UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, DC 20460

OFFICE OF CHEMICAL SAFETY AND POLLUTION PREVENTION

August 24, 2011

MEMORANDUM

SUBJECT: Transmittal of Meeting Minutes of the FIFRA Scientific Advisory Panel Meeting held May 24-26, 2011 on Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools

TO: Steven Bradbury, Ph.D.
Director
Office of Pesticide Programs

FROM: Fred Jenkins, Jr., Ph.D. 8/24/11
Designated Federal Official
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

THRU: Joseph Bailey 8/24/11
Acting Executive Secretary
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

Frank Sanders 8/24/11
Director
Office of Science Coordination and Policy

Attached, please find the meeting minutes of the FIFRA Scientific Advisory Panel open meeting held in Arlington, VA on May 24-26, 2011. This report addresses a set of scientific issues associated with "Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools".

Enclosure

cc:

Stephen Owens
Louise Wise
William Diamond
William Jordan
Vicki Dellarco
Margie Fehrenbach
Keith Matthews
Donald Brady
Jack Housenger
Tina Levine
Joan Harrigan-Farrelly
Lois Rossi

Richard Keigwin
John R. Fowle
Jennifer Mclain
Elaine Francis
Patricia Schmieder
Nancy McCarroll
Stephen Nesnow
Timothy F. McMahon
Kevin Crofton
Daniel Villeneuve
Enesta Jones
Vanessa Vu
Douglas Parsons
OPP Docket

FIFRA Scientific Advisory Panel Members

Kenneth Portier, Ph.D. (FIFRA SAP Chair)
Janice Chambers, Ph.D., DABT, Fellow ATS
Stephen Klaine, Ph.D.
Carey Pope, Ph.D.
Martha Sandy, Ph.D
Daniel Schlenk, Ph.D.

FQPA Science Review Board Members

Mitchell Cheeseman, Ph.D.
Robert Denver, Ph.D.
J. David Furlow, Ph.D.
Roberta Grant, Ph.D.
Richard Greenwood, Ph.D.
William Hayton, Ph.D.
Nelson Horseman, Ph.D.
David Jacobson-Kram, Ph.D.
Teresa Leavens, Ph.D.
Bette Meek, Ph.D.
Reynaldo Patino, Ph.D
Terry Schultz, Ph.D.
Geoffrey Scott, Ph.D.

SAP Minutes No. 2011-04

A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools

May 24-26, 2011
FIFRA Scientific Advisory Panel Meeting
Held at the
Environmental Protection Agency Conference Center
Arlington, VA

NOTICE

These meeting minutes have been written as part of the activities of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP). The meeting minutes represent the views and recommendations of the FIFRA SAP, not the United States Environmental Protection Agency (Agency). The content of the meeting minutes does not represent information approved or disseminated by the Agency. The meeting minutes have not been reviewed for approval by the Agency and, hence, the contents of these meeting minutes do not necessarily represent the views and policies of the Agency, nor of other agencies in the Executive Branch of the Federal Government, nor does mention of trade names or commercial products constitute a recommendation for use.

The FIFRA SAP is a Federal advisory committee operating in accordance with the Federal Advisory Committee Act and established under the provisions of FIFRA as amended by the Food Quality Protection Act (FQPA) of 1996. The FIFRA SAP provides advice, information, and recommendations to the Agency Administrator on pesticides and pesticide-related issues regarding the impact of regulatory actions on health and the environment. The Panel serves as the primary scientific peer review mechanism of the Environmental Protection Agency, Office of Pesticide Programs (OPP), and is structured to provide balanced expert assessment of pesticide and pesticide-related matters facing the Agency. FQPA Science Review Board members serve the FIFRA SAP on an ad hoc basis to assist in reviews conducted by the FIFRA SAP. Further information about FIFRA SAP reports and activities can be obtained from its website at http://www.epa.gov/scipoly/sap/ or the OPP Docket at (703) 305-5805. Interested persons are invited to contact Fred Jenkins, Jr., Ph.D., SAP Designated Federal Official, via e-mail at jenkins.fred@epa.gov.

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by EPA, as well as information presented by public commenters.

TABLE OF CONTENTS

PARTICIPANTS 4

INTRODUCTION 7

PUBLIC COMMENTS.....8

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS.....9

DETAILED PANEL DELIBERATIONS AND RESPONSE TO CHARGE.....21

REFERENCES 84

SAP Minutes No. 2011-04

A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools

May 24-26 2011
FIFRA Scientific Advisory Panel Meeting
Held at the
Environmental Protection Agency Conference Center
Arlington, VA

Kenneth M. Portier, Ph.D.
FIFRA SAP Session Chair
FIFRA Scientific Advisory Panel
Date: AUG 24 2011

Fred Jenkins, Jr. Ph.D.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: AUG 24 2011

Federal Insecticide Fungicide and Rodenticide Act
Scientific Advisory Panel Meeting
May 24-26, 2011

PARTICIPANTS

FIFRA SAP Chair

Kenneth M. Portier, Ph.D., Managing Director, Statistics & Evaluation Center,
American Cancer Society,
National Home Office, Atlanta, GA

Designated Federal Official

Fred Jenkins, Jr., Ph.D., FIFRA Scientific Advisory Panel, Office of Science Coordination and Policy, EPA

FIFRA Scientific Advisory Panel Members

Janice Chambers, Ph.D., DABT, Fellow ATS
Director, Center for Environmental Health Sciences, College of Veterinary Medicine,
Mississippi State University, Mississippi State, MS

Stephen Klaine, Ph.D.
Professor
Department of Biological Sciences
Director Clemson Institute of Environmental Toxicology
Clemson University
Clemson, SC

Carey Pope, Ph.D.
Professor, Head & Sitlington Chair of Toxicology
Department of Physiological Sciences
Center for Veterinary Health Sciences
Oklahoma State University College of Veterinary Medicine
Stillwater, OK

Martha Sandy, Ph.D.
Senior Toxicologist and Chief
Cancer Toxicology and Epidemiology Section
Office of Environmental Health Hazard Assessment
California Environmental Protection Agency
Oakland, CA

Daniel Schlenk, Ph.D.
Professor of Aquatic Ecotoxicology & Environmental Toxicology
Department of Environmental Sciences
University of California, Riverside
Riverside, CA

FQPA Science Review Board Members

Mitchell Cheeseman, Ph.D.
Acting Director
Office of Food Additive Safety
US Food and Drug Administration
College Park, MD

Robert Denver, Ph.D.
Professor
Departments of Molecular, Cellular & Developmental Biology and
Ecology & Evolutionary Biology
University of Michigan
Ann Arbor, MI

J. David Furlow, Ph.D.
Associate Professor
Section of Neurobiology, Physiology and Behavior
University of California, Davis
Davis, CA

Roberta Grant, Ph.D.
Toxicology Division
Chief Engineer's Office
Texas Commission on Environmental Quality
Austin, TX

Richard Greenwood, Ph.D.
Professor Emeritus
School of Biological Sciences
University of Portsmouth
Portsmouth, UK

William Hayton, Ph.D.
Professor Emeritus
College of Pharmacy, School of Medicine and Public Health
Ohio State University
Columbus, OH

Nelson Horseman, Ph.D.
Professor
Departments of Molecular and Cellular Physiology
University of Cincinnati
Cincinnati, OH

David Jacobson-Kram, Ph.D.
Associate Director of Pharmacology and Toxicology
Center for Drug Evaluation and Research
US Food and Drug Administration
Silver Spring, MD

Teresa Leavens, Ph.D.
Research Assistant Professor
Department of Population Health and Pathology
College of Veterinary Medicine
North Carolina State University
Raleigh, NC

Bette Meek, Ph.D.
Associate Director Chemical Risk Assessment
McLaughlin Institute for Population Health Risk Assessment
University of Ottawa
Ottawa, Canada

Reynaldo Patino, Ph.D.
Research Scientist/Professor
U.S. Geological Survey
Texas Tech University
Texas Cooperative Fish and Wildlife Research Unit
Lubbock, TX

Terry Schultz, Ph.D.
Professor Emeritus
Department of Comparative Medicine
College of Veterinary Medicine
University of Tennessee
Knoxville, TN

Geoffrey Scott, Ph.D.
Director
National Oceanic and Atmospheric Administration
National Center for Ocean Science
Center for Coastal Environmental Health & Biomolecular Research
Charleston, SC

INTRODUCTION

The Federal Insecticide, Fungicide and Rodenticide Act Scientific Advisory Panel (FIFRA SAP) has completed its review of a set of scientific issues associated with Integrated Approaches to Testing and Assessment Strategies (IATA): Use of New Computational and Molecular Tools.

Advance notice of the consultation meeting was published in the *Federal Register* on March 23, 2011. The review was conducted in an open Panel meeting held in Arlington, VA, on May 24-26, 2011. Dr. Kenneth M. Portier chaired the meeting and Dr. Fred Jenkins Jr. served as the Designated Federal Official.

EPA's Office of Pesticide Programs (OPP) is committed to improving and transforming its approaches to pesticide risk assessment and management via enhanced IATA. OPP views this as a critical time to prepare the program to take advantage of rapidly advancing science and emerging technologies.

The goal is to improve OPP's ability to assess hazard and exposure, and to ensure that pesticides are safe and effective when used according to the label. The purpose of this SAP Review was to seek guidance on OPP's vision, initial efforts, and plans to adopt IATA. OPP requested the SAP's input on plans to maximize use of existing data from similar compounds, inclusion of information from new computational and *in vitro* predictive models of toxicity hazard, and from exposure modeling to target the *in vivo* toxicity testing that is necessary to assess and manage chemical risks appropriately. OPP plans to build on an established foundation of using a variety of tools in a tiered testing and assessment framework by systematically adding new tools, methodologies, and an advancing understanding of key events in toxicity pathways. Two case studies were used to illustrate the use of these approaches.

The first case study illustrated the use of genomic information to assess risk using the fungicide propiconazole as an example. The second case study illustrated the use of adverse outcome pathway information to assess risk using the antimicrobial triclosan as an example. The purpose of both case studies was to demonstrate the approaches used to develop a mechanistic basis to support efficient and defensible risk assessment.

The purpose of this consultation was to seek the SAP's advice on the: 1) articulation of OPP's vision to make the risk assessment process more efficient and informative on a sound scientific basis, 2) scientific principles that should be applied to gauge the reliability and robustness of AOPs, 3) utilization of OECD principles for (Q)SAR expert systems, the IPCS MOA Framework and other criteria to reliably integrate predictive tools early in the process, 4) the building on traditional approaches in concert with the new Toxicity Testing in the 21st Century approaches to speed the development and discovery of AOPs and their subsequent acceptance, 5) concepts, considerations and factors that OPP should take into account in terms of developing and applying measures of success, and 6) OPP's utilization of toxicogenomic technology and

the AOP to inform the risk assessment case studies of propiconazole and triclosan respectively.

Opening remarks at the meeting were provided by Dr. Steven Bradbury: Director of OPP and Dr. John R. Fowle III: Deputy Director of Health Effects Division, OPP, EPA. Presentations were provided by Dr. Elaine Francis: National Program Director for Pesticides and Toxics Research, Office of Research and Development (ORD), EPA; Dr. Daniel Villeneuve: Aquatic Toxicologist ORD, EPA; Dr. Patricia Schmieder: Chief Molecular and Cellular Mechanisms Research Branch, Mid-Continent Ecology Division (MED), National Health and Environmental Effects Research Laboratory (NHEERL), ORD, EPA; Ms. Kristie Sullivan: Pesticide Program Dialogue Committee; Ms. Nancy McCarroll: Geneticist, OPP, EPA; Dr. Stephen Nesnow: Senior Research Scientist, Integrated Systems Toxicology Division, NHEERL, ORD, EPA; Dr. Timothy F. McMahon: Senior Toxicologist, Antimicrobial Division, OPP, EPA; Dr. Kevin Crofton: Integrated Systems Toxicology Division, NHEERL, ORD, EPA; Dr. Jennifer McLain: Deputy Director, Antimicrobial Division, OPP, EPA

PUBLIC COMMENTS

Written and/or oral statements were provided by:

Catherine Willet, Ph.D. on behalf of the Ethical Treatment of Animals

Kristie Sullivan on behalf of the Physicians Committee for Responsible Medicine

James Swenberg, D.V.M., Ph.D. of University of North Carolina Chapel Hill, NC

Marty Bernstein, Ph.D. on behalf of himself

Robert Finking, Ph.D. on behalf of BASF Chemical Company

Syngenta Crop Protection on behalf of themselves (including Amber Gomez, Ph.D., Jay Goodman, Ph.D., Richard Peffer, Ph.D. DABT, Curt Omiecinski, Ph.D., and Barbara S. Shane, Ph.D., DABT,)

Wendelyn Jones, Ph.D. on behalf of Croplife America

Summary of Panel Discussion and Recommendations

Charge Question 1a: *Does the Panel believe that OPP's strategic vision clearly articulates a sound scientific basis to making the risk assessment process more efficient and informative? To what extent have we described the critical components (i.e., knowledge bases and databases of existing information, predictive tools (e.g., (Q)SAR (Quantitative Structure-Activity Relationship), in-vitro methods, High Throughput Screening (HTS) assays and Adverse Outcome Pathways (AOPs)) and the progression of events (i.e., research and the use of research) to support the evolution from the current testing and assessment paradigm to an improved IATA of targeted tiered testing based on an understanding of AOPs (i.e., the molecular and cellular events leading to adverse effects at the whole organism level)? Are there any key science issues that the Agency has not captured that are necessary elements to consider in an IATA approach?*

The Panel agreed that the Agency clearly articulated a sound scientific basis for utilizing the National Research Council's (2007) recommendations regarding "21st Century Toxicity Testing" in a manner that makes the risk assessment process more efficient and informative. They also expressed favor for the Agency's proposed use of the AOP methodology to support its vision for employing Integrated IATA strategies. The Panel believed that the Agency adequately described the critical components & modern technologies needed to support the evolution of IATA. They concluded that the Agency's articulated vision of toxicity testing and assessment organized around the AOP approach is a sensible and logical approach which should improve the efficiency of risk assessments. They also concluded that the Agency has obviously identified the necessary steps and has established a foundation for the transition from traditional toxicity testing to testing using 21st Century techniques and approaches. The Panel noted several challenging issues and/or questions that should be addressed under the topics of: 1) Exposure assessment (relating environmental exposures to biomolecule exposure), 2) Dose response (biomolecule exposure to pathway perturbation), 3) Repair / homeostatic mechanisms (off-setting mechanisms that limit apical toxicity), 4) Interspecies differences in AOPs (pathway differences), 5)Absorption, Distribution, Metabolism, and Elimination (ADME) data (accurate data needs), 6) Delayed responses (dissociation in time due to protein synthesis for example), 7) Variable exposures durations (what is the exposure metric?), and 8) Synergism (AOPs that cross-talk).

Other issues mentioned by the Panel included the need for: 1) harmonization of terminology as the field matures to ensure everyone is working from the same foundations; 2) mapping how AOPs will contribute to more efficient tiered priority setting and assessment strategies, and 3) an approach to linking mechanistic data to *in vivo* adverse outcomes of regulatory interest. They pointed out that while there are known and unknown uncertainties in the current *in vivo* toxicity test methods, over three decades of experience with these tests provide a level of comfort that is not present with current IATAs with a different set of known and unknown uncertainties.

Charge Question 2a*: Given that there could be hundreds of AOPs and that varying levels of AOP information are needed depending on the decision being made (i.e., the tier of the risk assessment), please discuss principles for scientific acceptance (i.e., tools for gauging the reliability and robustness) of AOPs and the use of AOPs to move away from chemical by chemical approaches.*

The Panel expressed agreement with OPP's forecast that hundreds AOPs will need to be identified in the coming years and evaluated for plausibility and soundness. Some Panel members suggested that each AOP be considered a hypothesis or theory, the value of which will depend on the strength of the available supporting scientific evidence and the extent to which the AOP is experimentally tested and found to be consistent with empirical data. They remarked that an AOP should be considered a "living document", that is, one that is continuously updated as new knowledge becomes available.

The Panel provided detailed discussion on the principles of scientific acceptance, on evaluating the reliability and robustness of AOPs, and the transition away from individual chemical to chemical categorical approaches through the use of AOPs. These principles are derived from a recent OECD workshop entitled "Using Mechanistic Information in Forming Chemical Categories Series on Testing and Assessment" (OECD, 2011). These principles specify that an AOP should be based on a single, defined Molecular Initiating Event (MIE), linked to key events and a stated *in vivo* adverse outcome(s) with a qualitative assessment of its experimental support that includes documented identification of the MIE and molecular site of action, the key cellular responses, the target tissue/organ(s) and key tissues involved in organ responses with both physiological and anatomical assessments, and, if required, key population responses. The assessment should include documented identification of the MIE, other key events, and response-to-response relationships required to scale *in vitro* effect(s) to *in vivo* outcomes.

The Panel discussed how the assessment of the evidence in support of an AOP should include criteria based on the International Program on Chemical Safety (IPCS) MOA framework (Boobis, 2006) that addresses: 1) the concordance of dose-response relationships, 2) the temporal concordance among key events and the adverse outcome, 3) the strength, consistency, and specificity of association between the adverse outcome and the initiating event, 4) the AOPs biological plausibility, coherence, and the consistency of experimental evidence, and 5) an identification of any uncertainties, inconsistencies and data gaps. Referring to the OECD Workshop on Using Mechanistic Information in Forming Chemical Categories (OECD, 2011), the Panel also pointed out that confidence in the AOP is ascertained by answering the following questions: 1) How well characterized is the AOP?; 2) How well are the initiating and other key events causally linked to the outcome?; 3) What are the limitations in the evidence in support of the AOP?; 4) Is the AOP specific to certain genders, life stages, age classes, etc.?; and 5) Are the initiating and key events expected to be conserved across taxa? They remarked that answers to these questions will come from mechanistic investigations using specifically designed animal models, by comparisons among chemicals suspected of having common MIEs and periodic re-evaluation and re-formulation of AOPs. Finally,

there was general agreement among the Panel that required weight of evidence for any application of the AOP framework will be greatest at this early stage of implementation and paradigm shift but that weight of evidence required should decrease as supporting data, experience and resulting confidence in the approach increases with use.

