minimal risk is indicated and more detailed risk assessment data may not be triggered for conventional, biochemical, and microbial pesticides. The uncertainty factor used in assessing risk from an acute or sub-acute

assessment will include a factor for species extrapolation and for "acute to chronic" extrapolation.

Finally, the endpoint that is selected is important. Standardized endpoints are needed because they can be replicated between laboratories and the relevance to assessment endpoints is clear. For this reason, endpoints for acute and sub-acute assessments typically emphasize mortality (e.g., derivation of the LC_{50}) and secondarily consider growth, and if the species life history is short in duration, achievement of particular growth stages such as pupation.

Iterative or Tiered Approach

The use of a common testing framework allows comparison between data collected at different times or locations, or with different test materials. In the case of pesticides, this framework consists of different "tiers" of tests. Tiers are used as screening tools that distill and focus risk considerations. Lower-tier studies, such as Tiers I and II in EPA's OPPTS Guidelines, represent worst-case exposure scenarios that incorporate a safety factor and relatively simple testing procedures. Tier I testing of non-target invertebrates is intended to evaluate potential toxic effects on representative non-target organisms (e.g., honey bees, parasitoids and predators). In lower tiers, realism in terms of exposure pathway or level is usually relatively low; organisms are grown and tested in the laboratory using artificial diets or plant material in small containers. However, lower tiers also allow tighter control over experimental variables and exposure conditions, resulting in a greater ability to produce statistically reliable results at relatively low cost.

If an effect is observed in a Tier I study, this suggests a need to more deeply analyze and characterize the existing data, or to conduct higher-tier studies. Conversely, if no adverse effects are seen in Tier I studies with a certain species, then additional testing of that species, or those taxa or processes that it represents, may not be needed. It is possible that, if "worst case" conditions apply, Tier I studies should be regarded as sufficient to demonstrate acceptable risk.

Higher-tier studies encompass more complex laboratory, greenhouse, semi-field or field studies, and can more realistically capture potential exposure to the insecticidal protein in the field should lower-tier, worst-case studies reveal unacceptable toxicity under simpler conditions. Testing is usually done on a larger temporal and/or spatial scale than in the lower tier studies. Adequate replication may be difficult to achieve, extraneous variables may be difficult to control, and appropriate control treatments may be difficult to devise, leading to the possibility that no differences were detected because of lack of statistical power, rather than because the insecticidal protein had no effect. In addition, sampling and analysis can be labor intensive and expensive. For these reasons, higher-tier studies, and field studies in particular, are often carried out only when justified by the results of lower tier tests and when appropriate test organisms and methodologies are available.

Movement from lower to higher tiers is driven by the need and the ability to test hypotheses resulting from the problem formulation and by the need for additional data to satisfy regulatory requirements. The decision to move to a higher tier of testing is first and foremost determined by the results of lower tier tests. The underlying principles of

the tiered testing procedure that drive transitions to higher tiers include: (i) the degree of uncertainty associated with test results or the concern associated with a specific environmental hazard, (ii) the ability to mitigate the concern through product management and deployment, and (iii) the scientific ability and existence of resources (time, money, and expertise) to conduct the higher-tier study. Some hypotheses can be tested quickly by conducting relatively simple assays, while other hypotheses can probably never be tested because of the lack of resources or technical limitations. The goal in designing an appropriate testing framework is to satisfy these risk assessment needs within the bounds of current capability and without unnecessary regulatory burden.

For plants expressing insecticidal proteins, the following tiered study framework for non-target arthropod testing is proposed:

Tier I — Laboratory tests of selected non-target species using exposure levels representing at least 10x the highest Expected Environmental Concentration (EEC). Insecticidal protein mixed with artificial diet is the preferred test compound. The most appropriate protein is similar to that expressed *in planta.* The protein is often produced microbially and may be in an activated form where *in planta* data indicate this to be appropriate.

Tier II — Laboratory tests using plant material alone or mixed with artificial diet. For example, arthropods may be exposed to the pesticidal protein using pollen or leaf discs from the genetically modified plant. Because plant materials are used, exposure levels generally reflect 1x the EEC.

Tier III — Long-term laboratory and/or semi-field tests. Examples of long-term laboratory tests include full life-cycle tests and "tri-trophic tests." Extended laboratory or semi-field tests may be conducted under greenhouse conditions. Controlled semi-field tests employ cages or other techniques to provide some measure of experimental control under simulated or restricted field conditions.

Tier IV — Field tests. These tests may use plots of a hectare or more, which can be distributed across an area or region. Studies may involve looking at specific groups of sentinel organisms or census studies.

Tiered approaches are commonly used by regulatory agencies to efficiently make decisions on low-risk scenarios. For example, a well-defined tiered testing scheme is used within EPA's OPP for the regulation of conventional and biological crop protection chemicals in the areas of ecological and human health risk assessments. In each case, lower tiers represent worst-case exposure scenarios, while scenarios in higher tiers are more refined and realistic. Lower-tier tests conducted in the laboratory are designed to be conservative in nature. Passing lower tiers may indicate to a reviewer that there is little or no risk of acute toxicity to the tested organisms or to related organisms. By extension, and because of the exposure concentrations used, the lower-tier results also may indicate that long-term exposure (and thus, chronic testing) is not necessary to further characterize risk. Failing lower tiers does not necessarily indicate the presence of

an unacceptable risk in the field, but it may trigger the consideration to collect additional exposure and/or effects information to evaluate if the risk is acceptable or unacceptable. Higher-tier tests better represent environmental realism in terms of exposure scenarios and lower safety factors are generally used when evaluating results from such tests because the test endpoint and assessment endpoint are more closely linked than is the case with lower-tier studies.

There are pros and cons associated with the use of different tiers in evaluating potential non-target effects of insecticidal proteins expressed in crops. Laboratory tests can be conducted efficiently within a more controlled, rigorous environment than field tests. Test conditions and the health of organisms are easily regulated in the laboratory, and information can be gathered on cause and effect relationships. However, laboratory tests often assess mortality after short-term or partial exposure rather than the longer-term exposure that may be more representative of field conditions. Although Tier I tests are useful in an ecological risk assessment of insecticidal proteins, laboratory tests may not fully represent the level or form of harm that may occur in the field (EPA-SAP 2002). Proper consideration of Tier I results within the context of reasonably anticipated exposures can help to determine whether further evaluation of the harm that may occur in the field could be determined through higher-tier testing.

An established tiered framework for insecticidal proteins expressed in crops is useful because it serves to provide a common language for studies; it potentially eliminates studies that are not needed; it recommends studies that may provide additional information; and it allows comparisons among data sets gathered over time and space, and with different pest control technologies. The tiered testing framework is intended to be flexible to address changing testing and assessment needs on a case-specific basis. Thus, a tiered framework offers many advantages in terms of testing for potential effects of crops expressing insecticidal proteins on non-target invertebrates.

TIER I TESTING OF NON-TARGET INVERTEBRATES

Selection of Indicator Organisms

Tier I testing is the first step in an ecological risk assessment of pesticides. The surrogate concept using indicator organisms is a fundamental part of hazard testing within an ecological risk assessment. Testing of all potentially exposed invertebrates will never be possible and thus the selection of appropriate indicator species for laboratory testing is very important. Logical consideration and defined criteria may be useful when identifying appropriate individual indicator species (e.g., Dutton *et al.* 2003).

Non-target invertebrates identified for testing may be representative of those that are important in the crop of interest. It is impossible to test all species that are potentially present and the subset selected may represent different habitats (e.g., below the soil surface, soil surface, plant canopy), ecological functions (e.g., predator, parasite or decomposer), and taxonomic groups (e.g., relationship to the target pest). The species tested may be chosen based upon their ecological and economic importance, and their consistent performance in Tier I tests.