The Panel concluded that being able to reliably group chemicals based on chemical and metabolic properties is critical to moving away from chemical by chemical assessments. Transparent descriptions of plausible progression of effects in each AOP will be critical to forming functional groupings. Development of these descriptions may, in the short run, require running *in vivo* experiments simultaneously with *in vitro* studies, and require the development of methods of comparing multiple toxicological profiles to identify areas of difference and commonality between AOPs and further understand and quantify uncertainties with the models (model pathways).

Charge Question 2b: *OPP plans to evaluate new tools and information using available criteria and methods such as the OECD principles for (Q)SAR expert systems and the EPA-IPCS (International Program on Chemical Safety) MOA framework. Please discuss these and other criteria the Agency can use to reliably integrate predictive tools earlier in the process.*

Panel members reviewed how the Organization of Economic Cooperation and Development (OECD) principles and the EPA-IPCS (International Program on Chemical Safety) MOA framework can be used to reliably integrate predictive tools into the risk assessment process. The OECD principles for considering a QSAR for regulatory purposes specify that it should have a defined endpoint and domain of applicability with an unambiguous algorithm, and use appropriate measures of goodness-of-fit, robust algorithm fitting and have high predictability. Lastly, the principles specify that a QSAR should ideally have a mechanistic interpretation. The Panel suggested that although flexibility in use of these criteria is allowed, applying the OECD principles in a transparent manner is important to the acceptance and utility in the EPA-IPCS framework. Thus, there needs to be a clear understanding of how each model was developed so that "fitness for purpose" and weight of evidence for the proposed AOP can be considered.

The Panel pointed out that developing predictive tools that can be effectively incorporated into the assessment process at an early phase requires being able to predict internal exposures with some accuracy. It is difficult to predict internal exposures because internal concentrations vary with time depending on the species and strain of animal used, the route of administration, the nature of the carrier solvent (if one is used), and in some cases the initial dose regime. Developing *in silico* models to predict *in vivo* toxicity is also difficult, especially for longer-term health effects, because identifying and modeling the apical event involves understanding both the pharmacodynamic and pharmacokinetic phases of poisoning, as well as their interactions. The AOP approach should facilitate the development and use of a series of models capable of reliably predicting specific processes or key events in the toxicity pathways related to a chemical

or class of chemicals. Iteratively testing these pathways will allow refinement of our understanding of the critical pathways and lead to identification of the critical MOA.

Charge Question 2c*: How can the foundation of knowledge generated through historical and on-going reductionist (i.e., making use of information from lower levels of biological organization to understand adverse outcomes), hypothesis-driven, biological experimentation be effectively mined and integrated with the capabilities of omics (including toxicogenomics, transcriptomics, proteomics, and metabolomics) and high-throughput screening to speed the discovery and development of AOPs and their subsequent acceptance as a credible foundation for predictive risk assessment?*

The Panel noted that predictive toxicology must be based on an understanding of the underlying chemical and biological mechanisms behind an observed toxicological response. Integration of experimental data and newer omic and HTS technologies is currently being done by biomedical researchers in academia and pharmaceutical companies. These researchers are investigating the genetic and biochemical basis of disease and are attempting to identify molecular targets that may be responsive to tailored drugs or that can be used in screening for adverse health conditions such as cancer, metabolic syndrome, and hypertension. It is clear from the propiconazole case study that Agency researchers have a good grasp on how to integrate traditional biological experimentation with newer technologies. A continued interaction of Agency scientists with academic and pharmaceutical researchers is a key to the Agency being able to use this newer technology in identifying AOPs. In addition, the Panel suggested that increased interactions between toxicologists and basic scientists should increase access to new and emerging approaches to analyzing and modeling data from new technologies thereby offering up opportunities to advance the science and application of AOPs. The Agency is encouraged to continue to seek the advice of expert panels of scientists to review highly technical information from specific research areas to integrate diverse findings and analyze and provide guidance on controversial findings.

Accelerating the development and use of AOPs may be facilitated through the use of technology (i.e., wiki-based on-line environments) to improve the effectiveness of communications among scientist from diverse areas of science that don't naturally communicate with each other. This improved interaction is critical to the kind of synthesis of knowledge about chemical interactions, metabolism, systems biology, and ecology needed for AOP development. Similarly it is important that data from HTS, omic scans, and chemical properties continue to be available in open databases or libraries. These environments along with formal training programs will be critical to ensuring sufficient graduate students and post-doctoral fellows enter the field.

The Panel recommended that the Agency also consider supporting research that develops hypotheses on AOPs starting from what is currently known about human biology and human disease. The Panel provided examples of a couple of recent studies where gene-centered databases are mined based on information gleaned from human disease pathway and chemical toxicity pathway knowledge to hypothesize novel AOPs.

Charge Question 3: *OPP is at the beginning stages of contemplating metrics to demonstrate success in achieving our vision. What is the Panel's initial thinking regarding the concepts, considerations, and factors OPP should take into account as we evaluate our progress? What methodologies are best suited to measure success in achieving our vision? For example, what methods would you recommend to quantify the development of AOPs and the qualitative and quantitative application of AOPs in risk assessment?*

The Panel agreed with the Agency that achieving even the near-term objectives to implementing AOPs in risk assessment will take a decade or longer. It is essential that any AOPs that might form the basis of a risk assessment be scientifically defensible and must relate to a critical toxicological effect, and not just an auxiliary phenomenon, that is part of the most sensitive, or one of the most sensitive, systems impacted by the action of the chemical.

Any AOPs used should be consistent with epidemiological or case report data that link the chemical or chemical class to specific adverse health outcomes. Initial AOP identification should be done on chemicals with rich data sets, ones where solid experimental designs have been employed in generating the data considered critical to specifying the AOP. Concordance of AOPs should be verified by *in vivo* results, likely requiring more animal testing in the short term. For novel chemicals, pathways identified must be pertinent to the chemical. Progress will include demonstration that AOP's are grounded in a solid understanding of the normal biochemistry and physiology of the organism accompanied by a well-considered assessment of the key events that are impacted, including the dose-response associated with impacts on these key events. Additionally, there must be assurance that the data sets upon which the AOP's are based are carefully designed using appropriate dose/exposure levels and that comparisons within the experimental groups including controls are valid.

A major part of program evaluation will be peer review. This includes peer review of the initial AOPs, including data sets used, rationale for AOP identified, and consistency of data sets with postulated pathways. AOPs will be continually updated and refined as new knowledge is incorporated into our general understanding of key pathways and processes, so peer review will need to be repeated occasionally.

Progress should be evaluated in the context of how quickly the strategic vision translates to a tactical application. The Panel suggested that one measure of program success would be buy-in by stakeholders in the process. Another measure might be verification by multiple institutions of the AOP through well-designed and coordinated data collection and analysis. This approach addresses directly the issue of consistency of findings that is important to utilization of the AOP framework. Success stories can be an important measure of progress as well as critical to maintaining momentum and enthusiasm.

The Panel also suggested that EPA may wish to examine other similar types of hazard and risk based approaches such as the Hazard Analysis and Critical Control Points (HACCP) process used in assessing food safety.

The Panel suggested that the Agency consider establishing a timeline for AOP development with key milestones against which it can measure progress. Suggested milestones included application of the strategy to key taxa to testing in (1) mammals, (2) non-mammalian vertebrates, and (3) invertebrates. Key milestones for omics integration should be included. Other key milestones include establishment of standards for the use of positive and negative controls in genomics, transcriptomics, proteomics and metabolomics studies.

Case Study 1 Charge Question 1: *Please comment on the use of omic and related technologies employed to develop and link Key Events to create the MOA/AOP of propiconazole. Was the resultant MOA/AOP for propiconazole logical and scientifically sound?*

The Panel concurred that the use of omic and related technologies to develop and link key events to create the AOP of propiconazole demonstrated how powerful and useful these technologies can be. The benefits of discovery–based, hypothesis-generating omics studies are clearly demonstrated in this case study. The Panel believed that overall, the proposed AOP for propiconazole is logical and scientifically sound, with the caveat that so far, the AOP best describes only propiconazole at the doses used and in mice. However, the proposed key events should continue to be tested experimentally and modified and refined as needed.

One advantage of the omics data in the propiconazole case study was that it allowed for more insights into mechanisms of toxicity. Specifically, the AOP defined a series of key events which were based on apical biological observations. The Panel outlined the following strengths in using omics to develop an AOP for propiconazole:

1. The evidence for CAR/PXR activation as a key initiating event was well reasoned and adequately supported by the omics data.
2. A strong case was also made for regulation of cell cycle genes as a key event leading to liver hyperplasia, and further, that altered cholesterol metabolism is an important contributing factor inducing this key event in the AOP.
3. A perhaps unanticipated new role for alteration of retinoic acid (RA) metabolism as a key event was uncovered by the integrated omics approach (including support from direct measurement of hepatic RA levels).
4. Metabolomics revealed changes in protein and lipid oxidation, and altered glutathione metabolism, that supported the transcriptomic responses implicating increased oxidative stress in response to propiconazole.

Despite excellent progress on the use of omics to delineate an AOP for propiconazole, the Panel felt the following points still need to be adequately addressed moving before forward:

1. The necessity of all of the key events hypothesized for propiconazole to induce tumors versus the sufficiency of inclusion of just one or two events was not clear and should be demonstrated.
2. There were varying levels of evidence to support each of the key events proposed for the propiconazole AOP.
3. Significant concerns were also raised by multiple Panel members and public commenters regarding the effect of propiconazole on mutagenesis, as assayed using the BigBlue® mouse model.
4. Differences between the key events underlying phenobarbital vs. propiconazole AOPs were not convincingly demonstrated.
5. Some Panel members raised the issue that identification of key events from *in silico* pathway analysis of microarray data was still subject to considerable individual interpretation
6. A recently published study, which is relevant to the evaluation of the MOA and also may have implications for addressing the topics transferring hazard information across chemicals (Case 1, Question 3) and of species differences (Case 1, Question 4), was not included in the information presented to the SAP .

Case Study 1 Charge Question 2: *Are there other approaches/technologies that could be used to develop and link Key Events to develop MOAs/AOPs for propiconazole and/or chemicals?*

The Panel agreed that other technologies could be used to develop and link key events. While there are new techniques that hold promise in advancing toxicological knowledge, especially along the cellular response pathway, results from these omics and HTS investigations should fit in a logical fashion with existing *in vivo* and *in vitro* data. While genomics data help to uncover toxicological mechanism, HTS is more likely to enhance identification of the chemical space associated with a particular key event. Upcoming new technologies mentioned included deep sequencing of whole and individual genomes and increased availability of post-translational epigenomic assessments. Periodic evaluations of technologies will be just as necessary as the periodic re-evaluation of proposed AOPs. Many of today's omic technologies will be obsolete within a few years. The Agency will need to be flexible and prudent in its investments in this area. The Panel suggested that the Agency continue to focus on a top-down (i.e. apical outcome to molecular events) approach because it has the advantage of less costs and better use of limited resources.

Several Panel members observed that knowledge of physical characteristics, reactivity of chemical moieties and the ability to search for similarities in structure among chemicals with known outcomes (e.g., using MetaPath), is a very powerful tool in understanding the toxicity potential of uncharacterized chemicals. The development of such knowledge bases is very important, both as the foundation for the development and use of expert systems and in identifying targets for screens to ensure that new and useful information will be provided to fill gaps in that knowledge. Still, toxicological principles driving dose selection, frequency of exposures, timing of exposures, timing of assessments relative to exposure(s), etc. need also to be considered fully in determining

environmentally relevant altered pathways that can lead to toxicity.

Case Study 1 Charge Question 3: *Please comment on the general utility of AOPs/MOAs or Key events to read across chemicals to provide a more defensible weight of evidence support for the traditional risk assessment approaches that OPP will continue to use in the near- and mid-term?*

The ability to use an AOP developed for one compound for other similar chemicals ("read across chemicals") for risk assessment is highly desirable because it reduces cost, effort, time, and test animals. Similarly, the identification of AOPs/MOAs or key events is important in traditional risk assessment because it supports a more predictive and realistic assessment. An important advantage of the MOA/AOP approach to read across is that it allows categorization of chemicals based on toxicological similarity. Categorizing chemicals based on both MIE and first or early key events provides stronger weight of evidence than if membership is based solely on the MIE because it provides a higher level of confidence that a particular AOP has been initiated. The utility of the MOA/AOP/key event framework lies in its ability to categorize chemicals based on toxicological similarity.

The utility of the AOPs/MOAs is increased by assigning to each AOP a statement of the confidence that the AOP is active for that chemical or chemical class. The confidence statement must identify and integrate relevant data, analysis, interpretations and conclusions in a transparent report. Traditional toxicity assessment (i.e., hazard-identification and dose-response assessment) typically involves the prediction of a safe dose or concentration or the identification of an appropriate point of departure (POD). The need to identify the adverse effect and the appropriate POD based on that adverse effect remains even under the proposed new paradigm. The information gained from the AOP/MOA approach will help identify the targeted animal testing needed to identify the appropriate POD for the most relevant, critical adverse health effect (i.e., the adverse effect that occurs at the lowest concentration). In addition, an understanding of the MOA will allow the use of more information and better methods of extrapolating below the POD and also allows the use of data-derived uncertainty factors instead of default factors.

Case Study 1 Charge Question 4*: Please comment on the potential use of omic and related technologies to inform AOPs in non-mammalian taxa for use in ecological risk assessment?*

The predominate use of omic and related technologies in ecological risk assessments (ERAs) to date has been in characterizing molecular and cellular consequences of stressor exposures. The large number and diversity of non-mammalian species suggests that a correspondingly larger effort will be required by the research community and by the Agency to get to the point where they can be effectively utilized in ERAs.

MOAs and AOPs are already used in studies of the acute toxicity of agro-chemicals. By design, herbicides are toxic to plants and insecticides are toxic to insects. For ecological risk studies, this knowledge can and is used to focus testing on receptors and life stages of interest. Genomic information on an increasing number of taxa is becoming available. This information will be useful in ERAs, especially when the focus populations are endangered species in which testing is prohibited. Linking molecular effects to apical endpoints in the population will be very difficult and more model systems where this has been successfully done are needed. Major challenges that remain to be overcome have been listed by the Panel and relate to the wide range of homeostatic responses, translation of *in vitro* responses to organisms, fluctuating environmental conditions, differential energetic responses, adaptive responses, organism variability, unknown similarity of responses among non-mammalian and mammalian species, life-stage differences in response, and quantification of dose-response.

Suggestions were provided by the Panel on areas where the Agency might initially focus with increased likelihood of success. This includes using fish models as source of information on links between omics measurements and reproductive/growth outcomes. For invertebrates, the Panel suggested focus on chemicals where the hazard is unique and hence more easily tracked (e.g. chitinase inhibitors) and linked to pathways more easily measured with omics or HTS methods. Finally, it was noted that specifically for propiconazole, information about AOPs in non-mammalian taxa is insufficient to allow generalizations. The Panel suspects that omics technology will be effective in determining if the proposed MOA for propiconazole, or triclosan or any other xenobiotic can be used in ERAs.

Case Study 2 Charge Question 1: *Use of the adverse outcome pathway concept to characterize the impact of triclosan on the thyroid hormone system promotes a research and risk focus on toxicodynamic and kinetic differences in thyroid homeostasis between (1) the study species (e.g.., rats) and the population of interest (e.g., pregnant women/fetus and young children) and (2) potential species similarities and/or differences in the MIEs. Please comment on scientific knowledge and types of data considered to be informative in selecting appropriate factors for interspecies differences.*

This issue goes to the heart of the rich knowledge of vertebrate thyroid hormone pharmacology. The vertebrate hypothalamo-pituitary-thyroid axis formed in the earliest

jawless vertebrates and has been evolutionarily conserved for more than 450 million years. Deep and specific knowledge exists for thyroid hormone physiology, pharmacology, toxicology and metabolism. The potential for adverse outcomes due to thyroid disruption is very large given the critical role that this hormone plays in neurological development. As noted in the issue paper, when key molecular events are conserved, then a mechanistic-based understanding of an AOP may be extrapolated across species (US EPA, 2011). The hypothalamo-pituitary-thyroid axis is composed of a number of discrete steps at the molecular level, any one of which could be the site of the MIE that ultimately alters thyroid hormone activity.

Issues that arose in the discussion of triclosan and its effect on the thyroid hormone system included:

- Lack of evidence regarding whether the triclosan-induced decrease in serum T4 in the rat is due to interference with the T4 radioimmunoassay (i.e., an artifact).
- The unexpected lack of change in serum TSH in rat in the presence of reduced T4, which is inconsistent with the proposed AOP.
- Failure to find reduced T4 in triclosan-exposed humans, which is inconsistent with interspecies scalability of the AOP.