The identification of a set of relevant test systems for a given crop expressing an insecticidal protein is first directed by the information that is gathered during product development. Usually, a series of laboratory screening tests are conducted to identify insecticidal activity against target species. Further tests are typically performed to investigate the spectrum of activity of the protein against species closely related to the target pest, as well as against other insect orders. Information from these studies can guide the selection of relevant surrogates for non-target ecological effects testing. For example, if a *Bt* protein is discovered to have insecticidal activity against the western corn rootworm (*Diabrotica virgifera virgifera* LeConte), data typically will be gathered on the activity of that protein against closely related species (e.g., northern corn rootworm, southern corn rootworm). This would suggest that some representative non-target Coleoptera are appropriate candidates for Tier I testing. Demonstration of insecticidal activity in specific taxa therefore allows a more focused and logical approach to regulatory non-target arthropod testing.

For each proposed test organism, the following criteria are relevant:

- *The mode of action and specificity of the insecticidal protein and the impact of that protein on non-target species closely related to the target pest.* In the rootworm-specific example, these might include lady beetles and ground beetles that occur in the same taxonomic order (Coleoptera) as the target species, or beetle species, belonging to the same family (Chrysomelidae) as the target pest.

- *Exposure based on habitat and field abundance.* Where possible, test organisms should represent species that are abundant in the crop and have a relevant route of exposure to the insecticidal protein. For example, testing insidious flower bugs, lady beetles and ground beetles are relevant for insecticidal proteins produced in corn because these insects are abundant and important predators in cornfields. However, species exposed to the insecticidal protein are not always appropriate choices for indicator organism because it is also important to consider practicality of testing in the laboratory (see bullets below). It is critical that surrogate species provide relevant data on the hazard to exposed species.

- *Ecological and taxonomic diversity*. Test organisms may include a broad range of arthropods, particularly economically or socially beneficial species, that represent diverse habitats (e.g., below the soil surface, soil surface, plant canopy, aquatic) and ecological functions relevant to and representative of the cropping system of interest (e.g., pollinators, predators, parasitoids, decomposers).

- *Ability to conservatively estimate field exposure*. A conservative Tier I test can be designed by first identifying the potential routes of exposure of the indicator species to the insecticidal protein in the field. However, even when exposure cannot be estimated accurately, the concentration of the insecticidal

protein in the plant will provide a worst-case estimate of the Environmental Exposure Concentration (EEC).

- *Whether a suitable test system exists for laboratory analysis*. Test organisms adaptable to a laboratory bioassay system and suitable protocols are needed for testing. When feasible, the organism life stage that is most susceptible to the insecticidal protein is tested. Protocols typically include information on test endpoints, positive/negative controls, acceptable control mortality, sample sizes and statistical power analyses. For a number of species, standard testing protocols are available that could be adapted for the testing of insecticidal proteins.

- *Susceptibility to chemical pesticides typically applied to the host crop*. Choosing test organisms that are susceptible to chemical insecticides used on the host crop— the use of which may be reduced by the introduction of the insecticidal protein — provides additional insight into the relative risk associated with the protein.

Test organisms are typically chosen case-by-case according to the potential for exposure to the insecticidal protein (taking into account the crop and the region of introduction), as well as the ability to test the organism in the laboratory (EPA-SAP 2000). In cases where a representative exposed non-target organism cannot be adequately tested in the laboratory, a closely-related indicator that can be easily reared may be substituted (EPA-SAP 2000). Dietary toxicity tests may not be necessary if it can be demonstrated that large-scale commercialization of the insecticidal protein will not pose a hazard to close relatives of the target pest (EPA-SAP 2000) or if there is a known history of safe use.

EPA's OPPTS 885.4340 *Non-target Insect Testing Tier I Microbial Pesticide Test Guidelines* that are currently referenced for ecological risk assessment of Bt proteins recommend testing three beneficial insect species in the laboratory, including representatives of parasites and predators known to attack the target pest or share the same ecological habitat (http://www.epa.gov/pesticides/biopesticides/regtools/guidelines/oppts_885_4340.htm). However, more species may be tested for novel proteins that do not have a history of safe use. Less testing may occur for proteins with a history of safe use and existing Tier I non-target organism data. Representative indicator species are typically chosen from primary ecological functional groups that may be exposed to an insecticidal protein in the field. Functional groups may include (1) beneficial natural enemies (e.g., predators and parasitoids), (2) pollinators (e.g., bees, syrphids, lepidopterans), (3) decomposers (e.g., Collembola, earthworms, nematodes, mites, psocids), (4) non-target herbivores (e.g., monarch butterfly), and (5) herbivores that can serve as alternative prey for key natural enemies (EPA-SAP 2002). Standardized indicator species are chosen based on the reliability, repeatability, and cost effectiveness of suitable laboratory tests (EPA-SAP 2002).

Indicator organisms tested for currently registered lepidopteran-active Bt proteins (e.g., Cry1Ab, Cry1F, Cry1Ac and Cry2Ab) have typically included a lady beetle, green lacewing, and a parasitic hymenopteran (*Nasonia vitripennis*). Some of these insect species were chosen because of their history of testing against microbial pesticides, or the availability of colonies and the ease of testing in the laboratory. However, the adequacy of the test organisms to represent organisms relevant to the agro ecosystem where the transgenic crop will be deployed warrants consideration.

Lady beetles are generalist predators that commonly occur in cultivated crops including corn and cotton fields. Therefore, they are considered an appropriate indicator species to be tested in the laboratory. Studies have been conducted with *Hippodamia convergens* (convergent lady beetle) and *Coleomegilla maculata* (pink-spotted lady beetle; twelve-spotted lady beetle). *C. maculata* may be considered for future lady beetle tests because it is more of a generalist predator and tends to feed on pollen more readily than *H. convergens* (EPA-SAP 2001). Testing may also focus on lady beetle larvae because larvae consume more pollen in their diet than adults, and insect larvae tend to be more susceptible than adults to Bt proteins. However, adult tests and other lady beetle species are considered case-by-case, and additional lady beetle tests may be appropriate for coleopteran-active proteins, such as Cry3 proteins.

Green lacewings (Chrysopidae, particularly *Chrysoperla carnea*) occur in many crops such as corn and cotton, but to a lesser extent than other generalist predators such as lady beetles (Coccinellidae), big eyed bugs (*Geocoris* spp.), and the insidious flower bug (*Orius insidiosus*), sometimes called the minute pirate bug (Steffey *et al.* 1999). In some of the dietary toxicity tests conducted with green lacewing larvae, the *Bt* protein was presented to lacewings in a moth egg (*Sitotroga* sp.) diet. However, the *Bt* protein may bind to the surface of the moth eggs, resulting in limited exposure to lacewings that feed with piercing-sucking mouthparts (EPA-SAP 2002). Therefore, it is inappropriate to test the activity of insecticidal proteins, such as a *Bt* protein, by incorporating the Bt protein into a moth egg diet. Furthermore, because green lacewings do not consume much pollen in the field (EPA-SAP 2002) and are not exposed to Bt proteins *via* consumption of aphids (Head *et al*. 2001), other generalist predators should be considered for Tier I tests. If lacewing larvae are tested, a synthetic diet such as those reported by Hasegawa *et al*. (1989) could be used as a delivery system.

The August 2002 FIFRA SAP acknowledged the importance of considering the effects of PIPs on parasitoids. Parasitic Hymenoptera toxicity tests conducted thus far have used *Nasonia vitripennis,* an endoparasitoid of dipteran pupae. However, *N. vitripennis* is not associated with corn pests and does not typically occur in cornfields, and better surrogate parasitoids may exist. Other parasitic Hymenoptera (e.g., ichneumonids, braconids, mymarids, scelionids) that attack larval or adult hosts could be considered for testing when the target crop is corn. However, evaluating parasitic Hymenoptera in the laboratory may not be necessary when an adequate number of other non-target insects are tested and the

exposure of Hymenoptera to the insecticidal protein in the field is minimal because of their feeding habits.

For coleopteran-active insecticidal proteins, non-target Coleoptera may be emphasized in the risk assessment. Carabids (ground beetles) and staphylinids (rove beetles) have been identified as beetle families with important ecological and economic roles within agro ecosystems. Predaceous carabids that are abundant in cornfields and could be suitable for Tier I testing include *Bembidion quadrimaculatum*, which has a high reproductive rate, and *Pterstichus melanarius*, which is a larger species. *Amara* spp. and the smaller *Pterostichus* spp. also are beneficial organisms in cornfields and could be considered for Tier I testing (EPA-SAP 2002). In addition, the predatory staphylinid, *Stenus flavicornis*, occurs in cornfields and would be an appropriate indicator species (EPA-SAP 2002). Although certain carabids and staphylinids have been identified as appropriate indicator organisms for Tier I testing, there are currently no validated test systems for these species. Appropriate protocols for some of these species are currently being developed by government (e.g., USDA/ARS), university, and industry scientists. Designing protocols for other species in addition to those mentioned here (e.g., representative Chrysomelidae for coleopteran-active proteins) may also be considered for the assessment of future PIPs.

In addition to predaceous and parasitic invertebrates, Collembola and earthworms have been tested as representative decomposer organisms, and *Daphnia* have been used as representative aquatic invertebrates. Although Collembola and earthworms have been recommended as appropriate decomposer organisms, they may not always be appropriate indicator organisms and excluding these tests may be acceptable on a case-by-case basis (EPA-SAP 2000). Alternative representative decomposer organisms (e.g., the common pillbug, *Armadillidium vulgare,* or mites) may be considered. Organisms such as mites and pillbugs readily consume decomposing plant material and may be more relevant indicator organisms than earthworms.

Daphnia magna, commonly referred to as the water flea, is easily cultured and tested in the laboratory, and has been used routinely for Tier I acute toxicity testing. However, daphnids feed on algae and are unlikely to be exposed to insecticidal protein expressed in plants so alternative aquatic indicator species, such as certain dipteran larvae (culicids (mosquitoes), chironomids (midges), caddisflies, scavenger beetles, or predaceous diving beetles) could be considered case-by-case.

Tier I Testing Endpoints

Ecotoxicological endpoints are measures of response within a test that can be related to the assessment endpoints. They are derived from the specific measurements (e.g., of mortality or developmental rate) taken during the conduct of an experiment. Biological endpoints include binary and continuous responses. A binary response is recorded as response/no response (e.g., dead or alive) and a continuous response is recorded as a measured response (e.g., weight) or as a count (e.g., number of young produced). Typical endpoints estimated from toxicological tests are the LD_{50}, LC_{50}, $EC_{50,}$ and the no observed adverse effect concentration (NOAEC). Appropriate mortality-related endpoints are the LC_{50} and LD_{50}, which are the lethal concentration and the lethal dose, respectively, at which 50 percent of test organisms are estimated to die. The EC_{50} is the concentration in which a 50 percent impact on development or weight is estimated to occur. The NOAEC is the greatest concentration of a substance, found by experiment or observation, which causes no significant adverse effect when compared with the control. A NOAEC is based on detectable adverse effects on parameters such as morphology, growth, development, or lifespan of the target organism under defined conditions of exposure. The results for endpoints in chronic tests considering survival and growth are reported as an EC_{50} for growth inhibition, a NOAEC for growth inhibition, an LC_{50} for survival, or a NOAEC for survival.

The nature of possible exposure routes of the representative test organism to the insecticidal protein should be considered when determining the endpoints for use in laboratory or field bioassays. While mortality is the primary indicator of toxicity in Tier I testing, sub-lethal effects such as developmental delays, growth inhibition and measures of reproductive capacity can also be recorded as potential indicators of an adverse effect. Tests that do not allow the designated endpoint to be reached are not a useful part of the risk assessment (EPA-SAP 2002, 2004).

In selecting appropriate endpoints for Tier I testing, the following points warrant consideration:

- Focus on mortality, with possible inclusion of other field-relevant (biologically relevant) measures that are predictive of impacts on populations in the field.

- If testing cannot occur at a sufficiently high concentration to represent a conservative estimate, more sensitive endpoints (e.g., growth inhibition) or an alternative testing approach (e.g., a Tier II study) may be needed.

- The endpoints will vary to some extent depending upon the nature of the protein and its mode of action (and how exposure will occur), and the nature of the test system.

- Measures of reproduction are not generally appropriate for Tier I studies, but might be appropriate in higher tier studies.

Studies can include additional observations to determine whether endpoints capture all relevant effects.

Routes of Exposure

Insecticidal proteins, like Cry proteins, have no dermal contact toxicity and must be ingested by a susceptible organism to be effective. Thus, direct dietary exposure is required to evaluate the toxicity of insecticidal proteins against non-target test organisms. In these dietary exposure tests, known concentrations of a protein are incorporated into a diet substrate, from which test organisms ingest it. In nature, this direct dietary exposure represents the primary route of exposure to insecticidal proteins for organisms that feed upon transgenic plants. For higher trophic levels, such as predators and parasitoids, the route of exposure to insecticidal proteins in nature may be *via* consumption of intermediary host or prey species that have fed on transgenic plant tissues or treated diet substrates. This type of indirect exposure (i.e., *via* tri-trophic interactions) is referred to as secondary exposure. Evaluation of secondary exposure (*via* higher-tier tests) typically occurs when Tier I testing shows potential hazard to the tested surrogate organism.

Tier I exposure focuses on microbial-derived protein incorporated into an artificial substrate (diet) or combined with plant or host material to give an adequate safety margin over expected exposure. If it is not possible to expose the test organism to microbial-derived protein, a Tier II test using transgenic plant material is typically conducted. In this case, it may be necessary to augment the transgenic plant material with microbial-derived truncated protein to attain adequate levels of exposure. Foods used by test species in their relevant habitat are appropriate substrates for consideration in laboratory tests (EPA-SAP 2002, 2004). To ensure adequate exposure throughout the test, verifying the level and activity of toxin in test material and exposure to all life stages is recommended (EPA-SAP 2002, 2004).

Direct routes of exposure are appropriate for Tier I testing because secondary exposure generally will be much less than direct exposure (EPA-SAP 2000, Head *et al*. 2001, Raps *et al*. 2001, Dutton *et al*. 2002). The maximum amount of an insecticidal protein that an organism, such as an insect herbivore, may contain is limited by the total volume of its alimentary canal, the rate of intake, digestion/degradation and excretion of transgenic plant tissues, and its susceptibility to the insecticidal protein. Head *et al.* (2001) and Raps *et al.* (2001) also suggested that some sap-feeding insect herbivores, such as corn aphids that feed within the phloem, may not ingest detectable amounts of plant-produced Bt proteins, thus predators and parasitoids feeding on these non-target herbivores will not be exposed. Focusing on direct routes of exposure during Tier I testing rather than secondary exposure is preferred for other reasons as well. The magnitude of secondary exposure is often subject to the influences of various ecological and behavioral factors, such as the fate and/or forms of the insecticidal protein in the prey digestive

systems and/or tissues, and the feeding behavior of the prey, predators and parasitoids. Thus, it is difficult to standardize testing methods for evaluation of secondary exposure hazard. Furthermore, many insect predators and parasitoids have the ability to selectively feed on healthy *versus* unhealthy (e.g., intoxicated) prey. This type of feeding behavior would make it difficult to assess accurately the risk of secondary exposure and determine a maximum hazard concentration. In addition, feeding on intoxicated prey as the sole diet (e.g., 'no-choice') in a Tier II study will not allow the investigator to discriminate between any effects of the insecticidal protein *per se* and the effects of a sub-optimal diet (Romeis *et al*., 2004). However, tri-trophic tests under a worst-case exposure scenario, such as feeding larvae of *Chrysoperla carnea* exclusively with spider mites that have fed on the insecticidal protein, may have some value as higher-tier tests (Dutton *et al.* 2002).

Duration of Exposure

A study examining potential risk to a non-target species is typically of sufficient duration for the desired endpoint to be detectable when a hazard exists. A short exposure (e.g., <96 hours) generally is known as an acute test and a lifetime exposure is known as a chronic test. Test durations between one and four weeks do not fall neatly within either category, and may be considered sub-acute or sub-chronic tests.