Scientific knowledge and types of data that would be informative in selecting appropriate factors for interspecies differences included:

- Published models of thyroid kinetics in several species; e.g., a computational model of the human thyroid system, a biologically based dose-response model for dietary iodide in rat, longitudinal studies of thyroxine kinetics in pregnant and non-pregnant women, a PBPK model in rabbit of iodide in pregnant females and fetuses, pharmacokinetic models of thyroxine in male and female beagle dog, and rat, and a genetic study of thyroid hormone transporters in humans that identified polymorphisms and their influence on thyroid hormone levels.
- Knowledge that is available from studies of the effects of pharmaceuticals, mutations, and disease processes on thyroid hormone physiology and pharmacology has potential to inform and strengthen the AOP.
- A scaling relationship for the dosage (exposure) metric for species and body size is needed, with body weight to the ¾ power suggested as the default.
- Whether the same key events and AOP apply in the particular species, with the same dose-response relationships; e.g., one MOA for reduced serum [T4] involves competition for binding to thyroid transport proteins, the specificity of which shows considerable interspecies variability.
- Use of *in vitro* liver metabolism studies and interspecies differences in transcriptional regulation of key enzymes to inform interspecies differences in triclosan and thyroid hormone clearance.
- The possibility for independent or concurrent activation or inhibition of other pathways that impinge on thyroid hormone homeostasis.

- Development and use of a PBPK model for triclosan across life stages and species.
- Use of chemical specific adjustment factors when a PBPK model is unavailable for interspecies scaling of triclosan exposure.
- Use of clearance concepts and interspecies scaling relationships for clearance to quantify and scale exposure to triclosan across species.
- Metabolite pharmacokinetic relationships for instances when the MIE involves a metabolite of the parent toxicant.

Case Study 2 Charge Question 2: *Please comment on the factors most responsible for the variability of toxicokinetics and toxicodynamics of compounds affecting thyroid hormone homeostasis among humans (particularly susceptible subpopulations such as pregnant women and children).*

The Panel concluded that the main contributing factors to variability both in natural ranges of T4, T3, and TSH in human populations, as well as responses to replacement therapy with T4 or combinations of T4 and T3 are currently poorly understood in both humans and animal models. The following factors responsible for the variability of toxicokinetics and toxicodynamics of compounds affecting thyroid hormone homeostasis among humans were identified and discussed:

- Known effects of pregnancy on drug pharmacokinetics were summarized to inform possible effects of pregnancy on chemical exposure and thyroid hormone system kinetics.
- Known changes in the capacities of drug clearance pathways during gestation and during maturation to adulthood were summarized to inform possible changes in chemical exposure and thyroid hormone system kinetics in fetus and children.
- Recent progress has been reported on the influence of specific gene polymorphisms on individual human variability in thyroid axis set points and responses to administered thyroid hormones, with one interesting linked candidate as the Type II deiodinase (D2).
- A great deal of individual variability may reside in liver or other extrathyroidal metabolism as well as differential tissue uptake, but this remains in question at the moment.
- The key pharmacokinetic parameter controlling exposure of the site of toxicity to unbound (free) chemical is generally the intrinsic unbound clearance, which is controlled primarily by the activity of metabolism enzymes and the unbound renal clearance. Person-to-person differences in volume of distribution, plasma protein binding, and clearing organ blood flow generally do not affect systemic exposure to the unbound chemical, even though they affect other aspects of the chemical's pharmacokinetics.
- When a metabolite of the parent chemical is involved in the MIE, person-to-person and subpopulation differences in the fraction of parent chemical dosage converted to metabolite and the clearance of the metabolite would lead

to differences in the dosage of parent chemical that produced a particular response.

- For some chemicals and some routes of administration inter-individual and subpopulation variability in the bioavailability of the chemical could be an important factor.

Case Study 2 Charge Question 3: *Benchmark dose analysis involves selection of a benchmark response (BMR). The consequences of a perturbation of a key event depend on the magnitude, timing, and duration of the perturbation, which in turn depends on dose. Subtle perturbations of key events may be damped out and have little effect on downstream biology due to the operation of homeostatic processes. Please comment on the current understanding of the magnitude of perturbation of thyroid hormone levels in humans that may be required to lead to adverse outcomes.*

The Panel noted that while a 20% decrease in T4 levels was chosen as the BMR for the triclosan case study, the Agency provided little information to help answer the question as to whether that decrease in T4 levels following triclosan exposure would lead to adverse neurodevelopmental effects. For the Agency to be able to use precursor key events as the basis for a POD for a risk assessment, the quantitative relationships must be understood between precursor events or key events within a causal path leading to a disease or adverse outcome. The thyroid system can be ideal for investigating perturbations that possibly lead to adverse outcomes, because of the body of knowledge concerning its physiological regulation and response to drugs, mutations and malnutritions. However, in the case of triclosan, the experimental data does not seem to fit the typical pattern expected from perturbations on the thyroid system, since TSH levels are unchanged despite the decrease in T4 following exposure to triclosan. While the free hormone concentrations are important to know, a more critical barometer to outcome is the state of the feedback loop that controls endogenous thyroid function (i.e. TSH/TRH levels). Because the pathology (triclosan toxicity) does not appear to map correctly on the physiology (thyroid hormone feedback loop), the pathway for the adverse outcome appears to be missing key information, and thus the Panel recommended that the adverse outcome pathway needs to be revised and refined.

DETAILED PANEL DELIBERATIONS AND RESPONSE TO CHARGE

A. Charge Questions About OPP's Strategic Vision "Integrated Approaches to Testing and Assessment Strategy: Use of New Computational and Molecular Tools"

Charge Question 1*: Does the Panel believe that OPP's strategic vision clearly articulates a sound scientific basis to making the risk assessment process more efficient and informative? To what extent have we described the critical components (i.e., knowledge bases and databases of existing information, predictive tools (e.g., (Q)SAR, in-vitro methods, HTS assays and AOPs) and the progression of events (i.e., research and the use of research) to support the evolution from the current testing and assessment paradigm to an improved IATA of targeted tiered testing based on an understanding of AOPs (i.e., the molecular and cellular events leading to adverse effects at the whole organism level)? Are there any key science issues that the Agency has not captured that are necessary elements to consider in an IATA approach?*

Panel Response

Articulation of OPP's strategic vision

The Panel concurred that the Office of Pesticide Programs' (OPP) strategic vision clearly articulated a sound scientific basis for utilizing the National Research Council's (2007) recommendations regarding "21st Century Toxicity Testing" in a manner that makes the risk assessment process more efficient and informative. Given that the Agency makes more than 5,000 regulatory decisions each year involving thousands of active and inert ingredients, it was abundantly clear to the Panel that efficiency improvement is essential for supporting the Agency's regulatory process. The Panel agreed that vision of toxicity testing and assessment articulated in the draft document, organized around the AOP methodology is a sensible and logical approach towards improved efficiency. The Panel commented that utilizing data from rapidly and inexpensively performed *in silico* and *in vitro* technologies appears to be the most logical way to address the need to improve efficiency, and stressed that overall the Agency has conducted the necessary steps and established an impressive foundation for the transition from traditional toxicity testing to testing using 21st Century techniques and approaches. Given the Agency's acknowledgment that this will be a continual process that will entail "both an evolution and revolution", the Panel made several recommendations to help the Agency to validate the process as it develops. The Panel's input and recommendations also addressed the various challenges that the Agency would encounter while developing this process.

The Panel was in favor of the Agency's articulated use of the Adverse Outcome Methodology to support its strategic vision. The Panel discussed how the AOP provided a structured basis for documented, plausible, and testable processes by which chemicals induce molecular perturbations (primary lesions) that lead to a series (syndrome) of associated biological responses (secondary lesions) at the subcellular, cellular, organ, tissue, and whole animal and when, applicable population levels. AOPs represent a

synthesis of known chemical interactions, metabolism, systems biology, and ecology (OECD, 2011) reflecting the reality that chemicals interact at the molecular level of organization. This interaction often leads to a cascade of events that may or may not culminate at the organism or population level. Therefore, adverse effects observed *in vivo* are the result of many chemical biological interactions as well as the molecular structure of the chemical. The Panel noted that when the molecular perturbations of the chemical are closely linked to observed responses *in vivo*, a correlation model can be derived between the whole animal effect endpoint and chemical structures (One early example is the organophosphosphorus insecticides, which initiate action by chemically reacting with acetylcholinesterase; the OP's as a group have similar apical toxicity endpoints). However, they noted that such direct linkages are not common among long-term complex, multi-biological level adverse effects.

The Panel also asserted that OPP's strategic vision makes the risk assessment process more informative, particularly since it highlights the importance of determining and understanding a chemical's mode of action (MOA) and the key steps in its MOA. They explained that an understanding of the MOA and its key steps including exposure to a chemical and the subsequent adverse apical endpoints after exposure will help a toxicity assessment to be more predictive for humans. The Panel further elaborated that based on the chemical's MOA analysis:

- The most appropriate dose metric can be chosen to conduct a dose-response assessment.
- A decision can be made on whether assumption of a threshold or non-threshold dose response is supported mechanistically.
- An evaluation of whether the adverse effect is relevant to humans can be conducted.
- An assessment can be done on whether children (or other groups) may be more sensitive than adults to the relevant adverse effect.
- An understanding of MOA coupled with knowledge of toxicokinetics/toxicodynamics will indicate when information obtained from one route of exposure is relevant to another route of exposure, though possible point-of-entry effects would need to be identified for each route of exposure.
- Early events will be studied at doses relevant to the general population (i.e., toxicity assessments would not have to be based on high-dose testing in animals).
- Information from toxicity assessments will be more predictive of risks to the general population.

Specific comments on the Agency issue paper

Overall, the majority of the Panel thought the Agency's issue paper was well written and provided a compelling explanation of the strategic vision. However several panel members noted some specific issues that should be addressed to improve clarity of the strategic vision and approach and are discussed in the following paragraphs. In addition, one Panel member mentioned that the notions of problem formulation and relevance-to-human framework were mentioned only briefly in the vision statement and recommended that the Agency more prominently address both of these concepts.

The Panel noted that overall the Agency did a good job of defining terminology. However, they suggested that the Agency may want to make efforts to clarify certain definitions that were referred to by the public commenter's. Specifically, the Panel pointed to one public comment indicating that the definitions of the terms "apical endpoint" and "key event" endpoints were confusing. The Panel also mentioned another public comment that suggested that the Agency replace the term "adverse outcome pathway" with the term "abnormal perturbation pathway". The commenter believed that abnormal perturbation pathway would more appropriately reflect the elements of a toxicity pathway that may impact normal physiological responses in a way that may not necessarily be adverse.

Description of the critical components/modern technologies to support OPP's evolution of IATA

The Panel noted that the "Integrated Approaches to Testing and Assessment" (IATA) strategies concept was not very well defined in the Agency's issue paper. Thus, they recommended that the Agency better articulate the definition of IATA in their issue paper and also that the Agency illustrate the likely nature of a tiered testing strategy for the case studies. However, based on the Agency's presentations at the public meeting the Panel was able to largely interpret IATA as a progressive, tiered-evaluation approach that starts with a hazard-based hypotheses about the plausible toxicological potential of a pesticide or group of pesticides (i.e., a chemical-category). Based upon this hypothesis, existing exposure and toxicity information is combined with computer modeling and data from new diagnostic assays (*in vitro* and -omics) to target further information needs specific for a chemical or members of a chemical-category. The Panel also deduced that the AOP was the Agency's proposed means of approaching IATA.

However, the Panel concurred that the Agency did a very good job of describing the modern technologies that are proposed to support OPP's evolution of IATA. The Panel was particularly impressed with the Agency's efforts in creating database systems to support to their IATA strategies. They emphasized the importance of ensuring that 1): the information being accumulated in these databases is of a high quality, and 2) that the information is being utilized in a standard format similar to what has been achieved with the MetaPath knowledge base. They also mentioned that these databases will facilitate the capture of emerging patterns that can be used to formulate hypotheses and rules.

The Panel noted that as the Agency proceeds to utilize these modern technologies there will be several challenges to address. For instance, one Panel member noted that the introduction of modern technologies has led to a huge growth of new terminology. They further explained that much of this terminology is used with general acceptance in narrow science subfields. However, this terminology sometimes can be confusing when used in audiences of mixed backgrounds and expertise. Thus, they remarked that as this new paradigm/IATA progresses the harmonization of this terminology will become more critical.

The Panel pointed out that historically, integration of new science (e.g. molecular screen and omics) and modern structure-activity (e.g., OECD Toolbox profilers) into the regulatory process has been done in an iterative fashion. Iterative integration of IATA will facilitate the co-evolution of toxicity assessment, and regulatory and stakeholder acceptance. The AOP concept, while similar to the MOA framework, is a new science so its integration into regulatory toxicology will also likely be done in an iterative fashion.

In order to use all relevant mechanistic data from the various levels of biological organization to support assessments, the Panel noted that there must be a means of linking this information to *in vivo* adverse outcomes of regulatory interest. To accomplish this, OPP is embracing the concept of an AOP as the conceptual framework to organize knowledge in order to identify linkages at different levels of biological organization so that more efficient, hypothesis-driven approaches to chemical testing and assessment can be developed and supported. The Panel observed that key to this is the premise that by establishing the causal pathway OPP will be able to screen more reliably for potential adverse effects using data collected from lower levels of biological organization.

The Panel agreed that the AOP including MIEs provides an extremely helpful construct for engaging the broader scientific community in advancing the agenda to develop essential and more predictive tools in chemical health risk assessment (e.g., the MIE; MOA; key events, etc.); this includes communities considering predictive quantitative structure activity analysis tools for data-poor substances (MIE) and those in the research and assessment communities considering data-rich chemicals (key events/MOA). Thus, the Panel asserted that it is a critical area for investment, and key to being more predictive for risks in the human population across chemicals. The principal reason that predictive tools have not been extremely informative in human health risk assessment relates to their empirical nature and lack of mechanistic underpinning (i.e., delineation of the molecular initiating and key events).

The Panel noted that there are large numbers of chemicals and far fewer AOPs, so that once an AOP is understood it becomes possible to identify other chemicals that perturb elements of the AOP with relatively inexpensive and rapidly performed techniques such as HTS, SAR, and microarrays to identify altered pathway and cellular processes. In addition to cost savings from the AOP approach, the Panel noted that this approach will enable risk assessments to account for: 1) simultaneous exposures to multiple chemicals, and 2) and additional species and sensitive subpopulations.

To facilitate planning and perhaps also communication of evolving methodologies in a very challenging and progressive area, one Panel member suggested that it would be extremely helpful to "map" at early stage how the AOPs are envisaged to contribute to more efficient tiered priority setting and assessment strategies, relevant to other evolving tools. It was recognized that this "mapping" of evolving tools would represent only a "snapshot in time" based on what is understood currently and that it will necessarily evolve as understanding increases. However, it could serve as a helpful planning tool in focusing resources on most discriminating aspects. It might also improve more common understanding of interrelationships.

The Panel pointed out that past experience indicates there is the potential to be more discriminating through early consideration of exposure versus hazard-based prioritization to address the question of under what conditions of exposure would hazard testing be minimally investigated. The Agency's issue paper states the Agency is currently participating in an International Life Sciences Institute (ILSI) Research Foundation project to develop a Threshold of Toxicological Concern (TTC)-based approach for the evaluation of antimicrobial pesticide active ingredients (http://www.ilsi.org/Europe/Pages/TF_ThresholdToxicological.aspx). The TTC draws very generically on available hazard data to indicate a level of exposure that would be "without concern" for limited groupings of substances. However, there are potentially other options, including surrogates of exposure potential such as physical chemical properties and use of profiling, both of which have proved to be useful; these options may also inform the evolution of green chemistry approaches.

One concern expressed was that IATA on the surface appears to be a repackaging of the technologies (e.g., *in vitro* testing and QSAR,) which to date have not lived up to their promise. The Panel further pointed out that several decades of alternative methods research have produced better diagnostic methods and important insights into the mechanisms of toxic action. However, these efforts have not produced a significant number of useful alternative test methods. This reality reflects the fact that measuring effects at the *in vitro* level reduces the domain of measurements with respect to *in vivo* risk. Similarly, more than three decades of efforts have produced few QSAR models useful for regulatory applications. What makes this effort more likely to succeed is the *a priori* use of the concept of the AOP, which includes hypotheses-based data gathering and toxicity testing, and tiered analyses.

The Panel called attention to the point that there are inherent uncertainties in the current *in vivo* toxicity test methods. While some of these uncertainties are known, others have yet to be identified or characterized. However, they noted that the level of comfort with the uncertainties of the *in vivo* approach is strengthened by the more than three decades of experience with these methods and the data which they provide. The Panel advised that there will be uncertainties in AOP-based IATAs, and it will be important to characterize the nature of these uncertainties and objectively determine whether they are more or less acceptable than those of the *in vivo* tests that they are designed to augment and eventually replace. The Panel concluded that it is important to realize that uncertainties will be a part of toxicity testing and assessment for the

foreseeable future. The challenge will be to identify these uncertainties, quantify them and develop science-based strategies to address them.

Critical key science issues not captured by the Agency's IATA approach

The Panel noted that although the Agency clearly conducted the necessary steps and has established an impressive foundation for the transition from traditional toxicity testing to testing using 21st Century techniques and approaches, they identified several challenging issues and/or questions that should be addressed. They include:

1. Exposure assessment. What approaches are envisioned for quantification of exposure of the initial biomolecule (the primary target) within a living organism given chemical concentration in diet, water, workplace, etc.? (See Agency's issue paper p. 15 for a discussion regarding how the: SHEDS Multimedia "simulates exposures through residential exposures over different time periods").

2. Dose response. What approach will characterize the relationship between the initial biomolecule exposure to a chemical and the degree of perturbation of pathway elements? Will there be a threshold of perturbation below which apical toxicity is not observed, and if so, how will that be identified? How will idiosyncratic responses (usually immune mediated) be incorporated; e.g. when 0.01% of population shows a different response or the same response at a much lower exposure?