The duration of Tier I maximum hazard dose (MHD) tests for insecticidal proteins currently is based upon the guidance for microbial pest control agents (MPCAs) (EPA, 1996); tests are required to be as long as is practically possible (the maximum being 30 days). This requirement for the longest practical test duration stems largely from two generally accepted assumptions: (1) microbial pesticides typically have a much slower mode of action than synthetic chemical pesticides; and (2) longer tests allow observation of the effects of delayed toxicity on mortality, growth, and development, and thus allow more conservative risk estimators to be used. In some cases, the observation interval may be longer than the duration of exposure if natural exposure is expected to be short in duration but delayed effects are possible.

EPA guidelines that were developed for microbial pesticides recommend the following durations for Tier I testing:

- Non-target Insect Testing - Tier I guideline recommends a 21-30 day test or that the test be terminated when 20% mortality is reached in the negative control group (OPPTS 885.4340).

- Non-target Honey Bee Testing - Tier I guideline recommends at least 30 days of observation after dosing larvae (OPPTS 885.4380).

- Freshwater Aquatic Invertebrate Testing - Tier I guideline recommends at least 21 days of observation for organisms such as daphnids. If pathogenicity is observed on day 21, observations should continue until recovery, mortality or moribundity occurs (OPPTS 885.4240).

- Earthworm Subchronic Toxicity Test - guidelines recommend observations at defined time-points over 28 days for earthworms. However, this guideline was designed for chemical pesticides.

In cases where the test organism life history and/or the laboratory rearing conditions do not allow for the pre-defined length of test period, the EPA guidelines (EPA 1996) allow the use of 20% control mortality as a criterion to terminate the test. The use of 20% control mortality gives some flexibility in conducting studies with non-target organisms that are not routinely used in regulatory toxicity studies. However, when the test conditions do not permit prolonged survival, the test might be terminated prior to detection of any insecticide-related toxicity. The 2004 FIFRA SAP concluded that "*control mortalities as high as 20% raise a concern of an error with the protocol and that tests either need to be repeated or that some adjustments in protocols are required*." Therefore, if control mortality exceeds 20% before symptoms related to exposure to the insecticide would be expected, the test may not be valid and may need to be repeated. Defining maximum control mortality is valuable because it ensures that the test conditions are appropriate for the test species and it provides a minimum test quality standard. The appropriate level of control mortality will depend upon the choice of test organism and the protocol design and should be determined on a case-by-case basis.

The 30-day test outlined in the OPPTS guidelines is recommended to detect latent pathogenicity or reproductive effects (e.g., in daphnids). However, these guidelines were developed to test microbial pesticides and pathogenicity/infectivity is not an issue for insecticidal proteins. Test durations for studies presented to EPA's 1999 SAP included 2 to 21-day daphnid, 8 to 15-day honey bee, 14-day earthworm, 28-day Collembola, 21-day lady beetle, 15-day parasitic wasp, and 7 to 9-day green lacewing larval studies. These test durations were based on a summary of all of the data reviewed by EPA at that time. After reviewing the regulatory data requirements for non-target organism testing for *Bt* crops, the SAP indicated that the test durations presented by the EPA were appropriate except that the daphnid test should be longer than 2 days (EPA-SAP 2000). The SAP recommended test durations for Bt Cry proteins should be a minimum of 5 days, and preferably 7 to 14 days. However, the SAP (2000) also emphasized that, in specific situations such as testing honey bee brood, a longer exposure period would be needed.

The following factors warrant consideration when planning Tier I tests:

- Test duration should reflect the biology of the system (e.g., susceptible stage, generation time), the nature of the protein, and the test concentration selected. Because the duration will differ among species tested, the length of the test will need to be adjusted based on the biology of the test organism.

- The life cycle of the plant species and/or the length of time the insecticidal protein is expressed by the plant or within a specific plant tissue may be considered. With chemical and/or microbial applications, non-target organisms are usually exposed to a spray or soil application that may decay rapidly, but with insecticidal proteins expressed in plants, the protein may be produced by the plant over a longer period of time. However, in most cases, it will be impractical and unnecessary to have the test run for the life of the plant or the length of time the insecticidal protein is produced. Therefore, other exposure factors may also be considered in determining the test duration.

- Mode and speed of action of the specific insecticidal protein must be considered. Typically, a duration of 7 to 14 days (where technically possible) is reasonable for assessing hazard for *Bt* proteins. This duration is appropriate for the *Bt* Cry protein mode of action; *Bt* Cry proteins target the digestive system of susceptible insects and generally take no more than 3 to 5 days to cause visible symptoms in exposed, susceptible insects.

- A higher test concentration may allow for a shorter duration. For example, acute exposure to a MHD (see Selection of Test Concentrations section) allows for extrapolation to chronic exposures.

- It is preferable to periodically change test diet and verify concentrations of active toxin.

- The endpoint is typically compatible with the test duration. For example, if the endpoint is based on mortality, short-duration exposures (e.g., 7 days) will be adequate for most test organisms. If growth or development time is the endpoint, longer exposures (e.g., 14 to 30 days) may be needed to adequately detect an adverse effect.

- Shorter duration testing should be indicative of broader impacts (mortality and sub-lethal effects) over a longer time period.

Nature of Controls

Controls are typically included in experiments as indicators of the suitability of the test system and for comparison to the data generated for the treatment(s) of interest. For non-target invertebrate testing, several specific attributes of the test system are of interest. Negative controls are included to assess the suitability (health) of test organisms and the test conditions (e.g., temperature and diet) and/or to evaluate potential effects of the matrix or formulation in which the test protein is delivered. Positive controls are included to confirm that the test organism is exposed to the test protein and, in some cases, to assess the sensitivity of the test organism to a standard toxicant.

Negative controls should be included in non-target invertebrate tests to allow background (non-specific) effects to be distinguished from insecticide-induced effects. Negative controls are also useful for assessing if a bioassay is unsuitable for evaluating the test protein because of excessive background effects. Negative controls may include formulations (excluding the protein of interest), non-transgenic plant tissue matched to the relevant transgenic tissue, or preparations in which the PIP protein has been inactivated.

Positive controls generally are not required for non-target invertebrate testing (EPA-SAP 2000), but are often useful. If exposure to the test protein cannot be substantiated through other means, it may be necessary to include positive control substances in the bioassay. Positive controls are intended to: (1) show that the toxin is active in the form provided to the non-target invertebrate. This can be done by a sensitive target bioassay (see Head *et al.* 2001 for an example for the secondary route of exposure); (2) show that the bioassay works and is able to detect an effect. This control is of major importance in the protocol development/validation process; and (3) provide a reference by which the sensitivity of the test organisms can be compared to other similar bioassays using the same species.

A difficulty with this approach has been the lack of a universally accepted ingested (e.g., stomach) poison for use as a positive control. Alternatives for substantiating exposure may include exposing the test organism to dyes under the desired test conditions, followed by dissecting the insect gut and examining its contents. Weighing test organisms, diets, or food before and after exposure (e.g., Romeis *et al*., 2004), or simply observing feeding may also be used to substantiate exposure. A plant-derived lectin with non-specific oral insecticidal activity and a similar mode of action to the protein also may be considered for use as a positive control in future studies.

Points to consider when making decisions on selecting positive controls:

- Positive controls confirm ingestion of the protein and demonstrate the potential to detect a response in the test system.

- Positive controls need not be run every time but are typically run often enough to give confidence in the test system, and are particularly important when the system is novel or the application of the system is novel.

- Test substance characterization of some sort is needed, though it can take various forms.

Selection of Test Concentrations

Currently in the United States, Subdivision M Tier I studies are conducted as no-choice tests at a MHD, usually 10-100x the maximum expected exposure concentration in the field (154A-23, p 147). The reason for conducting these studies at a highly conservative MHD is to allow extrapolation from the surrogate species tested to potentially more

sensitive species within the same taxonomic and/or functional group, and to address both direct (i.e., consumption of plant material) and indirect (i.e., multi-trophic exposure) exposure routes. For insecticidal proteins, it would be appropriate to define the MHD as 10x the protein expression level in the plant via the expected route of exposure. For example, Babendreier *et al.* (2004) calculated the amount of corn pollen that a honey bee larva could digest during development.