3. Repair/ homeostatic mechanisms. In places, the AOP descriptions in the Agency's issue paper seem to imply that once the chemical initiating event occurs, the AOP operates like falling dominos with the last domino being apical toxicity. However, repair mechanisms and negative feedback within homeostatic mechanisms in the pathways may counteract particular pathway steps, so at some point the dominos may stop falling if the chemical initiating event is insufficient to overcome the repair/ homeostatic feedback. A complete AOP description should identify these off-setting mechanisms and include exposure metric limits below which the AOP does not progress to apical toxicity. This point is addressed in paragraph 3 of the preamble to the charge questions.

4. Interspecies differences in AOPs. As AOPs are developed and elaborated through specific applications, it will be important to identify and incorporate qualitative and quantitative pathway differences.

5. Absorption, Distribution, Metabolism and Elimination data (ADME). The Agency has not done enough to collect chemical-specific data on absorption, distribution, metabolism and elimination (ADME) of chemicals (i.e., toxicokinetics /toxicodynamics) although the Agency has developed a metabolism/metabolite database. ADME data are essential, especially to relate the following:
   - correlation of *in vitro* toxicity with *in vivo* toxicity

- exposure route extrapolation of toxicity assessments (it is important to be able to differentiate point of entry effects from systemic effects so different routes of exposure can be better evaluated)
- species extrapolation of toxicity data (i.e., animal relevance to humans)
- sensitive subpopulations versus general population

ADME data was briefly mentioned in the IATA strategy, but more discussion would be informative. The need for accurate ADME data will be the biggest challenge to implementing OPP's strategy. The Agency will need to determine what databases are available to provide the essential information; what type of *in vivo* or *in vitro* testing is required to provide reliable chemical-specific ADME data; and whether QSAR models are available that would provide reliable information on ADME. One example on which the Agency should focus is the dose-response extrapolation from *in vitro* to *in vivo*. As highlighted by the triclosan example during the Agency's presentation, the *in vitro* cell culture systems could not test concentrations greater than 2 µM triclosan (579 ng/mL) because of observed cytotoxicity, while in the *in vivo* dosing studies, the pups had plasma concentrations of triclosan in the range of 34 to 138 µM (1000 to 40000 ng/mL). This disparity in exposure-response must be better understood to facilitate accurate extrapolation. Development and use of microfluidics and other dynamic *in vitro* culture systems will be useful for modeling and extrapolating kinetics and effects from *in vitro* to *in vivo* scenarios.

7. Delayed responses. The proposed strategy does not address indirect responses where the MIE is dissociated in time from apical toxicity. For example, in Case study 1, propiconazole interacts rapidly with CAR/PXR transcription cofactors to induce CYPs, but protein synthesis takes some time. Other AOPs may involve prolonged time delays that separate the chemical initiating event from the apical toxicity (e.g., maternal DES exposure and cancer in daughters at puberty). The Agency should consider how their IATA strategy will include approaches to identify and deal with time delays.

8. Variable exposures durations. The IATA approach does not address how exposure regimes of varying duration will be quantified (i.e., by using an exposure metric such as Concentration x Time).

9. Synergism. The IATA will need to consider when two or more chemicals simultaneously stimulate an AOP, how will the individual inputs to the AOP be integrated to provide an over-all assessment of toxicity. For example, chemicals acting on the same AOP independently may not stimulate the AOP sufficiently to trigger apical toxicity, but the aggregate stimulation could lead to apical toxicity. As noted in the Agency's issue paper, AOPs are integrated into complex biological networks, and a scenario for chemicals to interactively produce a particular apical toxicity could involve separate chemical initiating events that stimulate AOPs that "cross talk".

Charge Question 2a*: Given that there could be hundreds of AOPs and that varying levels of AOP information are needed depending on the decision being made (i.e., the tier of the risk assessment), please discuss principles for scientific acceptance (i.e., tools for gauging the reliability and robustness) of AOPs and the use of AOPs to move away from chemical by chemical approaches.*

Panel Response

Principles for Scientific Acceptance

The Panel agreed with OPP in predicting that there are likely to be hundreds of AOPs that will be identified in the coming years. As AOPs are proposed, each will need to be evaluated for plausibility and soundness. As articulated by one Panel member, a proposed AOP should be considered to be either a hypothesis or a theory, depending upon the strength of the available scientific evidence supporting the AOP, and the extent to which the AOP has been experimentally tested and found to be consistent with empirical data. In this regard, the Panel believed that an AOP can be thought of as a "living document". In order to gauge the reliability and robustness of an AOP, the Panel suggested that the AOP should be based on the following two principal items (OECD, 2011):

1) Single, defined MIE and key events that is linked to a stated *in vivo* hazard outcome(s).

2) Summary of the experimental support for the AOP which includes:
    a) assessment of the level of qualitative understanding of the AOP.
    b) assessment of the experimental evidence or data.
    c) statement of confidence in the AOP.
    d) assessment of the level of quantitative understanding of the AOP.

For the summary of experimental support for the AOP (principle 2), the Panel further outlined criteria for each. The Panel suggested for the assessment of the level of qualitative understanding of the AOP (principle 2a) the Agency should include documented identification of the following (OECD, 2011):

1) MIE and molecular site of action.
2) Key cellular responses.
3) Target tissue/organ(s) and key tissue or organ responses.
4) Key organism responses; both physiological and anatomical.
5) Key population responses (if required).

For the assessment of the experimental evidence or data (principle 2b), the Panel pointed out that the Agency's issue paper and the two case studies hypothesize that, for a given AOP, the initial molecular event leads to the next key event, with a subsequent cascade of events in the causal pathway to the apical adverse outcome of regulatory interest. However, the Panel noted that the linkages between events must be supported

by data, and the hypotheses must be tested experimentally in order to establish the AOP with some degree of confidence. The Panel also noted that this confidence is typically directly associated with the "weight of evidence" (WoE). Therefore, according to Boobis *et al*. (2006), the principles for scientific assessment of the experimental evidence or data to support AOPs are:

1) Concordance of dose-response relationships for key events and adverse outcomes.
2) Temporal concordance among the key events and adverse outcome.
3) Strength, consistency, and specificity of association of adverse outcome and key events.
4) Biological plausibility, coherence, and consistency of the experimental evidence.
5) Identification of any uncertainties, inconsistencies and data gaps.

The Panel mentioned that the statement of confidence in the AOP (principle 2c), maybe ascertained by addressing the following questions (OECD, 2011):

1) How well characterized is the AOP?
2) How well are the initiating and other key events causally linked to the outcome?
3) What are the limitations in the evidence in support of the AOP?
4) Is the AOP specific to certain sex, life stages, age classes etc?
5) Are the initiating and key events expected to be conserved across taxa?

And finally, the Panel recommended that the assessment of the level of quantitative understanding of an AOP (principle 2d) should include documented quantification of the following (OECD, 2011):

1) MIE.
2) Other key events.
3) Response-to-response relationships required to scale *in vitro* effect(s) to *in vivo* outcomes.

As noted by one Panel member, in assessing the quantitative nature of an AOP, in addition to confidence in the quality of the data, other factors such as demonstrated dose-response relationships, relevance of the tested concentrations, and the duration and pattern of exposure (e.g., area under the concentration-time curve or the maximum concentration) need to be taken into account. An example of the advantages of this is provided by the metabolomic study of Mahle *et al*. (2011), where multiple measurements of response over the dose-response surface allowed the inclusion of lower doses in the study and the characterization of the response of the test animals to the model compound α–naphthylisothiocyanate.

The Panel acknowledged that these principles were addressed to some extent in the case studies but not in the text of the Issues paper. They also mentioned that the

above stated principles are consistent with the general need to ensure the reliability and compatibility of the available information. As noted by one Panel member, the work undertaken by the Agency in establishing a knowledge base for metabolic transformations of xenobiotics could provide an exemplar for the accumulation of data from other areas.

The Panel asserted that it is important to recognize that chemicals often elicit multiple effects in biological systems and can work through multiple mechanisms of toxic action and be associated with multiple apical adverse effects. As noted by one Panel member, PCBs, which induce cancer, also cause developmental toxicity, immunotoxicity, and several other adverse health effects. Accordingly, they remarked that just because a chemical has been associated with one AOP does not prevent it from being associated with additional adverse outcomes and AOPs. They elaborated out that this point was the reasoning behind the aforementioned principle 1 for gauging the reliability and robustness of an AOP, which states that it must have a single, defined MIE linked to a stated *in vivo* outcome.

The Panel noted that consideration of the principle 2b, which required that the AOP entail an assessment of the experimental evidence or data , is likely to vary between a data rich (MOA)and a data poor (predictive AOP methodologies) scenario. Similarly, the extent of required reliability and robustness for AOPs depends on the nature of the decision being made and its likely impact (i.e., priority setting or screening versus in depth assessment or registration).

The Panel pointed out that the defensibility of an AOP will necessarily evolve from consideration of the information on a MIE, other key events, and the apical outcome and will be based on an assessment of the WoE for each event. They further elaborated that what is considered sufficient WoE will be dependent upon the context in which it is being applied. They also noted that more WoE will be required for decisions with greater potential impact. It is likely that regulatory decisions supporting risk management will require consistency across several levels of biological organization within an organism including anchoring to apical effects, especially in the initial stages as the Agency transitions from assessments based predominantly on identification of hazard (i.e., qualitative AOPs) to more predictive approaches (i.e., quantitative AOPs). In contrast, in priority setting, consistency of output of predictive tools (i.e., QSARs, expert systems, *in silico* models) may suffice. Moreover, targeted testing for priority setting may not require complete mechanistic underpinning or ground truthing against apical endpoints.

The Panel believed that it is also important to recognize that the required WoE for any application is likely to be greater in the early stage of the shift from the current paradigm of *in vivo* testing and assessment to a scheme of targeted testing based on a mechanistic understanding of the toxicity of the target chemical. They further noted that as experience and resulting confidence in new approaches increases, it is likely that the WoE required for use in a particular application will decrease. Although the WoE of input is always critical, initial use of new predictive approaches or tools, particularly those which have not been "grounded" against apical endpoints, seems advisable,

particularly in priority setting as a way of increasing understanding and confidence in their use.

The soundness of an AOP can be tested in a variety of ways. One way includes classic mechanistic investigations such as: 1) blocking the occurrence of a key event and demonstrating that subsequent key events (or the apical adverse outcome) do not occur, or 2) enhancing the occurrence or magnitude of a key event and demonstrating an enhanced apical outcome (greater severity, reduced latency, etc.). As demonstrated in the case studies, explicit molecular screens and/or selected omics techniques can serve as "markers" of particular key events. The Panel suggested that another means of testing the proposed AOP is 1) to compare data for chemicals that are considered to cause the same MIE, the same key events or the same adverse outcome, and 2) to assess whether the data for each chemical is consistent for the event(s) proposed by the AOP. An example would be to determine whether there are there other conazoles which have key events that are the same as seen with propiconazole?

The testing of hypothesized causal links between key events in an AOP and on-going critical evaluation and incorporation of new information into the data sets and knowledge base that quantifies an AOP, will result in an on-going refinement and modification of that AOP; hence the term "living document". As noted by one Panel member, as a result of further experimentation and hypothesis testing, some proposed AOPs may undergo significant revision from their initial structure, and others may be determined to be incorrect and will be discarded. The Panel pointed out that this would require periodic re-evaluation of AOPs and incorporation of new information as work proceeds towards developing a library of well-established AOPs.

As noted by some of the Panel members, before data can be used in the development of an AOP there should be a framework for accepting the data based on a set of standards for data formats, and definitions of quality criteria for the data, and a standard regarding the minimum amount of information that should be considered. For example they mentioned that, there is a need to have access to information regarding the quality of the analytical measurements, and of the conditions in which *in vitro* and *in vivo* measurements have been conducted. Objective criteria need to be in place for dealing with contradictory data. However, after careful screening of the information, it should be possible to identify studies that are suitable for AOP development.

The Panel remarked that in instances where there are potentially alternative AOPs, the one which is more representative of the data should be considered. Where dose response relationships have been demonstrated at pharmacokinetically feasible concentrations in *in vitro* assays, the AOP is likely to be more useful in making decisions about the relevance of the pathway. Confidence in individual AOPs will increase with repeatability which is a good measure of robustness.

Some of the major challenges to the approach lie in the prediction of pharmacokinetic behavior, since this can vary markedly between species and within humans and in the establishment of the relationship between the behaviors of systems *in vitro* with those *in vivo*. As indicated by a number of Panel members, the

pharmacokinetics of a compound (particularly if it is non-polar) are likely to be markedly different between animals that use air as a respiratory medium and those that use water. These issues were also brought up under the previous section titled "*Critical key science issues not captured by the Agency's IATA approach"* in the Panel's response to Charge Question 1.

Regarding QSAR models, it was noted that progress is still needed in establishing models that are reliable across different chemical classes, especially in cases where the structural conformation necessary to interact with a receptor has not been established. Similarly, process is needed to enable the "lumping" of AOPs based on different strains and species of organisms, and life stages. This point relates to the importance of the first aforementioned principle for gauging the reliability and robustness of an AOP (Principle 1: (develop AOPs with a single, defined MIE that is linked to a stated *in vivo* outcome) for gauging the reliability and robustness of an AOP.

Use of AOPs to Move Away From Chemical By Chemical Approaches

It was mentioned that AOPs could be used in an effective manner to help OPP move from a "chemical-by-chemical" approach to a chemical-category based approach to assessing chemicals. As noted by several Panel members, AOPs are likely to be useful in predicting the potential for less well-tested chemicals to produce a particular adverse outcome. For example, if a chemical is shown to elicit the initial molecular event of an AOP and some of the other key events, the chemical may be predicted with some confidence to produce the apical adverse outcome, even if the definitive *in vivo* test has not been conducted. This is the basis for the chemical-category approach.

The Panel remarked that the chemical categories approach would entail a grouping of chemicals based on their chemical and metabolic properties. *In vitro* effects and *in vivo* hazard endpoints from the category members that have been tested could then be extrapolated to the untested members of the same category. In order to transition the use of AOPs from a chemical by chemical approach to a chemical categories approach, it would be critical to define the categories. However, without a transparent description of a plausible progression of effects in the AOP, it is difficult to reliably group chemicals into categories and subcategories based on their similarity in toxicological behavior. Thus, the Panel noted that the applicability domain or chemical space of the category should be ascertained by addressing the following questions (OECD, 2011):

1) Which chemicals trigger and which do not trigger the MIE in the AOP?
2) What chemical features increase/decrease the probability of a chemical being associated with an AOP?
3) Are there similar key events caused by the chemicals that could tie them to a common AOP?
4) Are there differences among the chemicals that could lead to sub-categorization?

The Panel recommended that in order to facilitate the transition to this new paradigm, traditional *in vivo* studies used to identify hazards may have to be conducted in parallel with alternative methods. The Panel recommended that it would be clearly desirable to collect data at several levels of biological organization (e.g., apical and molecular) to develop and support the hypothesized AOPs, particularly during the initial stages of the transition. As noted by one Panel member, this can be thought of as "short term pain for long term gain". One other aspect that needs to be explicitly considered in developing AOPs is how best to address their uncertainties relative to the more traditional hazard-identification-based approaches. They pointed out that failure to consider uncertainty would likely pose a significant barrier to the adoption of a new paradigm.

Lastly, the Panel noted that the development of AOPs will facilitate the delineation of toxicological mechanisms associated with primary and secondary lesions. The former are less difficult to model, but there will be a large number of AOPs with one for each type of receptor. However, for some types of toxicity, secondary lesions as well as resulting downstream toxicological outcomes may be similar. For example, different groups of neurotoxicants cause primary lesions in different ways (e.g., modification of ion channel kinetics, disruption of synaptic enzymes or transporter systems, or acting as agonists or antagonists of neurotransmitters). Some of these cause similar secondary lesions resulting from disruption of normal functioning of the central and peripheral nervous systems (e.g., increased or decreased frequency of action potentials), and where this is sufficient (in intensity and/or duration), they cause similar symptoms of *in vivo* toxicity. If similar secondary lesions result from the diverse primary lesions, then it may be possible to combine data from a range of classes of toxicants. The Panel advised that it will be important to develop methods of comparing multiple toxicological profiles to identify areas of difference and commonality between AOPs, which present information in preferably three or less dimensions to enable interpretation and understanding. They should also characterize uncertainties associated with the models.

***Charge Question 2b**: OPP plans to evaluate new tools and information using available criteria and methods such as the OECD principles for (Q)SAR expert systems and the EPA-IPCS (International Program on Chemical Safety) MOA framework. Please discuss these and other criteria the Agency can use to reliably integrate predictive tools earlier in the process.*

Panel Response

As explained by the Panel, the OECD QSAR principles originated from the European Chemical Industry Council (CEFIC) and the International Council of Chemical Associations (ICCA) March 4-6, 2002 workshop entitled "Acceptance of (Q)SARs for Human Health and Environmental Endpoints". The purpose of these principles was to establish a set of ideals which would facilitate the consideration of QSAR models for regulatory purposes. These principles state the following requirements in order for a QSAR to be considered for regulatory purposes:

1) A defined endpoint.
2) An unambiguous algorithm.
3) A defined domain of applicability.
4) Appropriate measures of goodness-of-fit, robustness and predictivity.
5) Ideally a mechanistic interpretation.

The Panel further explained that whilst these principles are flexible, transparency of QSAR models is critical to enable their application. Accordingly, there needs to be a clear understanding of how each model was developed so that they can be meaningfully considered in a "fit for purpose" and "weight of evidence" scheme such as proposed by the AOP approach. The Panel elaborated on each QSAR principle, as follows:

- QSAR Principle 1 (a defined endpoint) ensures clarity in the endpoint being predicted by a given model. Since a given endpoint could be determined by different experimental protocols (e.g., species) and under different experimental conditions (e.g., duration, exposure scheme), it is important to identify the experimental system that is being modeled. For example, the defined endpoint for a "no-observed-effect concentration" derived in from a given regulatory test guideline, relates to a specific effect within a specific tissue/organ under specified conditions.

- QSAR Principle 2 (an unambiguous algorithm) ensures transparency in the algorithm that generates predictions of an endpoint from information on chemical structure and/or chemical properties. Without this information, the performance of a model cannot be independently ascertained, nor predictions reliably reproduced. The need to transparently discriminate between models based on empirical association and those that are mechanistically based (e.g., models with identified MIE in an AOP) cannot be overemphasized. The former can play only a very limited role in regulatory decisions and, for this reason, continued development of AOPs is critical.

- QSAR Principle 3 (a defined domain of applicability) expresses the fact that QSARs are reductionist models which are inevitably associated with limitations in terms of the types of chemical structures, chemical properties, and mechanisms of action from which the models can generate reliable predictions. In the AOP approach, there will be an expanded biological space which will need to be considered alongside the chemical space. It is not straightforward to establish representativeness of data and the applicability domain in high dimensional space, since for more than three dimensions (variables) it is not possible to use plots to visualize outliers, or areas of the data space where there are no measurements.

- QSAR Principle 4 (an appropriate measure of goodness-of–fit, robustness and predictivity) ensures no loss of distinction between the internal performance of a model (as represented by goodness-of-fit and robustness) and the predictivity of a model (as determined by external validation). This principle largely a relates to statistics. The inclusion of noise (i.e., over fitting) can be a major problem when developing QSAR models. Objective criteria need to be in place to define (and when possible explain) outliers. It is important that the data are screened to avoid the inclusion of redundant data.

- QSAR Principle 5 (mechanistic interpretation) is critical for regulatory acceptance. While the intent of Principle 5 is to ensure that some consideration is given to: 1) the possibility of a mechanistic association between the descriptors used in a model and the endpoint being predicted and 2) ensure that this association is documented, this principle has taken on the connotation of mechanistic understanding of the whole model as models with mixed mechanisms of action tend to have lower predictivity. Therefore, it is important to recognize that not all QSAR models will have mechanistic underpinning. Models based on analyses of "patterns" or statistical associations still have value. However, their output needs to be weighted differently. For example, such models may be used as part of a priority setting for further testing or as a very early tier of an IATA.

As noted by the Panel most QSAR models involve an individual biological, outcome but where a set of biological responses is involved, it may be necessary to use appropriate multivariate methods to explore these and to extract a clearly defined response to be modeled. For example, methods such as canonical correlation analysis that can handle two sets of variables may be useful. Similarly, statistical shape analysis (e.g., superimposition) can provide measures of deviations of several multivariate maps from a consensus and the scatter of individual chemicals between maps. In this way, it may be possible to explore the relationships between historical data and new information.

In regards to the EPA-IPCS (International Program on Chemical Safety) MOA framework, the Panel remarked that the assessment of the evidence in support of a MOA is largely based on the afore mentioned Weight-of-Evidence (WoE) approach of Bradford Hill criteria (Boobis et. al, 2006).

The Panel noted that both the OECD QSAR principles and the Hill Criteria outlined in the IPCS framework are well accepted means of evaluating QSAR models and MOAs/AOPs, respectively, used by OPP. Some tools, for example EcoSAR, EPIsuite, and METAPATH, have histories of use and acceptance within the USEPA, as well as, in some cases, other regulatory bodies. Additional development to simplify application of these principles and associated criteria as a basis for broader understanding in considering the adequacy of predictive models and their weight-of-evidence is recommended.

As stressed by the Panel, it is essential to point out that the selected endpoints for the next generations of predictive models will not be those generally associated with regulatory assessment previously. Rather, they are likely to be models for key events along a particular AOP. As a result, these models will not only have to meet the OECD principles but they will also be evaluated for their ability to predict a particular step along an AOP. The integration of new *in vitro* assays and associated information earlier in the IATA process means that it is not realistic to develop test guidelines for each *in vitro* assay before it will be used in some type of assessment. However, the ability of each individual assay or dataset to predict a particular step along an AOP should be evaluated.

Alternative methods (either testing or non-testing) typically target specific cellular or physiological responses and, as such, preclude validation with *in vivo* data by a one-for-one approach. The AOP approach is designed to allow for the use of a battery of assays and subsequent databases that are selected to target particular steps along an explicit pathway. As such, each assay/dataset in a suite of information would inform the next tier of the IATA. The scientific justification of an alternative method or dataset should focus on comparing the test outcome to what is known about the underlying biology as described in the AOP, and thereby, aid the decision-making process. Not all key events in an AOP or all tiers in an IATA may need to be satisfied in order to make a regulatory decision.

Charge Question 2c*: How can the foundation of knowledge generated through historical and on-going reductionist (i.e., making use of information from lower levels of biological organization to understand adverse outcomes), hypothesis-driven, biological experimentation be effectively mined and integrated with the capabilities of omics (including toxicogenomics, transcriptomics, proteomics, and metabolomics) and high-throughput screening to speed the discovery and development of AOPs and their subsequent acceptance as a credible foundation for predictive risk assessment?*

Panel Response

The Panel members interpreted this question in varying ways, but the main responses focused on current and developing technologies to assist in identification and validation of key events in AOPs, and the most effective means to comb the existing literature for existing knowledge on molecular initiators (e.g., agents that cause molecular initiating events [MIEs]) and key events. Predictive toxicology makes use of data-gathering and observational processes to develop hypotheses and models that can be used to make informed predictions about adverse effects of chemicals, even when there are little available experimental data. The goal is to base predictive toxicology on an understanding of the underlying chemical and biological mechanisms behind an observed toxicological response. Although it is rare to completely understand the chemical and biological mechanisms by which a chemical elicits its observed effects, there are data for many chemicals and chemical categories which allow for the development of an AOP which in turn can be used to justify the most advantageous testing. One confounding factor is that the current paradigm from the MIE to adverse outcome does not appear to be amenable to effects where the MIE occurs during developmental stages of an organism. Instead, the paradigm is more applicable to the adverse outcome that may not occur until much later in life. Some thought will need to be given regarding how to handle adverse effects with a significant time lag between exposure and eventual effect.

In considering these general issues, it became readily apparent that there are many parallels to challenges faced by biomedical researchers in academia and pharmaceutical companies investigating the genetic and biochemical basis of disease to identify molecular targets of drugs in cancer, metabolic syndrome, hypertension and other human diseases. In addition, understanding the molecular basis of side effects of otherwise useful drugs can lead to more refinement and modification for more clinically powerful compounds. Therefore, increased interactions between the Agency and basic and applied researchers in industry and academia will be essential to exploit new knowledge gained from these efforts, and to facilitate the development and refinement of new screening methodologies. In addition to the successes that are presented in meetings and published in the open literature, a means to learn from failures will also need to be established, for instance in cases where a large investment in omics technology led to unfruitful lines of inquiry. Lastly, research into complex biological problems such as developing AOPs or understanding disease drives new technology development. Conversely, new technology development drives what kinds of questions one can ask, and will continue to do so.

With these introductory comments in mind, the following approaches should be considered by the Agency as it proceeds with developing and validating AOPs, beginning with a discussion of existing and emerging technologies (Part A) and continuing with suggestions for mining existing and developing databases and expert systems (Part B).

Part A- Existing and Emerging Technologies

1. Determine the relative importance of the proposed key events in AOPs, both for molecular initiator and downstream mediators:

   The propiconazole case study provided an excellent example of how existing scientific knowledge and on-going reductionist hypothesis-driven biological experimentation can be integrated with information generated through toxicogenomic approaches to discover new AOPs. Discovery-based omics approaches enable the evaluation of global changes in gene, protein, and endogenous metabolite expression in response to chemical exposure, and can lead to the identification of novel effects and novel AOPs. As illustrated by the propiconazole case study, discovery-based omics studies can generate hypotheses regarding possible key events involved in the particular adverse outcome of interest. These hypotheses must then be examined and developed further, based on the body of existing scientific knowledge. To further explore the link between a possible key event with the outcome of interest, hypothesis-driven experiments will often be necessary to elucidate the proposed AOP. Examples of such approaches are outlined below, and in following sections.

   For example, rapid advances in molecular biology have facilitated means to identify key events in normal developmental pathways and physiological states, as well as in disease states. The use of small interfering RNAs (siRNAs) in cell lines and even in intact animal tissues provides a "reverse genetics" approach to knock-down the expression of a gene product and evaluates its role in a particular response pathway. These loss-of-function experiments are then complimented by rescue and gain-of-function experiments by over-expression of the gene product in question. Of potential interest to the Agency is the current progress in applying siRNA screens to identify key genes in a particular process (Bakal and Perrimon, 2010; Paul *et al*., 2011) where genome-wide or targeted libraries of short double stranded RNAs or vectors built to express siRNAs in cells are transfected into mammalian cells in culture (or even whole organisms in the case of *C. elegans*).

   While cell culture systems may be useful surrogates for HTS and for more clearly identifying or possibly discovering key events, the effects of chemicals (or representatives of a chemical class) in question likely will need to be evaluated *in vivo*. Use of knockout mice have been useful already for identifying initiating events in toxicity studies, most notably for this Panel the CAR knockout mouse model. Currently, the International Knockout Mouse Consortium (IKMC) (http://www.knockoutmouse.org/) is developing strains of mice in which each gene in the genome is either replaced with a reporter gene such as β galactosidase

or modified such that key exons are flanked by loxP sites for tissue-specific and/or inducible knock out of a particular gene (for Intriguingly, increased speed and ease of identification of mutant alleles, which previously was routinely done only in invertebrates or in zebrafish, has led to several recent successful efforts to identify key genes in disease pathways or physiological pathways via forward genetic screens in mice (Zhang *et al*., 2009). In some cases, the laboratory rat, rather than the mouse, is a more suitable test organism as has been the documented case for many physiological responses. Recent reports have demonstrated the possibility of gene knock-down or inhibition with transgenic shRNA vectors (Herold *et al*., 2008), or vectors that encode targeted zinc-finger proteins linked to transcriptional repressors or nucleases (Cui *et al*., 2011). In addition, embryonic stem-cell-based homologous recombination has been successfully achieved in the rat model (Meek *et al*., 2010). Therefore, the decades-long-advantage gap that the mouse model has enjoyed over the rat for targeted genetic studies may be narrowing.

2. Understand how a particular gene and its product(s) are regulated in response to multiple conditions:

   Once an initiating or key event in a proposed AOP is unequivocally determined, the next important step is to understand how the important gene products are regulated – including hormonal, environmental, circadian and dietary effects on expression and modification, localization and activity of the protein in question. For example, it will be very important to know the relationship between the transcriptome and proteome in AOPs, since some exposures may generally suppress translation, or expression of the mRNAs but be otherwise unlinked to increases in protein expression and activity. Thus, key events in chemical responses may not be transcriptionally mediated, and there will be an increasing need to examine global post-translational modifications. This includes examining the phosphoproteome (Yang *et al*., 2010) or other modification changes, such as acetylation, methylation, ubiquitination, sumoylation and so on, in comprehensive screens (Shi et al., 2011; Seyfried *et al*., 2008; Mischerikow and Heck, 2011). Again the challenge here will be to link those changes to relevance in the particular AOP. Furthermore, ongoing efforts to develop a protein interactome to identify all of the protein's potential partners (Wu *et al*., 2009) and how post-translational modifications impact those interactions (Pless *et al*., 2011), will be useful as a tool to augment interpretation of results of more traditional "omic" approaches.

3. Continue to invest in the development of cell culture models of *in vivo* responses:

   There should be continued investment in establishment of appropriate cell culture surrogates (wherever possible) to facilitate identification or validation of key events in developing AOPs, such as outlined in the siRNA approaches noted above. Well-characterized continuous cell lines or reproducible and robust primary cell culture systems will continue to be invaluable for development of

HTS assays. Increased input from scientists developing three-dimensional or co-culture systems will be extremely valuable in this case. Again, much of this kind of expertise is being developed for drug discovery and testing. For example, long-term primary cultures that respond in appropriate ways to various drug treatments can be achieved using new microfluidic cell arrays with continuous gravity driven flow of media (Lee *et al*., 2007). These and other emerging cell culture technologies should be continually evaluated by the Agency for appropriate applications in AOP development.

4. Understand cell and tissue autonomy issues in evaluating key events in AOPs:

   While cell culture models have proven their worth in many cases, initiation and progress of events in an AOP may involve the interaction of several cell types within a tissue, and effects in one organ under study (e.g., the liver) may then lead to profound effects in other tissues and organs (via increased metabolism of steroid or thyroid hormones, for example). Even when restricting one's focus to a single organ, the chemical in question may have different effects (or no effect) on one cell type relative to another. Identification of a set of cell proliferation genes that are up-regulated in that tissue does not reveal which particular cells are responding. Tissue damage is also often accompanied by infiltration by various immune cells with their own gene signature, but whether this is an adaptive or maladaptive response depends on the insult and the tissue in question. In some cases, it may be instructive to perform omics approaches on single cells or Fluorescence Activated Cell Sorting (FACS) sorted cells from the tissue or organ in question, or use laser capture microscopy to excise and analyze a specific cell type of interest within a complex organ (Okaty *et al*., 2011). Even the liver, usually presumed to be fairly homogeneous, is comprised not only of the predominant hepatocytes, but Kupffer cells, blood vessel endothelial cells, and others (Hoekstra *et al*., 2003).

   Finally, the same chemical may have differential and even opposite effects in different cell types. A useful lesson here can be learned from the so-called selective estrogen receptor modulators (SERMS), which can act as an estrogen receptor agonist in one cell type and an estrogen receptor antagonist in another (McDonnell and Wardell, 2010). These differential responses have been linked to different receptor isotype or ratios of coactivators and corepressors in a given cell type.

5. Understand the molecular basis of both strain and species differences in qualitative and quantitative responses to chemical exposure:

   An important unresolved question in the field is how predictable AOPs will be across species, from ecologically important organisms in the wild to laboratory fish, frogs, rats and mice, and ultimately to humans. Further, even within species, the molecular basis for individual and strain variation in chemical exposure responses remains elusive. Presumably, at least some of this variation may lie in

differential pharmacokinetics due in part to differences in route and period of exposure, biotransformation, and clearance rate. Some progress in exploiting advances in gene mapping for those genes responsible for strain or individual variation is being made (Dayan *et al*., 2009), as is a greater understanding of epigenomics in this area (Koturbash *et al*., 2011). Regardless, close ties with investigators experienced in these issues, across a broad range of organisms, will be essential going forward.

In terms of applying omics technology across species, though, advances in high throughput deep sequencing of mRNAs (RNA-Seq) (Wang *et al*., 2009) will eventually obviate the need to develop species-specific microarrays or rely on unpredictable cross-species hybridization. The successful application of RNA-Seq will need to be anchored to fully sequenced and annotated genomes for the species in question.

6. Develop HTS assays for a spectrum of potential key targets within an AOP:

   Once an AOP is hypothesized, both existing historical information and toxicogenomics information related to the MIE and/or other key events can be used to inform the design of high-throughput screening assays in order to identify chemicals likely to act via that particular AOP. In addition, Panel members discussed whether HTS assays may progress to the point of having real predictive power even *in lieu* of animal based assays, with some differing opinions on this point. As an example, however, decades of work on understanding the entire thyroid hormone (TH) pathway, from synthesis to cellular mechanisms of action, has identified multiple key gene products whose activity is or may be amenable to adaptability to HTS approaches including:

   - TH synthesis: Sodium/iodide symporter, thyroid peroxidase.
   - TH action: thyroid hormone receptor- binding, coactivator/corepressor interactions, transactivation (with attention to specific isotypes and splicing isoforms).
   - TH transport: Organic Anion-Transporting Polypeptide (OATP) and Monocarboxylate transporters (MCT) class cell membrane transporters, cellular export proteins, serum binding proteins.
   - TH metabolism: Type I, II, and III deiodinase activity; non-specific effects on liver P450 mediated metabolism.

Finally, while reporter gene assays have been a tremendous asset to basic transcription factor biology including development of HTS assays, newer methodology is emerging to allow determination of the expression of multiple genes at once and in a high throughput manner (Brenan *et al*., 2009; Spurgeon *et al*., 2008). Specific gene expression signatures, such as those uncovered via investigation of AOPs outlined by the Agency, may one day supplant reporter genes as readouts of altered transcription factor activity in HTS assays.

Part B- Mining Existing and Developing Databases and Expert Systems

Beyond continued incorporation of new technologies, improved methods for mining the existing literature for information useful for AOP development is a recognized need by the Agency. Existing historical scientific information can be used to speed the acceptance of a proposed AOP for use in predictive risk assessment. For example, training sets of chemicals shown previously (i.e., historical information) to cause the adverse outcome of interest, and one or more key events, can be further characterized with omics techniques to document other key events in the AOP. This work will inform the hypothesized AOP, and will likely lead to refinement of the AOP. As the AOP is refined and found to be supported by diverse sets of data on multiple chemicals known to induce the adverse outcome, the WoE supporting the validity of the AOP increases. In addition, these training sets of chemicals acting via a given AOP can be used to test and validate the utility of high-throughput screening assays designed to identify chemicals acting via that particular AOP.