If the LC_{50} is greater than the MHD, then there is a low probability of risk or adverse effects to the non-target organism(s). If greater than a 50% mortality/effect is observed at the MHD, then a more definitive dose-response study should be conducted. At the same time, data that establishes an LC_{50}, ED_{50}, or LD_{50} that is greater than the maximum hazard dosage or concentration often is adequate for the purpose of hazard assessment (OPPTS 885.4000, p. 3). The level of exposure should be based on a detailed quantification of the estimated environmental concentration (EEC) in the field (EPA-SAP 2004). No comparable guidance currently exists for chronic testing. The guidance in Subdivision E for chemical insecticides indicates that a 5x safety margin should be included in acute tests for terrestrial non-endangered species and a 1x margin for chronic tests.

Generally, non-target invertebrate Tier I studies with proteinaceous insecticides expressed in plants seek to achieve at or above 10x the reasonably anticipated environmental exposure concentration. If, however, comprehensive expression characterization later leads to a safety margin of <10x, the consequences are considered in terms of uncertainties in the ecological risk assessment. The Tier I study exposures are generally not less than the high-end exposure (the 90^{th} percentile of the characterized distribution of expression through a relevant route of exposure).

The effectiveness of the maximum hazard approach is dependent upon selecting species with a range of relevant physiologies relative to the mode of action of the test material (EPA-SAP 2004). It is important to make a distinction between MHD testing and risk characterization. A risk quotient (EEC/LC_{50} ratio) greater than 0.1 would trigger a risk characterization such as understanding how the LC_{50} compares to expected exposure under field conditions. A less than 50% effect at the MHD is taken to indicate minimal risk. However, a greater than 50% effect does not necessarily indicate the existence of unacceptable risks; instead, it triggers the need for a dose-response evaluation and a refinement of the exposure estimation.

Prior to selecting test concentrations, consideration of the following is recommended:

- The test concentration is designed to reasonably achieve in excess of the anticipated maximum expected exposure concentration (EEC) while accounting for variation in interspecies sensitivity and inter-plant variation in insecticidal protein expression level, as subsequently elaborated in the ecological risk assessment.

- Typically this would represent 10x the EEC *via* an environmentally relevant route of exposure (maximum protein expression can be used as a conservative estimate of EEC) for testing of purified insecticidal proteins. The MHD also could be represented as the High End Exposure Estimate (HEEE) of the plant material to which the organism may be exposed, which is based on an estimate of insecticidal protein expression levels in plant tissue at the upper end of exposure, preferably the 90th percentile.

- The safety margin may be less than 10x where uncertainty in the system is low, such as in a case in which many species are tested or tests are very sensitive, although the concentration used must exceed the EEC.

- The advantages of a MHD over a limit dose/concentration approach (as used with conventional chemical insecticides) include:
    <It does not use as much insecticidal protein test material
    <It does not involve unrealistically high concentrations
    < A determination of no effect equates with minimal risk

Statistical Design

Any test used to assess hazard must be sufficiently sensitive to detect treatment-related effects. The sensitivity of a test will be a function of the test system and its inherent variability, experimental design, and the level of replication. Two statistical approaches are commonly applied to non-target invertebrate toxicity tests. The first is analysis of variance (ANOVA), which is equivalent to a parametric t-test when only two groups are present (as in MHD). The ANOVA approach using software, such as SAS PROC GLM or MIXED, determines whether or not differences in mean response among treatments are greater than expected by chance. A second approach is to apply a proportions test (e.g., z test or SAS procedures PROC CATMOD or GLIMMIX) to binary responses, such as mortality, number of affected individuals, or number of individuals that pupated, to determine if the proportion of individuals exhibiting a response is significantly different from some pre-determined, hypothetical proportion (e.g., 0.5) or from some proportion exhibited by the control individuals. A third possible approach is the analysis of survivorship curves, though this is less often used because it requires substantially more effort to track the response of every individual.

There are two types of potential error that should be considered in selecting a statistical design. The first type of error, known as the type I error or α, occurs when we conclude that treatments are different when in fact they are not (reject the null hypothesis of no differences, H_o, in favor of the alternative hypothesis, H_a, when H_o is true). The second type of error, known as the type II error or β, occurs when we conclude that treatments are not different when in fact they are (fail to reject H_o when H_a is true). The ability to detect effects accurately, when they are present (i.e., not making a type II error), is referred to as the power of the experiment, and is improved when variability is reduced and/or replication is increased. By using approaches that increase the statistical power of non-target invertebrate tests, the chance of detecting a treatment-related effect is improved.

ANOVA (or t-test) is often used to test if the mean response within an experimental treatment is significantly different from that in a negative (e.g., untreated) control. H_o in this case is that the mean response within both groups is equal. Alpha (α) in these experiments is commonly set to 0.05, indicating that the probability of the observed effect size (e.g., difference in mean response) is considered significant if it is less than 5%. The power of the experiment, however, reflects the ability to detect a specified magnitude of difference with a specified effect size. Determining the power of the experiment requires that the variability in the data be estimated, as well as specifying what magnitude of difference is biologically significant.

When using ANOVA, the level of replication that is needed to achieve a given level of power (to detect a specified magnitude of difference with a given probability) depends on variability within treatment groups. The variability estimated from previous experiments can be used to design new experiments with the appropriate replication and sample size. Examples of appropriate sample sizes and replication for specific non-target invertebrate test systems can be found in Candolfi *et al*. (2000a). While these estimates are primarily for contact (chemical insecticide) exposure tests with conventional pesticides, they provide a guide for ingestion experiments as well because control variability should be similar in the two types of tests. One assumption of ANOVA is that the variability among observations within each of the treatment groups is equal. However, a t-statistic which adjusts for heterogeneous variances is available, as are non-parametric test alternatives, should a test for homogeneity of variance reveal statistically significant differences among the variances of different treatment groups. In addition, transforming the data may make the variances conform to the assumptions of ANOVA.

A proportions test provides an alternative statistical approach to evaluate whether an observed proportion of responding individuals is different from a hypothetical proportion. Non-target invertebrate hazard tests often are conducted at exposure concentrations several times higher than the maximum concentrations expected to occur under realistic exposure scenarios. This has customarily allowed an endpoint of 50% mortality to be used as a trigger for additional higher-tier testing. Lower levels of mortality under these conditions of extreme exposure suggest that population effects are likely to be negligible given realistic exposure scenarios. Thus, it follows that the observed proportion of responding individuals can be compared to a 50% effect to determine if the observed proportion is significantly lower than 50%. For example, using a binomial approach, a sample size of 20 individuals is sufficient to allow a treatment effect of 30% to be differentiated from a 50% effect with 95% confidence using a one-sided Z test. A one-sided test is appropriate because only effects of less than 50% indicate that further experiments are not needed to evaluate risk.

In summary, both ANOVA and binomial proportions tests provide valid options for statistically assessing the results of non-target invertebrate tests. If ANOVA is used, previously generated experimental data are needed to estimate the number of replicates required to achieve the desired level of power to detect biologically important effects. A good starting point for these estimates may be found in Candolfi *et al.* (2000). A binomial proportions test provides an alternative when effects on less than a predetermined proportion of a population exposed to excessive treatment concentrations

can be used to adequately conclude that risk to non-target invertebrate populations at realistic concentrations is negligible.

Points to consider in designing and analyzing a Tier I study:

- Necessary statistical power must be defined *a priori* for any study and then this can be used to determine the number of replicates and sample size needed in the Tier I test. This calculation also requires specifying the Type I error, estimating the experimental variance, and defining what level of effect must be detected.

- A reasonable aim is to detect a 50% difference with 80% power (see Candolfi *et al.*, 2000a). Unpublished data indicate that methods used to test ingested proteins have comparable levels of variability to those assessed in the EU by Candolfi *et al.* (2000a) for chemical pesticides, thus comparable criteria should be applicable in both cases.

- Sample sizes should be large enough to provide sufficient statistical power to detect differences that may exist.