The Panel recognizes the efforts to develop such databases as outlined and presented by the Agency at the meeting. These databases rely on expert systems for identifying key findings to varying degrees; continued involvement in literature review by basic and applied experts for inclusion of the best controlled and executed studies will be most useful. Annotation is encouraged whether GLP and "standard protocols" are followed or non-GLP studies that also include important fieldwork findings.

Overall, the Panel believes that the efforts by the Agency in this area are both noteworthy and appropriate, and the following comments and suggestions are provided for improvement and refinement of current approaches:

1. Computational biologists apply computer science, mathematics, and statistics to the study of biology. The field of computational biology is thought of as having two distinct areas:1) knowledge discovery, which includes data-mining and the elucidation of patterns from experimental data (an approach that is used in bioinformatics); and 2) simulation-based analyses, which use *in silico* approaches to develop predictions that can be tested *in vitro* and *in vivo*. Simulation-based analysis has direct relevance to AOPs and IATA.

2. Omics and high-throughput screening are experimental-based methods whose results can aid in identifying bioactivity signatures. Bioactivity signatures have the potential of being a useful way of forming chemical categories. Bioactivity signatures are likely to take the form of knowledge bases in which the data are organized in terms of ontologies that permit automated knowledge extraction from the data.

3. Critical to the development of AOPs is the ability to synthesize the required knowledge of chemical interactions, metabolism, systems biology, and ecology. All of these topics are large and diverse areas of science which are difficult to integrate due to the high degree of specialization in the respective fields. Since

AOPs represent current knowledge they should be thought of as living documents that will change with time. Technology must be developed and used that will allow for real time interactions of diverse sets of experts located throughout the world while at the same time keeping an audit trail of all past interactions. This is perhaps best met by using wiki-based technology.

4. Since MIEs are typically chemical interactions with receptors and enzymes, an inventory of possible MIEs should be developed and linked to Google-like technology that allows searching by use of key terms. Since key events are often induction or inhibition of protein expressions, an inventory of possible key events similarly should be developed.

   Related to this point, it is imperative that data generated via HTS as well as omic derived data (continue to) be deposited in public databases for ready access by the wider scientific community which can spur useful confirmatory or alternative interpretations of the data, and synergy with ongoing efforts in disease pathway studies and drug development. Likewise, access to HTS data for important proposed initiators (nuclear receptors such as the ER and AR, or enzymes like aromatase) using very large combinatorial chemical libraries can help define the chemical space of all classes of chemicals affecting that protein's activity.

5. The Panel's discussions thus far has focused mostly on starting from observations in animals treated with chemicals, and describing AOPs. Adverse outcome pathways can also be identified based on what we know from human biology and human disease. A recent paper by Gohlke *et al*. (2009) illustrates the power of combining information from human disease pathways with chemical toxicity pathways. Gohlke *et al.* used gene-centered databases to develop a network of complex diseases and environmental factors and key molecular pathways. These authors focused on a couple of groups of human diseases: metabolic syndrome and neuropsychiatric disorders. Their analyses generated several hypotheses regarding novel biological pathways leading to these diseases. Efforts like this, to use what we know about human disease and chemical toxicity, can be very valuable in discovering and developing new hypothesized AOPs, which can then be further explored and tested. The added benefit of using what is known about human disease to develop an AOP is the increased confidence one has regarding the human relevance of the proposed pathway. The EPA should encourage further development and use of this type of analysis, utilizing the knowledge bases EPA and others are populating with chemical toxicity information together with information on genes and pathways involved in human disease states, to explore the interactions and connections between disease and chemical toxicity pathways and networks.

6. Given the numerous examples where biomedical research in particular has experienced parallel challenges and concerns as omics technology is more increasingly embraced, enhanced interaction between basic scientists and toxicologists in multiple research environments will be needed. Existing Agency

and international workgroups including those of IPCS and OECD and satellite symposia at various meetings have been productive in encouraging relevant coordination, and should continue to be encouraged and supported by the Agency and its partners. Active participation in such working groups by both basic and applied scientists, working in governmental and academic laboratories as well as in industry (including both "Big Pharma" and pesticide and herbicide manufacturers), sharing both their successes and failures, would be of great value moving forward at this still relatively young stage of this effort.

Furthermore, the Agency could possibly benefit from looking across disciplines and engaging others such as engineers and computer scientists in addition to biologists. The AOP is in essence the engineering discipline of process control with perturbations that either result in the system returning to steady state, reaching a new steady state, or becoming unstable and therefore dysfunctional, principles dealt with on a daily basis by engineers. Other disciplines' approaches to modeling naturally occurring phenomena could be useful for modeling the biological pathways. For example evolutionary modeling with survival of the fittest may provide insights, or the recent advances in deep question and answer technology as showcased recently by the IBM Watson on Jeopardy (www-03.ibm.com/innovation/us/Watson). The human mind can be limited in its ability to parallel process in a way that can understand the various interconnectedness of the relevant biological pathways. However, the deep question and answer technology provides a way to interpret and reason through huge amounts of stored data to build confidence in the prediction of an answer to a question, which for IATA, will be predicting the potential for adverse outcomes from a wealth of omics and other in vitro data.

.

7. Lastly, the new IATA paradigm being developed by the Agency will require a new training regimen for graduate students/post-doctoral fellows entering the field. These individuals will need to be broadly trained biologists with a deep appreciation and understanding of genetics and new genomic tools and computational approaches. The same is true for biomedical trainees as well. Funding mechanisms for this new training ideal could then be developed across agencies with this common goal in mind.

Charge Question 3*: OPP is at the beginning stages of contemplating metrics to demonstrate success in achieving our vision. What is the Panel's initial thinking regarding the concepts, considerations, and factors OPP should take into account as we evaluate our progress? What methodologies are best suited to measure success in achieving our vision? For example, what methods would you recommend to quantify the development of AOPs and the qualitative and quantitative application of AOPs in risk assessment?*

Panel Response

Concepts, Considerations, and Factors to Take into Account as Progress is Evaluated.

The Panel focused their responses on issues that they considered important for assuring the quality, plausibility, and utility of the developed AOPs for risk assessments as the Agency progresses toward achieving their vision.

It is essential that any AOPs identified that might form the basis of risk assessments, be scientifically defensible. The AOP must describe a pathway that has key events that are known to be feasible in the normal biochemical and/or physiological processes. There is substantial merit in thinking about AOPs as being perturbations of normal processes and not new pathways that are being created by the presence of the toxicant. If the AOP is describing a pathway that is not just an over-stimulated or under-stimulated normal pathway (e.g., an accumulation of excessive amounts of some biochemical intermediate that then initiates an unusual pathway), then this cause-and-effect phenomenon should be confirmed in independent experiments to demonstrate feasibility. Therefore each AOP should have steps or key events that are part of a known, or at least reasonable, progression of events that could lead from the initial molecular target to the apical observation. There needs to be a solid understanding of the basis of normal physiology (e.g., nervous system function, reproductive processes, differentiation of tissues). The AOPs identified must not contradict any steps of normal biological processes, since they would need to be biologically plausible. Even if some steps are not known with certainty, the overall process must agree with what is known about the particular biology being considered. It is understood that some of these AOP's would be networks and not linear pathways. The AOPs identified must be consistent with any rate limiting steps or other considerations that indicate what magnitude of effects in any key events are likely to impact the overall progression of the pathway.

In order for an AOP to be utilized in risk assessment, it must define some critical toxicological effect and not just an auxiliary phenomenon, which does not contribute to the most sensitive, or at least among the most sensitive, systems impacted by toxicant action. Therefore, a dose-response consideration must be provided for any group of AOPs that are under development for a given chemical or class of chemicals, so that those pathways that are impacted at the lowest doses would be deemed the important pathways and those that are impacted only at higher doses would be deemed of limited or no importance. While there is merit in observing effects experimentally at high doses to identify hazard, if the doses at which such effects occur are extremely unrealistic, then

they are probably occurring at levels of toxicant that are greatly overwhelming the defenses and homeostatic mechanisms, and thus would not reflect what pathways might be functional at more realistic dose levels. Such high dose pathways may not be expected to occur and therefore are not representative of what the toxicant would do at more realistic levels. These realistic dose-response relationships might be more difficult to design and interpret with *in vitro* and omics data. Attention might be diverted away from some other pathway that might be operational at more realistic levels of exposure. If exposure information can be generated (or in the case of new compounds, estimated), it should be considered in design of *in vitro* or more limited *in vivo* studies to make certain that extraordinary toxicities are not blinding scientists to more relevant toxicities. If any AOP is identified using extremely high dose data, the pathway needs to be confirmed at a lower dosage and the concentrations or level of impact occurring at each key event needs to be confirmed so as to identify adverse impact that occurs at a given level of inhibition/binding/etc. In other words, how much of an impact on a key event is adverse and how much is just normal biological compensation and within the range of homeostatic mechanisms.

Any AOPs identified should be consistent with any solid epidemiological or case report data that link the chemical or chemical class to specific adverse outcomes. Caution is urged about critical evaluation of the validity of the linkages between a chemical and a toxic response identified in epidemiological studies or case reports since these could easily have insufficient reliable information for accurate conclusions. In addition EPA will need improved epidemiology approaches using enhanced remote sensing tools and geospatial analysis to further implement this new AOP/IATA approach. New GIS geospatial approaches should be developed in the future as this will help better relate epidemiology data on exposure/effects with MOA and AOP approaches.

There must be assurance that solid experimental design was employed in any of the data sets used for developing the AOP. Controls must be appropriate, and care must be given to make sure that controls and treated samples coming from treated animals are truly comparable. This would be particularly important in any developmental endpoints, where a developmental delay might be induced by the toxicant. If this were to occur, then differences might be observed between the controls and treated animals that are attributed to the effects of the toxicant when the differences might be only the differences in the effective "age" of the animals when sampled.

If a novel chemical is being assessed, it is essential that any pathways identified be pertinent to the chemical. This might be very difficult if the chemical was truly novel and there was little to no information on its effects available from *in vivo* experiments. Many changes can be observed in the omics technologies because of the immense amount of data generated. For example, the systems biology approach can identify numerous pathways that are different between controls and treated. Whether these are truly meaningful in the intact organism cannot really be known from only the omics data sets and the computational analyses of these. Follow-up with targeted *in vivo* experiments to verify any highly novel AOPs should be done before moving these omics-based conclusions into risk assessment.

Initial, AOP identification should be done on chemicals with rich data sets, as has already been initiated. Concordance of AOPs should be verified by *in vivo* results. Peer review of the initial AOPs, including data sets used, rationale for AOP identified, and consistency of data sets with postulated pathways, should be conducted before any AOPs are considered for use. Periodic testing of the knowledge bases with peer review workshops or "round robin" events evaluating compounds with complex AOPs should be performed. For example, data on compounds that have target organs other than the liver (e.g., paraquat or neurotoxicants) should be used for these evaluations. Currently, omics is targeted primarily on liver endpoints. It is likely that omic evaluations of other tissues will be necessary. Perhaps HTS should help in identifying tissues where additional omics can be carried out. In addition, "negative controls" should also undergo some form of evaluation. A "round-robin" or "blind" process can also be used to assess the precision of the AOP. In addition, given the similarities of omics interpretation with histopathological evidence, peer review of omics data with a "blind" evaluation should be implemented before compositing the data into a knowledge base.

The transition in testing to approaches based on omics and other high-throughput technology based methods needs considerable up-front time, coordination and team-work. The Agency will need to recruit other capable partners to work in concert on developing and standardizing approaches. Evidence of progress towards achieving the vision laid out could be provided by verifying that multi-institutional efforts can determine similar key events and adverse outcome pathways using the same known chemical with a recognized, well-defined apical adverse event for "grounding". Careful consideration of reagents and variables, (e.g., animal model and care, diurnal regime and seasonal effects dosing conditions, chemical source and purity, cell lines used, ambient conditions, control conditions, etc.) has to be maintained across such studies. However, if multi-institution confirmation of key events and pathways for known toxicants could be established early on in standardizing a large suite of approaches, it would be a major indicator of progress while, attaining converging analytical strategies. If members of other classes of chemicals with distinct defined toxic endpoints could be similarly evaluated for consistent findings, further confidence could be gained in strategies being developed. Further efforts in these same laboratories/agencies could then work forward by evaluating another chemical/chemical set but with unknown/unspecified apical endpoints. If these approaches are to address the bulk of uncharacterized chemicals in commerce with relatively little information known about their toxic potential, then confirming that multi-institutional efforts can determine similar key events and pathways is paramount in building confidence that these types of approaches have promise for reaching that goal. While it can require great effort to "wade" through the numbers of differentially expressed genes, transcripts, proteins or metabolites to couple with a known adverse outcome and, if the chemical is an unknown or it has an as undefined toxic endpoint, determining "relevant" omics changes will be much more complex. Obviously, these are time and resource heavy activities. Determining that multiple institutional efforts come up with similar answers on the same chemical would, however, be a big step in confirmation and validation of the overall strategy.

Data management and quality assurance are key issues in the process of creating AOPs. EPA has developed multiple standardized "methods" and "guidelines" for analytical chemistry and other processes needed for scientific evaluation. Quantitative and qualitative "standards" will be needed in creating these documents and procedures. HTS, QSAR, and omics guidelines and specific software method procedures should be developed that allow transparent and objective evaluation of data. For example, clear directions to conduct sequential GO analyses, KEGG or K-means clustering or any other pathway software should be vetted and delineated to reduce uncertainty of interpretation. Regarding data management, the model provided for Metapath, DER Composer, and Effectopedia, where international input and collaboration has been coordinated by setting criteria for data entry, retrieval and management, is advised.

In regards to registration data reporting requirements, there will be more reliance on modeling and HTS methods and less data on animal models and field monitoring data. Thus, there will be a need to have targeted data generation by registrants that help feed the models to improve them. Experience at the National Oceanic and Atmospheric Administration (NOAA) in weather forecasting indicates the importance of supportive observational networks measuring key model parameters to help feed the models to improve them. Thus, the Agency will need to have data reporting requirements from registrants that support models (e.g., metabolites, fate and effects). The Agency will need to define what may be the equivalent type of observational networks to better define and support IATA and AOPs. This will be of particular importance for *in vivo* data and field validation data required of registrants for ecotoxicology data.

Other factors EPA will need to take into account for this approach are as follows: (1) the need for more standardization of new methods selected in a rigorous and timely fashion, (2) the implications of this process for both data transmission and data storage requirements, and (3) reliance on expert Panels as needed to provide input on highly technical and very specific research areas to analyze controversial and divergent findings about the use of these new technologies and time and consistency of approach for AOPs in a generic sense. There will be an expectation that these HTS and AOP approaches will lead to a more streamlined process, given that AOPs are living documents, that may be subject to rapid change as more and more data are generated, analyzed, and developed within the AOP process. Thus there will be more continued updating and refinement of the AOPs as part of the standard approach. This process of continued learning regarding an individual pesticide will lend itself to broader stakeholder input and transparency as the process develops, refines and matures.

Establishment of this approach for mammalian testing should continue to be the primary focus; however, parallel efforts should be undertaken at other taxonomic levels in order to meet the full spectrum of ecotoxicology data requirements in transitioning to this approach. The fish research described by the Agency is a very good next step particularly since vertebrates (mammals and fish) will share many common modalities at the genomic, transcriptomic and proteomic level. There is less certainty of these common modalities on the metabolomic level. Research at NOAA on metabolomics has shown major differences between marine mammals and fish for metabolites in urine and

blood (Bearden, 2011; Borpujerdi, et. al, 2009; Schlock et. al. 2010; Morrison et. al, 2007; Viant et. al, 2009,). Metabolomics will need to continue to develop at parallel levels to become a meaningful part of this new paradigm.

In determining how these new methods using key events along an AOP for individual and classes of similar pesticides will be developed, EPA may wish to examine other similar types of hazard and risk based approaches. In Hazard Analysis and Critical Control Points (HACCP) for food safety those critical control points (CCPs such as Time – Temperature Controls) are identified where the importance of these factors in the conveyance of infectious food borne illnesses is known. For AOPs, these critical control points or key events/multiple key events along pathways will need to be defined throughout the lifecycle of chemical exposure and effects in much the same way as are the pathways of food production (harvest, transport, processing, cooking and consumption). Microbial testing is generally used to ensure the HACCP approach is working, and imperfect generalized indicators are often used to assure the health and well-being of the food supply and those who consume that food supply. Developing those new omics tests and general integrative indicators for chemical exposures/effects for AOPs will be a future challenge.

There are various modeling efforts to help support the omics approach that are in various stages of development and maturity. They will only become better as they are further developed, populated, used, refined and better integrated across AOPs. Their development is a very good step forward and allows for sharing of information among researchers. QA/QC measures of the data included in these shared databases are important. The Agency has carefully and thoughtfully considered these needs and is encouraged to continue the full development of essential database efforts.

The Agency will need to define the norms for controls for any new assay before using it on a wide scale manner in the AOP process. The Panel discussion in the case study with propiconazole clearly illustrated this as there were questions about what constitutes an adequate control. This was also raised in the public comments presented by industry. There may be a need to include some round robin testing among the Agency and other federal labs, universities, and others before implementing these new omic approaches to validate their soundness and reproducibility. Future testing of compounds could include those that are a “slam dunk” for not passing the registration process (e.g., pesticides that have been banned for use), a “slam dunk” for passing the registration process (e.g., high margin of safety) and one or two pesticides that may fall in between. This would be very insightful toward determining how useful this approach is and how does it compare cost wise and does it reach the consistency of results as do traditional lines of evidence. As these new assays are developed as multiple lines of evidence for an AOP, it may be important to consider having Panels of experts to review these results of wide scale usage of these assays in common practice in order to help EPA define their efficacy and reproducibility of results.

Under this new approach, AOPs become the conceptual model to solve complex problems involving multi-taxa, multiple levels of biological organizations. Attempts to

integrate these new approaches via both *in vitro* and *in vivo* testing protocols will provide more in-depth understanding of the risk assessment issues faced by the Agency. A timeline for implementation of new HTS methods and AOP/IATA approaches should be developed with some adaptive management approach that includes key milestones to adapt the framework. Significant milestones may include the following: (1) application of new technology to mammalian testing, (2) adaptation of the strategy to key taxa among non-mammalian vertebrates, and (3) adaptation of the strategy to key taxa among invertebrates. Ultimately a comparison of assessment of a compound using the current conventional approach versus the new IATA and AOP approach should be done at some step in the process; such a comparison would be very informative by providing a completed pathway to follow for future AOP development for other compounds. It would be important to do this for a model compound in several chemical categories, as well as with several model compounds in particular chemical-category to provide a spring board for further implementing this process. Comparisons made using both the current conventional ecotoxicology animal models and this new IATA/AOP framework should include: (1) comparisons of the ecotoxicology and human health risks using both frameworks, (2) cost comparisons, and (3) comparison of time savings.

In development of a process for implementation, the Agency should identify what key steps in the current registration process have the greatest data needs and attempt to apply this to help identify how these new omic methods can be most helpful in addressing such data needs. This approach will gain greater acceptance among registrants and will provide perhaps a more cost effective and holistic method to address these data needs. It would also provide a phased method of implementation within the current methods being used.

Development of omic tools for non- vertebrate species and marine mammalian species should and can be done by EPA in concert with other federal agencies and academia, that have begun to embark on the development of genomic, transcriptomic and metabolomic approaches to toxicology research. For example NOAA has developed dolphin cell lines that may provide new information in toxicological testing for these protected marine mammals. Developing omics technical working groups across EPA, other federal agencies and academia for identified major taxa needs, such as fish, shrimp, mollusk, polychaetes, amphibians/reptiles and marine mammals, should be developed and conducted in parallel with efforts on rodents and other mammals. The different omics taxa working groups should have annual workshops supported by EPA to share progress and results across taxa in developing AOPs as new methodologies are developed, adopted and implemented.

Standards for developing positive and negative controls for inclusion in new omics test protocols are needed, particularly with metabolomics. Agencies such as the National Institute of Standards and Technology (NIST) have in their mission the development of new standards needed for science, technology and commerce. A list of standards for genomics, transcriptomics, proteomics and metabolomics will be needed for application of this new EPA strategy. The absence of these standards may limit the use

of this technology for omics analysis as it will be difficult to compare data throughout the world, inhibiting the rapid adoption of these methods.

The Agency is to be commended for an excellent beginning to a focused effort to engage stakeholders in this new AOP /IATA approach by in both soliciting their input and by providing additional educational outreach to better inform key constituents. One key group the Agency should consider for public outreach is the ecotoxicology educators, particularly those involved in graduate education in toxicology and ecotoxicology. This AOP/IATA approach requires new training in more transdisciplinary skills including more in-depth use of QSARs in molecular testing end points and integrated modeling across increasing levels of biological complexity along with bioinformatic approaches to statistical analysis. Informing educators of these approaches will assure that graduate school curricula change to better reflect educational needs of the Agency as well as the recruitment of next generation researchers with needed talents to help further develop this new paradigm. EPA should continue to support both predoctoral and postdoctoral fellowships in this area of needed research skills by collaborating with academic partners to develop the competent work force that will be needed by Federal and state agencies, academia and industry to implement and develop this new research framework.

In summary, progress with developing AOP's can likely be achieved if: 1) if the AOP's are grounded in a solid understanding of the normal biochemistry and physiology of the organism, 2) there is a well-considered assessment of what key events are impacted and the dose-response associated with impacts on these key events, and 3) there is assurance that the data sets upon which the AOP's are based are carefully designed using appropriate dose/exposure levels and that comparisons within the experimental groups including controls are valid.

Methodologies best suited to measure success in achieving vision

OPP lists goals of the IATA strategy as increasing the efficiency, effectiveness, and accuracy of the testing and assessment process by enhancing their abilities to target the effects of concern and to base decisions on the information most relevant for the assessment. While the end goal and ultimate measure of success will be the ability to provide an estimate of risk from exposure to contaminant solely on the basis of *in vitro* and *in silico* techniques to predict the AOP, OPP will need to define short, medium, and long term goals to maintain progress in shift to the new risk assessment paradigm. It is important to landmark incremental steps in the process in order to keep the momentum focused on reaching the ultimate vision. The metrics used to measure progress and success will have to consider the competing priorities among OPP's stated goals, so that a metric that indicates increased efficiency also still assures that effectiveness and accuracy of the assessment have not been adversely affected, and vice versa.

The AOP concept can be viewed as a new model for generating the data/information for hazard assessment and utilizing it to generate knowledge for risk assessment. New models are often assessed using calibration and validation metrics. Briefly, calibration is determining how known data sets fit into the model while

validation is making predictions and then collecting the data to determine the accuracy of that prediction. In this case, calibration would be done by picking a chemical that is well characterized and working through the AOP to see if the same quantitative measure of risk is achieved. Further, was that result achieved faster or cheaper? Does it determine a different assessment of risk and if so, is it a better assessment? Validation is picking a new chemical, conducting the traditional and AOP approach concurrently, and determining the comparability of the results.

.

In contrast to some efforts that come before the SAP for guidance, there is no near-term deadline approaching for development and implementation of this new paradigm. This is a good, because the overall objectives are really monumental. However, a schedule will need to be developed that sets goals and benchmarks of progress. The Panel generally agreed that the transitional approaches being considered in this overall effort will take a decade or longer. During the initial phase of the transition to the new approach, it is likely that more animal testing will be necessary instead of reduced animal use (a stated goal) prior to further development of appropriate *in vitro* approaches. Moreover, use of animal models may continue to be needed for chemicals that act less at a cellular level but require intercellular signaling or complex tissue connections. Careful consideration of studies, recruitment of partners, harmonization and transparency of approaches, is essential for building stakeholder confidence and for educating all involved in the long-term benefits of these new approaches. Eventually, evidence of efficient and timely evaluation of multiple chemicals will be demonstrated by numbers, i.e., how many chemicals can be evaluated over a unit of time and cost expenditure, relative to past experience conducting traditional evaluations.

Progress should be evaluated in the context of how quickly the strategic vision translates to a tactical application. Progress should be measured in regard to how quickly information technology capabilities can be established that are needed to handle the information required to develop an AOP and move the AOP from a qualitative to quantitative assessment.

In section 1.3 of the issue paper, if the word pesticide was replaced by the word substance or chemical, the paragraph could be adopted by other offices, agencies and countries. So OPP should actively engage others in moving the paradigm shift forward. Success stories, even small ones less relevant to pesticides, are going to be important to measuring progress as well as maintaining the enthusiasm and momentum for the paradigm shift. The ability to develop inventories for MIEs, key events and apical endpoints of regulatory interest may be an early measurement of progress. The ability to identify HTS for key events for specific AOPs may be an intermediate measurement of progress. The development of *in vitro* databases for such key events may be a measurement of progress.

In conclusion, the Panel recommended that OPP regularly ask the following questions to determine how to measure and gauge the success of their ongoing IATA strategy:

1) How many regulatory decisions utilized SARs or QSARs to identify relevant pathways to risk? (If the numbers of decisions are rising, this would indicate that the methodologies are useful to the process).

2) When implementing IATA in a regulatory review, what is the average amount of time and the "variability" in the amount of time between initiation of chemical review and its risk decision? (If the amount of time is decreasing, this would indicate growing success).

3) How many decisions incorporate an AOP in the risk assessment framework for the chemical? (If the numbers of decisions are rising, this would indicate that the AOPs are useful to the process and that they should be utilized).

4) What is the number of decisions that occur within a chemical class rather than a chemical alone decisions?

5) What is the number of new standards incorporated into decision processes and how many of those standards are capable of being utilized by multiple Agencies?

6) What is the public perception and acceptance (measured via surveying) of the new technology over time? (Buy-in by stakeholders in the process is another possible measure of success.)

7) What are the numbers of studies conducted utilizing animals conducted over time? (This number should be expected to increase in the short-term and decrease over the long term).

8) How many added knowledge components are being captured? (For example the Agency should be counting the cumulative number of MIEs discovered over time. If the number of MIEs is rising, then the Agency will begin to saturate its understanding of them.)

B. Charge Questions for Case Studies

Case Study 1: "Use of Toxicogenomic Technology to Inform the Risk Assessment A Case Study: Propiconazole"

Case Study 1 Charge Question 1*: Please comment on the use of omic and related technologies employed to develop and link Key Events to create the MOA/AOP of propiconazole. Was the resultant MOA/AOP for propiconazole logical and scientifically sound?*

Panel Response

The use of omic and related technologies to develop and link key events to create the AOP of propiconazole demonstrated how powerful and useful these technologies can be. The benefits of discovery–based, hypothesis-generating omics studies were clearly demonstrated with this case study. A good intersection of transcriptomics, proteomic, and metabolomic data was observed and generally helped tie together and support the proposed AOP. Propiconazole was an appropriate choice for this test case since a wealth of toxicological data collected over several years already existed, and "read across" comparisons to other conazoles as well as to other compounds with known hepatic toxicities such as phenobarbital can be done. The Panel felt that overall, the proposed AOP for propiconazole was logical and scientifically sound, with the caveat that so far, the AOP best described only propiconazole, at the doses used, and in mice. However, the proposed key events should continue to be tested experimentally, and modified and refined as needed.

Based on previous whole animal testing and *in vitro* assays, a preliminary MOA/AOP for propiconazole induced liver tumors in mice was proposed that included:

1) Induction of CYPs and other enzymes,
2) Liver hypertrophy,
3) Mitogenic stimulus resulting in increased cell proliferation,
4) Higher levels of cell turnover,
5) Increased likelihood of spontaneous mutations, and
6) Clonal expansion of transformed cells to form altered foci, adenomas, and carcinomas.

Using a protocol that included dosing of propiconazole at carcinogenic levels, and examining microarray, proteomic, and metabolomics data at different days, the initial AOP was refined to identify multiple novel key events and a potential initiating event involved in the carcinogenic action of propiconazole including:

1) CAR/PXR activation leading to enzyme induction, increased all-trans retinoic acid metabolism/decreased all-trans retinoic acid levels, increased cell proliferation/altered cell signaling, and tumors.

2) Inhibition of CYP51 (lanosterol 14α-demethylase) leading to dysregulation of cholesterol biosynthesis/metabolism, increased mevalonic acid levels, increased cell proliferation/altered cell signaling, and tumors .
3) CAR/PXR activation leading to enzyme induction, ROS (reactive oxygen species) generation/oxidative stress, mutation, and tumors.
4) Suppression of apoptosis [MIE not yet identified].

One advantage of the omics data in the propiconazole case study was that it allowed for more insights into mechanisms of toxicity. Specifically, the AOP defines a series of key events. These key events were based on apical biological observations. These AOP key events allowed for hypotheses to be proposed based on gene expression profiles. For example, via the AOP the differentially expressed genes (e.g., CYP1a2, CYP2b, and CYP2c genes) were hypothesized as commonly induced via another key event (*i.e.*, CAR/PXR nuclear receptor activation). The differential gene expression reveals that these expressions are common to both CAR activation and CYP induction, thus linking these key events. These concepts were nicely demonstrated in more depth in slide 12 (Figure 1) from the Agency public presentation entitled "Integration of Toxicogenomic Technology to Inform Risk Assessment A Case Study: Propiconazole" with the key message contained in slide 30 (Figure 2) illustrating how using omics technology can assess key events.

Data used to derive key events in the propiconazole MOA: in vivo data from studies conducted under the same conditions used in the cancer bioassay

	In vivo					In vitro	In vivo and/or In vitro
Key Event	Apical	Transcriptomic	Proteomic	Metabolomic	Other		Biochem/Molecular
Propiconazole parent						X	
CAR/PXR activation		X					
Cyp2b &Cyp3a induction/activity	X	X					X
Increase hypertropy	X						
Increase oxidative stress		X	X	X			X
Induce mutations		X			X		
Increase hepatic atRA metabolism/Decrease hepatic atRA levels		X					X
Dysregulate cholesterol biosynthesis and metabolism		X	X	X		X	X
Increase mevalonic acid levels		X				X	X
Increase cell proliferation	X	X				X	
Decrease apoptosis		X					
Increase tumors	X						

Figure 1. Slide number 12 from the Agency public presentation entitled "Integration of Toxicogenomic Technology to Inform Risk Assessment A Case Study: Propiconazole" (Presented during the May 24-26, 2011 FIFRA SAP Meeting on Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools)

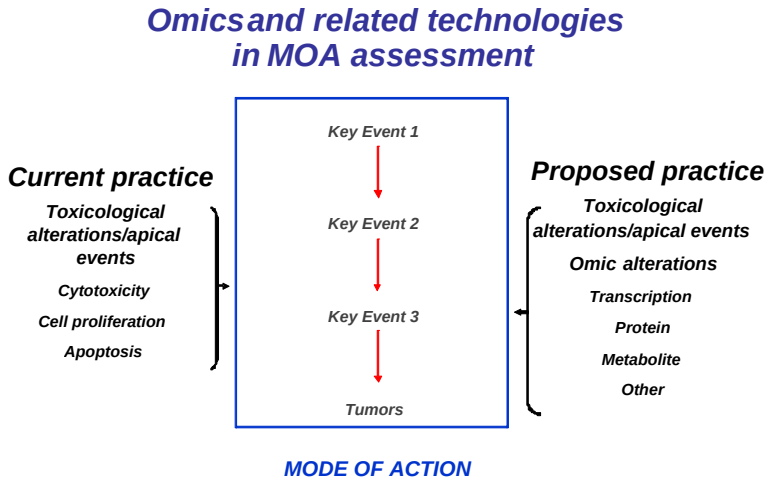

Figure 2. Slide number 30 from the Agency public presentation entitled "Integration of Toxicogenomic Technology to Inform Risk Assessment A Case Study: Propiconazole" (Presented during the May 24-26, 2011 FIFRA SAP Meeting on Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools)

In terms of the conclusions regarding the proposed propiconazole AOP based on omics technologies, the Panel felt there were several strongly supported aspects as well as some areas that needed further development.

A. <u>Strongest points in use of omics to develop an AOP for propiconazole:</u>

1. The evidence for CAR/PXR activation as a key initiating event was well reasoned and adequately supported by the omics data. The bioinformatic pathway analysis, with multiple algorithms strongly supported CAR/PXR activation as the initiating event, especially given existing data in the open literature demonstrating propiconazole induction of PXR regulated reporter genes. Unequivocal support, however, would be provided by CAR or PXR knockout mice.

2. A strong case also was made for regulation of cell cycle genes as a key event leading to liver hyperplasia, for which altered cholesterol metabolism was an important contributing factor. Functional data that are supported by ongoing experiments *in vitro* by the Agency have now implicated accumulation of mevalonic acid, which results from increased cholesterol biosynthesis and metabolism, as a driver of hepatocyte proliferation.

3. A perhaps unanticipated new role for alteration of retinoic acid (RA) metabolism as a key event was uncovered by the integrated omics approach (including support from direct measurement of hepatic RA levels). Retinoic acid is a well-known

inducer of cellular differentiation in multiple contexts, and support for altered RA metabolism as a key event is being followed up in cell culture experiments. Propiconazole induced reduction in RA levels in the liver, as a key event leading to hyperplasia. This key event may perhaps be tested by checking whether restoration of RA can fully or partially reverse the impact of propiconazole either *in vivo* or in the cell culture model.

4. Metabolomics revealed changes in protein and lipid oxidation, and altered glutathione metabolism, that supported the transcriptomic responses implicating increased oxidative stress in response to propiconazole. Whether propiconazole's role in inferred induction of reactive oxygen species then leads to liver tumorigenesis has not been unequivocally determined, however.

B. Questions that remain open:

Based on the propiconazole story as currently developed, it might be speculated that hazard identification for a particular chemical could be accomplished with omics technology alone. However, this approach only shifts the need for comprehensive screening results before assessments could be done. In addition, the omics revolution has put the empirical field of toxicology into hyper-drive by generating millions of new "glimpses" of biological activity. This has created a greater reliance on artificial intelligence algorithms to interpret the data faster. Without AOPs to guide the use of 'omic data, we run the risk of repeating the mistakes of the past where artificial intelligence algorithms along with thousands of molecular descriptors were proposed as the solution to toxicological QSARs.

Thus, despite excellent progress on the use of omics to delineate an AOP for propiconazole, the Panel felt the following points still need to be adequately addressed moving forward.

1. It is not clear at present if all of the key events hypothesized for propiconazole are required to induce tumors, or if just one or two are sufficient. This question could be explored by i) identifying other chemicals that appear to act through one or more of these AOPs (i.e., the chemical has been shown to cause the MIE and at least one or more subsequent key events), and evaluating whether they also induce liver tumors, and ii) identifying other chemicals that have been shown to induce liver tumors and have also been shown to cause the MIE (or an early key event) in one or more of the proposed propiconazole AOPs, and investigating whether these chemicals are active through all three, or only a subset of the propiconazole AOPs.

2. Further, there are varying levels of evidence to support each of the key events proposed for the propiconazole AOP. For example, the proposed key event of suppression of apoptosis needs to be better characterized. Is it truly a key event; i.e. is it a necessary step in the progression to liver tumors in this model, or simply hyperplasia, or neither? What is (are) the MIE(s) that leads to suppressed

apoptosis? Can the transcriptomic data be backed up by proteomics or TUNEL assays, or caspase activation at time points and doses analyzed?