TIER II TESTING OF NON-TARGET ORGANISMS

Tier II tests may be triggered if Tier I data show that the insecticidal protein is hazardous. Tier II tests may also be considered instead of Tier I tests when Tier I testing is impractical. At Tier II, the test species is more realistically exposed to the insecticidal protein. These tests may serve to identify false positives from Tier I, thereby providing an alternative to conducting more complex, higher-tier tests. Tier II tests follow the same procedures as Tier I, except that the test substance is plant material collected from insecticide-producing plants with expression levels that meet the criteria for commercially acceptable levels of efficacy. For example, in Tier II testing, test organisms may be exposed to the protein using pollen or leaf discs collected from plants in field studies. In addition to following the recommendations for Tier I testing, further consideration may be given to the nature of the plant material to be tested and the route of exposure under field conditions when designing a Tier II study.

Additional points to consider for Tier II testing are as follows:

- Reference control treatments, such as material from different transgenic or non-transgenic varieties, can be important for separating varietal effects from those caused by the insecticidal protein when non-transgenic isogenic material is not available.

- Testing of relevant plant tissues for levels of expression across a range of production environments is appropriate. If there is wide variation, plants from an environment with a relatively high level of expression is not appropriate. Non-target invertebrate tests are most appropriate when conducted at the high end of reasonably expected environmental exposure.

- Field or greenhouse collection and storage of plant materials follow a collection protocol that maintains the activity of the insecticidal protein until the Tier II test is initiated. Storage stability studies are appropriate to confirm stability of the pesticidal protein.

- Test duration may be extended depending upon desired endpoints, test material, test organism and methods.

TRIGGERS FOR TESTING BEYOND TIERS I AND II

A tier-based testing system links the planned and possible studies in some logical, complementary fashion. Higher-tier testing typically occurs when lower-tier tests indicate a potential risk above a threshold called a trigger. Triggers are identified prior to the conduct of lower-tier studies so that a criterion exists for additional testing where the data justify it.

FIFRA SAPs convened by EPA in 1999, 2000, 2002, and 2004, concluded that there are limitations to laboratory-based testing that may result in the need for field or semi-field testing. Those perceived limitations included: levels and routes of test material exposure used in the laboratory may not be realistic; effects assessments are made after short-term exposure of organisms, not lifetime exposure as it might occur in the field; organisms in the field are subject to supplementary stresses that have additive effects that may amplify impacts that occur under the optimal physical and biological conditions of laboratory tests; and laboratory tests cannot practically evaluate all species that are actually exposed to PIPs in the field. Some of these perceived limitations can be addressed within an appropriate tier-based testing approach that combines hazard data generated in the laboratory with field-based exposure data that place laboratory results within a field context. Others can not readily be determined within the context of regulatory testing and may need to be addressed by research which feeds information back to the regulatory process.

Prior to initiating higher-tier studies, the following warrant consideration:

- Triggers typically reflect a result that exceeds the established risk criteria to ensure that tests are adequately sensitive and conservative.

- The additional testing that is triggered may be broader Tier I testing and/or higher-tier tests.

- Higher-tier tests may replace Tier I studies if suitable surrogates or protocols are not available for Tier I studies (or if this course of action is preferable for other reasons).

- Higher-tier field or semi-field tests may be conducted when results of lower-tier laboratory studies indicate potentially unacceptable risks. In addition, field or semi-field studies may be appropriate if lower-tier studies do not indicate acceptable risk with sufficient certainty; however, additional lower-tier studies may be more suitable depending on the hypothesis to be tested.

- Fieldwork focuses on species that are expected to be sensitive and/or exposed to the insecticidal protein in question, and often include suitable positive controls to test the power of the experiment to detect differences among treatments.

- Certain limitations of field work exist so the appropriate applications of field data should be understood before tests are initiated.

- In some instances, scientific confidence in Tier I testing may be sufficient to remove the need for any further testing when no effects are detected at the MHD (EPA-SAP 2004).

- Higher-tier tests are particularly important for insecticidal proteins that are not amenable to Tier I testing and that have no previous history of environmental exposure. Such cases may lead to insufficient certainty about whether risks are acceptable.

TIERS III AND IV FIELD TESTING

Higher-tier field tests may be triggered when results of lower-tier laboratory studies indicate potentially unacceptable risk. In addition, field studies may be appropriate if lower-tier studies do not indicate acceptable risk with sufficient certainty. However, further lower-tier studies may be more suitable depending on the hypothesis to be tested.

In addition, field studies have their own limitations. Large plot field studies may result in significant within-plot variation leading to the need for additional sampling for precise estimation (EPA-SAP 2004). Therefore, the 2002 and 2004 SAPs recommended intermediate testing, such as extended laboratory tests with realistic substrates and exposure scenarios and semi-field tests, rather than census studies. Semi-field tests may involve individual organisms or multiple organisms in microcosms (small systems similar to larger systems in constitution, configuration or development), mesocosms (surrogates for real ecosystem that are sufficient in size to meet all of the components of interest), field cages, or contained arenas.

Semi-field tests provide a bridge between laboratory and field studies and can be used for gathering acute and chronic data on individuals and populations. Full-field tests may be avoided in cases where semi-field tests demonstrate lack of harm under worst-case field conditions. Large amounts of data gathered during full-field tests often lead to confusion and difficulty because of the low inferential power associated with data on rare or sporadically occurring organisms. Semi-field tests allow for the highest levels of field exposure, they are more cost-effective than full-field tests, and are more feasible to

conduct according to Good Laboratory Practice (GLP) standards (e.g., 40 CFR Part 160) than are full-field studies. These tests can also provide regional distributions of data and results beyond what might be expected from single-field investigations. Season-long exposure of insecticidal proteins in comparison with a program of conventional pesticide spraying is possible with semi-field tests, and interactions between treatments can be avoided. Semi-field tests may also serve as a complement to large-scale monitoring and can be used to determine if direct or indirect effects may be responsible for a perceived change or difference between field observations. Behavioral and sub-lethal impacts can be investigated and detailed observations are possible. For further details see Candolfi *et al.* (2000a and 2000b), Jepson (1994) and Jepson and Mead-Briggs (1992).

Basic features of semi-field tests to consider:

- Guaranteed exposure, if it is going to occur, and a higher likelihood of measuring effects, including mortality, because of confinement of the test system and reduced access of vertebrate predators to insect carcasses.
- Can readily relate results to those obtained in the laboratory.
- Can introduce laboratory-reared, or field-collected and marked individuals, to boost or homogenize sample sizes.
- Crop can be grown in cages, avoiding natural colonization, and allowing pests and beneficial non-target organisms to be added as desired. Immigration of previously unexposed, and emigration of exposed organisms, is prevented (a problem with attribution of treatment effects in field studies).
- Greater numbers of replicates are possible for improved statistical validity.
- Can be compared with multiple pesticides and other treatments.
- Can compare different organisms, different test concentrations, and controls.

Conducting field studies is considered case-by-case, based on the level of potential hazard and exposure, and goals may be adjusted as information and experience accumulate (EPA-SAP 2001, 2004). When field studies are relevant and useful for risk assessment, a sufficiently large scale and with an appropriate design and sampling regime are needed to account for the specificity and season-long expression of the transgene being evaluated. The nature of sampling, including the method, scale and timing, depends on the species being evaluated. An appropriate study design with respect to plot size, replication, sampling method, and sample size will provide a statistically valid test with enough power to detect treatment effects (EPA-SAP 2004). Appropriate positive and negative controls, as well as specified endpoints, also contribute to study design (EPA-SAP 2004). Although a positive control group is important to verify sample methods and show effects of alternative control practices, the 2004 SAP noted that organisms will move between plots and be exposed to the toxic chemical in the positive control plots. In addition, reinvasion of organisms into the chemically treated plots from adjacent plots after toxicity has declined may limit the ability to discriminate between treatments and will potentially negate effects that would persist on an agriculturally relevant scale. Because of test plot size and power of test considerations (in terms of experimental materials and locations), the conduct of Tier III-IV field studies is frequently not practical prior to commercialization of crops expressing insecticidal proteins.

It is important to consider crop environment, function and relationship to the target pests when choosing appropriate indicator organisms. Studies that examine all available above-ground taxa complicate analyses and result in data that are difficult to interpret for treatment-related differences. Although these studies may be useful in narrowing the focus of indicator organisms in future studies, many of the taxa will have too few individuals upon which to base conclusions (EPA-SAP 2004).

The August 2002 SAP made the following general recommendations for designing field tests on the abundance and diversity of non-target arthropods:

- Evaluate sites from a number of candidate locations, possibly in the previous season, to determine whether the organisms of interest are present and sufficiently abundant to provide a basis for statistical discrimination of small but significant effects.

- Use sampling methods of known efficiency and precision with consideration of within-plot variability when determining intensity and frequency of sampling.

- Choose a scale and experimental design that minimizes the risks of edge effects and reinvasion from untreated control plots, and which takes into account the dispersal rate and phenology of the organism of interest.

- Add additional plant genotypes that do and do not express the insecticidal protein, with a clear statement of the number of back cross generations that separate transgenic from non-transgenic genotypes. This is to understand the relative importance of background variation versus the insecticidal protein.

- Attempt to first identify non-target arthropods that might be at risk of toxicological impacts through laboratory studies that focus on representative species.

- Consider barrier/cage studies in the field as an intermediate choice between laboratory and full-scale field studies where effects are detected in the Tier I tests.

- Include a conventional insecticide to serve as a positive control.

- Focus sampling on critical time periods with respect to comparison and control treatments. For example, samples should be taken on the day prior to the application of the conventional insecticide, and then for several days immediately after, for example, on days one, three, five and seven.

- Analyze and interpret data only for those species that are sufficiently abundant, such that sampling precision is high.

Overall, field and semi-field studies are the most direct way to assess potential impacts of crops expressing insecticidal proteins on non-target organisms at a population level.

STACKED AND PYRAMIDED (COMBINED) TRAITS

Combined or stacked trait products are crops with two or more genes introduced with different or complementary modes of action and/or spectra of activity. For example, crops modified to contain two insect-resistance genes, or an insect-resistance gene and an herbicide tolerance gene, are considered stacked or combined trait products. The term "pyramid" is used to describe the special case where the multiple resistance genes present encode insecticidal proteins that target the same pests with possible overlap in their modes of action. For example, a corn or cotton plant containing a Cry1A protein and a Cry2A protein active against the same lepidopteran pest, such as European corn borer or tobacco budworm, is considered to have two "pyramided" genes.

It is possible that inserting two genes into a plant may result in different effects on non-target invertebrates than a single insertion. However, studies conducted thus far with Bt microbial formulations and PIPs have not observed synergistic or antagonistic interactions between different proteins known to be independently active (see http://www.epa.gov/scipoly/sap/meetings/2004/june/final1a.pdf). It is unlikely that a plant containing two different genes as a result of combining two single traits by breeding will result in an increased hazard to non-target arthropods compared to the sum of the effects of the single-trait parental varieties. Therefore, additional data on non-target arthropods beyond what are generated on the individual proteins may not be necessary, provided that the protein expression levels in the stacked genes plant are the same as in the single-gene plants and the susceptibility of target pests to the combined proteins is comparable to their susceptibility to the individual traits. If there is no difference in susceptibility to the combined *versus* the individual expressed proteins among susceptible target insects, then it is unlikely that there will be a difference in susceptibility among non-target organisms.

LITERATURE CITED

Babendreier, D., N. Kalberer, J. Romeis, P. Fluri, and F. Bigler. 2004. Pollen consumption in honey bee larvae: a step forward in the risk assessment of transgenic plants. Apidologie. 35: 293-300.

Candolfi, M.P., S. Blumel, R. Forster, F.M. Bakker, C. Grimm, S.A. Hassan, U. Heimbach, M.A. Mead-Briggs, B. Reber, R. Schmuck, and H. Vogt. 2000. Guidelines to evaluate side-effects of plant protection products to non-target arthropods. IOBC, BART and EPPO joint initiative. 158 pp.

Candolfi, M., F. Bigler, P. Campbell, U. Heimbach, R. Schmuck, G. Angeli, F. Bakker, K. Brown, G. Carli, A. Dinter, D. Forti, R. Forster, A. Gathmann, S. Hassan, M. Mead-Briggs, M. Melandri, P. Neumann, E. Pasqualini, W. Powell, J-N. Reboulet, K. Romijn, B. Sechser, T. Thieme, A. Ufer, C. Vergnet, and H. Vogt. 2000b. Principles for regulatory testing and interpretation of semi-field studies with non-target arthropods. Journal of Pest Science. 73: 141-147.

Candolfi, M.P., K.L. Barrett, P.J. Campbell, R. Forster, N. Grandy, M-C. Huet, G. Lewis,

P.A. Oomen, R. Schmuck, and H. Vogt. 2000c. Guidance Document on Regulatory Testing and Risk Assessment Procedures for Plant Protection Products with Non-Target Arthropods. Escort 2 Workshop. SETAC.

Dutton, A., H. Klein, J. Romeis, and F. Bigler. 2002. Uptake of Bt-toxin by herbivores on transgenic maize and consequences for the predator *Chrysoperla carnea*. Ecological Entomology. 27: 441-447.

Dutton, A., J. Romeis, and F. Bigler. 2003. Assessing the risks of insect resistant transgenic plants on entomophagous arthropods: Bt-maize expressing Cry1Ab as a case study. BioControl. 48: 611-636.

EPA. 1989. Subdivision M of the Pesticide Assessment Guidelines (published through the National Technical Information Service (NTIS) in 1983 (EPA–540/9–82–028).

EPA. 1996. Microbial Pesticide Test Guidelines: OPPTS 885.0001- Overview for Microbial Pest Control Agents. EPA 712-C-96-290 (http://www.epa.gov/oppbppd1/biopesticides/regtools/guidelines/microbial_gdlns.htm)

EPA. 1998. Guidelines for Ecological Risk Assessment. EPA/630/R-95/002F, April 1998 Final. United States Environmental Protection Agency. Washington, D.C.

EPA-SAP. February 4, 2000. Characterization and non-target organism data requirements for protein plant-pesticides. SAP report No. 99-06A for FIFRA Scientific Advisory Panel Meeting held December 8, 1999, held at the Sheraton Crystal City Hotel, Arlington, VA.

EPA-SAP. March 12, 2001. Bt plant-pesticides risk benefit assessments. SAP report No. 2000-07 for FIFRA Scientific Advisory Panel Meeting held October 18-20 at the Marriott Crystal City Hotel, Arlington, VA.

EPA-SAP. November 6, 2002. Corn rootworm plant-incorporated protectant insect resistance management and non-target insect issues. Transmittal of meeting minutes of the FIFRA Scientific Advisory Panel Meeting held August 27-29 at the Marriott Crystal City Hotel, Arlington, VA.

EPA-SAP. August 19, 2004. Product characterization, human health risk, ecological risk, and insect resistance management for *Bacillus thuringiensis* (Bt) cotton products. Transmittal of meeting minutes of the FIFRA Scientific Advisory Panel Meeting held June 8-10 at the Holiday Inn Ballston, Arlington, VA.

Federal Register. July 19, 2001. Plant-incorporated protectants; Final rules and proposed rule. 40 CFR Parts 152 and 174.

Hasegawa, M., K. Niijima, and M. Matsuka. 1989. Rearing *Chrysoperla carnea* (Neuroptera: Chrysopidae) on chemically defined diets. Applied Entomology and Zoology. 24: 96-102.

Head, G. P., C.R. Brown, M.E. Groth, and J.J. Duan. 2001. Cry1Ac protein levels in phytophagous insects feeding on transgenic corn: implications for secondary exposure risk assessment. Entomologia Experimentalis et Applicata. 99: 37-45.

Jepson, P.C. 1994. *Insects, spiders and mites* In: Handbook of Ecoxicology. Ed. Peter Calow. Blackwell Science, pp 299-325

Jepson, P.C., and Mead-Briggs, M. 1992. A discussion of methods used in semi-field studies to evaluate pesticide toxicity to beneficial invertebrates. In: Guidelines for testing the Effects of Pesticides on Beneficials. Ed. S.A. Hassan, Bulletin of IOBC/WPRS, 1992, XV/3, pp 4-18.

Newman, M.C. 1998. Fundamentals of Ecotoxicology. Chelsea, MI. Ann Arbor Press.

Raps, A., Kehr, J., Gugerli, P., Moar, W.J., Bigler, F., and Hilbeck, A. 2001. Immunological analysis of phloem sap of *Bacillus thuringiensis* corn and of the non-target herbivore *Rhopalosiphum padi* (Homoptera: Aphididae) for the presence of Cry1Ab. Molecular Ecology 10: 525-533.

Romeis, J., Dutton, A., and Bigler, F. 2004. *Bacillus thuringiensis* toxin (Cry1Ab) has no direct effect on larvae of the green lacewing *Chrysoperla carnea* (Stephens) (Neuroptera: Chrysopidae). Journal of Insect Physiology 50: 175-183.

Steffey, K., Rice, M.E., All, J., Andrew, D.A., Gray, M.E., and Van Duyn, J.W. (eds.). 1999. Handbook of corn insects. Entomological Society of America, Lanham, MD.

LIST OF ACRONYMS

APHIS – Animal Plant Health Inspection Service
BPPD – Biopesticides and Pollution Prevention Division
BRS – Biotechnology Regulatory Services
Bt – *Bacillus thuringiensis*
CFR – Code of Federal Regulations
EC50 – 50% Effective Concentration
ED50 – 50% Effective Dose
EEC – Expected Environmental Concentration
EPA - Environmental Protection Agency
FDA - Food and Drug Administration
FFDCA – Federal Food, Drug and Cosmetic Act
FIFRA – Federal Insecticide, Fungicide and Rodenticide Act
GE – Genetically Engineered
GLP – Good Laboratory Practices
HEEE - High End Exposure Estimate
IPM – Integrated Pest Management
LC50 – 50% Lethal Concentration
LD50 – 50% Lethal Dose
MHD – Maximum Hazard Dose.
NEPA – National Environmental Protection Act
NOAEC - No Observed Adverse Effect Concentration.
NOEC - No Observed Effect Concentration
OPP – Office of Pesticide Programs
OPPTS – Office of Prevention, Pesticides and Toxic Substances
PIP – Plant-incorporated Protectant
SAP – Scientific Advisory Panel
USDA – United States Department of Agriculture

GLOSSARY OF TERMS

Active Ingredient – in the context of PIPs, an insecticidal substance that is produced in a living plant and the genetic material necessary for the production of the substance (40 CFR § 174.3).

***Bacillus thuringiensis* (Bt)** - a group of rod-shaped soil bacteria that produce various insecticidal proteins.

Census Study - a field study where organisms from a wide variety of taxa are collected and identified (see targeted study).

Cry - a designation used to describe a class of crystalline proteins produced by *Bacillus thuringiensis* bacteria. These proteins are toxic to specific insect taxa, but are harmless to mammals and most beneficial insects.

Diet Substrate – in the context of laboratory-based toxicity testing, material into which test substance (PIP) is incorporated to enable oral exposure of test organisms to the test substance (see synthetic diet).

EC_{50} - the estimated concentration of a test substance that has a defined effect on fifty percent of the test population.

ED_{50} - the estimated dose of a test substance that has a defined effect on fifty percent of the test population.

Estimated Environmental Concentration (EEC) - the estimated concentration of the PIP expected within the ecosystem and to which nontarget organisms might be maximally exposed.

Exposure - the concentration or dose of the insecticidal protein encountered by an organism in the environment.

Good Laboratory Practices - standard criteria for conducting studies or tests that support, or are intended to support, applications for research or marketing permits for pesticidal products regulated by the EPA as defined in 40 CFR part 160.

High End Exposure Estimate (HEEE) - an estimate of exposure that is greater than the 90th percentile of all individuals in a defined population but less than the exposure for the highest percentile in that population (see estimated environmental concentration).

Hazard - the inherent toxicity of a PIP.

Indicator - Any biological entity or process whose characteristics show the presence of specific environmental conditions.

Indicator Organism - an organism selected because it is judged to indicate potential

effects to a larger group of organisms that it represents.

LC50 - estimated concentration of a test substance that has a lethal effect on fifty percent of the test population.

LD50 - estimated dose of a test substance that has a lethal effect on fifty percent of the test population.

Maximum Hazard Dose (MHD) – the dose chosen to represent an extreme exposure scenario, calculated using the HEEE or EEC and incorporating an additional safety factor.

Microbially-derived Protein – protein produced for toxicology testing using transformed micro-organisms such as *E. coli* or *Pseudomonas* spp. This protein is used as a surrogate for the PIP because it is more readily available.

Negative Control - a group of replicates in a test that are not exposed to the test substance, but may be exposed to the buffer or carrier in which the test substance is contained.

No Observed Adverse Effects Concentration (NOAEC) - the lowest concentration used in a test at which no adverse effects were observed.

No Observed Effect Concentration (NOEC) - the lowest concentration used in a test at which no effects were observed.

Non-target invertebrate - an invertebrate organism that is not intended to be affected by the PIP.

Plant-incorporated Protectant (PIP) - a pesticidal substance expressed in a plant to control a specific pest or pests.

Positive Control - a group of replicates in a test that are exposed to a substance known to have an effect (usually toxicity). This is used to show that the test organisms are exposed to the test substance during the test.

Pyramided Genes - insertion into a crop plant of two or more genes that each produces an insecticidal protein active against the same target species (see stacked genes).

Reduced Risk - reduced-risk pesticides have one or more of the following advantages over existing products: low impact on human health, low toxicity to non-target organisms (birds, fish, and plants), low potential for groundwater contamination, lower use rates, low pest resistance potential, or compatibility with Integrated Pest Management (IPM). From a regulatory perspective, a major advantage for reduced-risk pesticides is expedited registration review.

Risk Assessment – assessment of the probability that a harmful condition (hazard)

occurs under a given set of conditions, using both hazard and exposure data.

Safety Factor – a safety factor is applied to the EEC or HEEE to ensure a worst case exposure scenario during ecotoxicity testing. The safety factor targeted in Tiers I and II for PIPs is typically 1, 10 or 100 times the EEC or HEEE (see maximum hazard dose).

Screening Test - a simple test conducted to evaluate the efficacy of a PIP on different taxa. It is frequently done using a diet-overlay assay and provides semi-quantitative data.

Semi-field Test - a replicated test conducted on plants or plant parts in an enclosure in a field setting and incorporating typical environmental exposure.

Stacked Genes - insertion into a crop plant of two or more genes that produce multiple insecticidal proteins or other traits (for example insect-resistance and herbicide tolerance) (see pyramided genes).

Surrogate Organism - an organism selected for laboratory testing because it represents a taxonomic or functional group of organisms that should be addressed in the risk assessment.

Synthetic Diet - a diet that has been optimized for rearing or testing organisms in the laboratory or contained environments. Microbially-derived protein or plant materials expressing the PIP are incorporated into the diet for testing.

Targeted Study - a field study in which, based on previous information, sampling and identification are optimized to collect information on a few selected taxa (see census study).

Taxon/Taxa - a group of organisms within the same taxonomic classification; this may refer to all organisms within an order, family, genus, or species.

Tier I Tests - laboratory test using microbially-derived protein mixed with diet at exposure levels representing at least 10x the highest EEC in a no-choice design.

Tier II Tests - laboratory test using plant material alone or mixed with artificial diet at exposure levels that generally reflect 1x the EEC in a no-choice design.

Tier III Tests - long-term laboratory and/or semi-field test. This can be a full life-cycle test or a tri-trophic test, and may be conducted under greenhouse, screenhouse, or field cage conditions.

Tier IV Tests - full-field tests with larger plots that may be distributed across an area or region.

Trigger – specified criterion, expressed in terms of a test endpoint, that is used to determine whether additional testing is needed.

Tritrophic Study - study of interactions among three different trophic levels, i.e. a plant, a herbivore and a natural enemy (parasitoid or predator).

Trophic Level - an organism's place within a food chain, defined by the organisms it feeds upon.