3. Significant concerns were also raised by multiple Panel members and public commenters regarding the effect of propiconazole on mutagenesis, as assayed using the BigBlue® mouse model. Specifically, the effects of propiconazole on hepatic DNA modification seemed remarkably fast and potent, as compared to other known carcinogens that have well characterized genotoxicity. The Panel recommended independent analysis of the data including repetition of the assay with additional doses and time courses to resolve this critical point of contention regarding mutagenicity as a key event in the propiconazole AOP.

4. Differences between the key events underlying phenobarbital vs. propiconazole AOPs were not convincingly demonstrated to some panel members. Data presented to the Panel by the Agency suggested that propiconazole induces a distinct molecular signature relative to phenobarbital, but this contradicted previous studies in the open literature suggesting more similarities than differences (e.g. Dail *et al*., 2008). The inferred pathway differences might be explained by differential expression in controls due to age, sex, strain, circadian rhythms, or other physiological factors. Further, the doses of compounds used (and whether they were at or exceeded maximal tolerated doses) and potential differences in kinetics of responses, must be taken into consideration when comparing effects of compounds, particularly ones that are predicted to induce responses via the same initiating event. For example, while it would indeed be quite interesting if propiconazole specifically or conazoles as a class, but not phenobarbital, selectively altered certain responses in hepatocytes such as altering retinoic acid metabolism, as much attention should be paid to similarities in omic responses as differences. Regardless, coming to a consensus on the MIEs and the key events leading to an AOP, even simply among conazoles themselves, must be achieved since a major goal of using AOPs in IATA is to build predictive power, which may not have been fully realized in this case study. The case study also demonstrated the continued importance of depositing omics data in public databases, as the Agency has done, to allow independent analysis and agreement or disagreement in the field regarding the conclusions from such studies.

5. Along these lines, some Panel members raised the issue that identification of key events from *in silico* pathway analysis of microarray data is still subject to considerable individual interpretation. Therefore, the conclusions reached by one group may differ considerably in terms of which pathways are more important than others, and therefore worth considering as a key event.

6. A recently published study (Schlosser *et al*., 2003) was not included in the information presented to the SAP and is relevant to the evaluation of the MOA. The study also may have implications for addressing the topics of reading across chemicals (Case 1, Question 3) and of species differences (Case 1, Question 4). The investigators used "humanized" mice models (mice expressing the human

CAR/PXR receptors and not their own) to demonstrate that although liver cells from both humanized and wild-type mice show increased CYP enzyme activity after treatment with phenobarbital , only the wild-type cells were able to show a proliferative response. This finding suggests that the induction of hepatic CYP enzymes by phenobarbital or any other xenobiotic, such as propiconazole, is by itself insufficient to cause cell proliferation and tumor formation, although it could still be necessary. The Schlosser *et al*. study appears not to have included a positive control (e.g., treatment with EGF) to establish whether the modified liver cells still retained the ability to proliferate, but nevertheless their major findings are consistent with the fact that, according to the information presented to this Panel by the USEPA and in public commenters, neither phenobarbital nor conazoles seem to be able induce proliferation of liver cells in humans. Therefore, the results of the study with humanized mice can be interpreted as indicating that differences in the functionality of the CAR/PXR receptor between species may be the determining factor for whether cell proliferation will occur following activation of the receptor.

Data presented by public commenters to the SAP and some of the data in the humanized mouse study (Schlosser *et al*. 2003) also indicated that the induction of CYP enzymes by xenobiotics such as phenobarbital or propiconazole is several-fold or perhaps even 1-2 orders of magnitude higher in the mouse than the human liver. Therefore, a more generic MOA for xenobiotics including both propiconazole and phenobarbital and which would apply to both humans and rodents, could be envisioned where cell proliferation would depend on the quantitative relationship between the activation of CAR/PXR receptors and the induction of CYP enzymes. Namely, the higher level of CYP induction observed in mouse cells may be what induces cell proliferation in this species, whereas the lower level of induction in human cells is insufficient to do so. This scenario may be useful to consider an alternative to the MOA proposed in the Agency's issue paper.

Case Study 1 Charge Question 2*: Are there other approaches/technologies that could be used to develop and link Key Events to develop MOAs/AOPs for propiconazole and/or chemicals?*

Panel Response

The Panel agreed that there are other technologies that could be used to develop and link key events. In the current paradigm, altered apical events are used to identify and link key events. Historically, mechanistic toxicological investigations have been used in this way. An example of this type of approach would be blocking a key event and demonstrating that subsequent events or the *in vivo* adverse effect do not occur. In other cases, high-throughput screening (HTS) technology similarly can be used to demonstrate linkages between key events and to develop MOAs/AOPs. For example as shown in the estrogen-receptor (ER)-binding FIFRA SAP (US EPA, 2009a), ER-binding assay data that was measured *in vitro* in a liver slice assay was linked qualitatively and

quantitatively to target protein synthesis (vitellogenin), which lead to reproductive impairment.

The molecular approaches, as outlined in the case studies allowing characterization of multiple macromolecular changes following a chemical exposure, have revolutionized capabilities for understanding interconnections in signaling pathways leading to toxic outcome. Omics technologies may lead to identification of biomarkers that represent selected key events. Use of the identified biomarkers may allow one to skip events in the MOA but still use the AOP. As noted by the Agency, it will be crucial to use the best tool to address the hypothesis coming out of the AOP in order to achieve success.

Whereas these new techniques hold promise in advancing toxicological knowledge, especially along the cellular response pathway, results from these omics and HTS investigations should fit in a logical fashion with existing *in vivo* and *in vitro* data. While genomics data help to uncover toxicological mechanisms, HTS is more likely to enhance identification of categories of chemicals with similar properties associated with a particular key event.

While it is difficult to predict the future of the fast moving field of molecular biology, as noted by one member of the Panel, deep sequencing of whole and individual genomes is probably one of the technologies on the horizon. In addition, consideration of post-translational evaluations for epigenomic evaluations is likely to be a new technology of use with the new paradigm.

Periodic evaluations of technologies will be just as necessary as the periodic re-evaluation of proposed AOPs. While omic strategies are becoming more common today, they (or at least some of them) will likely be obsolete within a few years. Thus, OPP needs to be flexible with new technologies and prudent with regard to investments.

Regardless of the methods used, toxicological principles driving dose selection, frequency of exposures, timing of exposures, and timing of assessments relative to exposure(s) need to be fully considered in determining environmentally relevant altered pathways that can lead to toxicity.

Key events and MOA/AOP for chemicals which act directly at the cellular level can continue to be studied using traditional *in vitro* or subcellular assays using or derived from multiple cell types and a range of endpoints. Understanding key events and MOA/AOP for chemicals that do not act primarily at the cellular level (e.g., chemicals which modify intercellular signaling in complex organs such as the brain or among two or more interacting tissues) is a more complex task. However, the efficient evaluation of nonlinear interactions/patterns/webs/fingerprints, phosphorylation changes and other posttranslational effects, etc. among key and interconnecting events seems, at present, best served by incorporation of these new omics and related technologies.

Integration of tools is also critical. The top down (i.e., apical outcome to molecular events) approach and bottom up (i.e., MIE to apical outcome) approach are both appropriate strategies. However, since new technologies are rather expensive and given the limitation of resources, initial studies should focus on top-down evaluations. Moreover, it will be necessary to invest funds into technologies that are amenable to integration into IATAs.

At the present time, HTS testing without targets specified by an AOP or conducting biological studies without integrating PBPK or an exposure component may not be the most effective way of using limited resources to further a paradigm shift. As noted by one Panel member, knowledge bases which are currently or soon will be ready may allow PBPK subsidized with Metapath technologies to calculate target organ concentrations of parent and metabolites at specific temporal periods so that key events within specific target organs can be calibrated with dose and time.

Temporal aspects of exposure can have significant impacts on AOPs, as it is likely that MOA/AOPs may change depending on concentration at specific time points. For example, due to inhibition kinetics with CYP, propiconazole may have a "low dose" AOP that may lead to endocrine or reproductive impairment through aromatase inhibition, and a "high dose" AOP that may lead to hepatocellular carcinoma. Life stage and gender is also critically important in this evaluation.

Given the time that it will take to implement a coherent policy and procedure for introducing AOP into risk assessment paradigms, EPA scientists and others could begin to provide tissue samples from current dose-response and temporal toxicity studies that are providing apical endpoint data. These tissue aliquots should be archived for future omics or other molecular methods that will allow linkage between molecular and apical endpoints. It is imperative to take tissue samples at the same time points and doses in which an apical endpoint occurs.

As observed by several Panel members, knowledge of physical characteristics, reactivity of chemical moieties and the ability to search for similarities in structure among chemicals with known outcomes (e.g., using MetaPath) can be a very powerful tool in understanding the toxicity potential of uncharacterized chemicals. The development of such knowledge bases is very important, as they will allow the development and use of expert systems and the choosing of targets for screens to ensure that new and useful information will be provided to fill gaps in that knowledge.

As noted by a member of the Panel, this similarity of approach has been used in the SAR area (see Wu *et al*., 2010; Blackburn et al 2011). However, a word of caution was raised by another Panel member, who stated that history has shown that whenever there is a perception of complexity, there is a tendency to borrow statistical tools used in data exploration and network processes to construct predictive models. With the advent of larger and faster computers, as well as better bioinformatics software, pattern recognition and artificial intelligence tools have also been adopted by modelers. The hope here is that by adding hundreds of molecular descriptors, *in vitro* measurements or

protein responses will elucidate a multivariate model for complex biological phenomena. However, the results of the past two decades have shown that the use of statistical methods instead of toxicity mechanisms, while useful in fitting existing data, is not useful for predicting complex toxic endpoints.

From the propiconazole case study, it can be argued that the omics studies reveal two and possibly three potential AOPs leading to the apical endpoint of increased liver tumors. Testing this out will require an increased understanding of the quantitative relationships among key events within the pathways. Evaluating other conazoles may also be prudent to see if there is overlap, since more confidence is likely if the same key events occur with different members of the same chemical-category.

Case Study 1 Charge Question 3*: Please comment on the general utility of AOPs/MOAs or Key events to read across chemicals to provide a more defensible weight of evidence support for the traditional risk assessment approaches that OPP will continue to use in the near- and mid-term?*

Panel Response

The strategy of "read across" relies on the formation of categories of similar chemicals so that data from tested chemicals within each category can be used (read across) to assess the activities of other chemicals in the category that have not been tested for the apical endpoint. The ability to read across chemicals for risk assessment would be highly desirable for several reasons elaborated elsewhere (reduces cost, effort, time, and test animals). Furthermore, the identification of AOPs/MOAs or Key Events can be important in traditional risk assessment since it allows the toxicologist/risk assessor to perform a more predictive, realistic assessment. Also, one important advantage of the MOA/AOP approach to read across is that it allows categorization of chemicals based on toxicological similarity. Thus, for a given chemical, it is important to consider the scope and breadth of the existing information on that chemical, and indicate the level of confidence associated with the application of AOP information to read across. The output of an AOP should be a clear statement of the confidence, analysis, interpretations, and conclusions, and should be a transparent evaluation of all available evidence and relevant data.

Categorization of a chemical based on both its MIE and additional early Key Events would provide stronger weight of evidence than basing it solely on the MIE. This would provide a higher level of confidence that a particular AOP has been initiated. In fact, the utility of the MOA/AOP/Key Event framework will depend on the robustness of the scientific evidence that supports it based on the following questions (OECD, 2011):

1) Has the MIE for the particular chemical or class of chemicals been correctly identified?
2) Have the linkages between the MIE, key events across different levels of biological organization (cell to tissue to organ), and adverse outcome of interest been positively established?

3) Have the cross-linkages between multiple AOPs been adequately considered?

These are all important questions that will need to be considered for validation of a MOA/AOP. It is also important to recognize that a chemical may act through more than one MOA/AOP to induce a given adverse outcome. If a chemical has been tested for the potential to act through a limited set of AOPs, important actions of the chemical may be missed, and the chemical may be incorrectly classified and assessed. Similarly, a chemical can have many different biological actions, not all of which may be involved in inducing an adverse effect. For example, the data presented in the propiconazole case study illustrates that propiconazole interacts with several different nuclear receptors (e.g., CAR/PXR, PXR/RXR, PPARα), yet only some of these interactions (i.e., CAR/PXR) appear to be involved in the chemical's tumorigenic action.

The field of computational biology is thought of as having two distinct areas: knowledge discovery, which includes data-mining and the elucidation of patterns from experimental data (an approach that is used in bioinformatics), and simulation-based analyses, which uses *in silico* approaches to develop predictions that can be tested *in vitro* and *in vivo*. Simulation-based analysis has direct relevance to AOPs and IATA as envisioned by OPP. A significant amount of experimental work is needed to establish an AOP but once established, an AOP becomes central to acceptance of IATA. First qualitative, and then quantitative comparisons are made between data from experimental models and the *in vivo* situation. The qualitative AOP can guide the information requirements and testing (i.e., identification of relevant inputs and database development for these inputs), while the quantitative AOP can guide the evidence synthesis which allow for higher ordered regulatory decisions. Existing knowledge of chemicals and biological responses should be used to develop predictive models of toxicity that in turn facilitate the *in silico* analyses of new chemicals. When data gaps are identified, targeted *in vitro* and, if required, *in vivo* experiments are used to provide the required data. The predictive models, as well as the design of appropriate testing strategies (i.e., IATA), will benefit tremendously from the input of information from AOPs and systems biology. For this reason, the two fields should evolve in parallel, with knowledge generated in one area being used to inform decisions in the other.

The end product of a traditional toxicity assessment (i.e., hazard-identification and dose-response assessment) typically involves the prediction of a safe dose or concentration or the identification of an appropriate point of departure (POD) if a margin of exposure approach is used. There is still a need to identify the adverse effect and the appropriate POD based on that adverse effect (either a NOAEL/LOAEL or a BMDL). The information gained from the AOP/MOA approach will help identify the targeted animal testing needed to identify the appropriate POD for the most relevant, critical adverse health effect (i.e., the adverse effect that occurs at the lowest concentration). In addition, an understanding of the MOA will allow for the best method to extrapolate below the POD and also allows the use of data-derived uncertainty factors instead of default factors.

Case Study 1 Charge Question 4: *Please comment on the potential use of omic and related technologies to inform AOPs in non-mammalian taxa for use in ecological risk assessment?*

Omic and related technologies have an important place in ecotoxicology and ecological risk assessment. They are rapidly evolving and being applied in a number of areas. The predominant application to date has been for characterizing molecular and cellular consequences of stressor exposure. However, as a recommended by the Panel, because of the far larger number and wider diversity of non-mammalian species, a much larger effort may be required to apply these new frameworks in ecological risk assessment.

The EPA Office of Water convened an Science Advisory Board two years ago that asked a similar overall question in how the Agency could use AOPs to update Aquatic Life Criteria (ACL) (US EPAb, 2009). The last ACL document was updated in 1984 and requires testing in multiple species. FIFRA also requires multiple species testing (acute and chronic) for pesticide registration. Thus, use of AOP methods to target testing can likewise greatly reduce the number of animals tested and allow effective testing to specific species.

MOA is already used for acute toxicity of pesticides (i.e., herbicides are toxic to plants, and insecticides are toxic to insects). This is usually addressed in the problem formulation stage. Consequently, AOP data can be used in conceptual models in the problem formulation stage of the assessment. A greater understanding of the AOP can streamline testing to "receptors of interest" and "life stages of interest". If we know a compound is an ER agonist in vertebrates, then tests can be focused on chronic exposure and reproduction in vertebrates and limit acute toxicity studies.

As additional genomic information is obtained from taxa used in standard ecorisk evaluation (fathead minnow, *Daphnia*, algae), it should be possible to use the same processes as those used for the assessment of human health risk. Deep sequencing methodologies are making this possible. Endangered species should particularly be amenable to these studies as animal testing in these species is prohibited.

The key to making omics and related technologies relevant to ecological risk assessment is being able to calibrate the molecular response to apical endpoints. The most difficult component here of course will be linking molecular effects not only to apical endpoints of reproduction growth and survival, but also linking apical endpoints to the population level. Successful models that can be used as templates include those for salmonids (Ankley *et al*. 2008), and marine copepods (Cary *et al*., 2004; Chandler, 2004; Chandler *et al*. 2004). HTS with redundancy (e.g., multiple ER bioassays) can be used as screens for MOA/AOP for new chemical registrations with subsequent targeting for genomic evaluations and then targeted animal testing.

The Panel recommended that in order to utilize these approaches in ecological risk assessment and the AOP paradigm, major challenges need to be overcome which,

while significant, are not insurmountable. Building the scientific foundation to implement the AOP paradigm in non-mammalian species will include several scientific issues including the following:

- Defining the actual range of homeostatic responses (e.g., what is the true range of the controls?)
- Determining how to handle hormesis in the response variable.
- Translating *in vitro* responses to organisms in a variety of, or fluctuating, environmental conditions.
- Differentiating biochemical from energetic responses.
- Confronting the huge range of organism variability (including both plant and animal kingdoms)
- Determining if non-mammalian cell cultures respond similarly to mammalian cell lines.
- Extrapolating from *in vitro* to *in vivo*.
- Differentiating response as a function of life stage and seasonal variation
- Quantifying dose-response since some estimate of the threshold (both acute and chronic) is needed to compare with exposure estimates in order to facilitate risk characterization

This process also provides a mechanism to discern adaptive responses that permit survival versus those adaptations that are truly adverse and toxicologically important. This is of particular importance in defining the relationship between adaptive changes and disease process outcomes by establishing those that have direct and predictable relationships to specific chemicals and/or classes of chemicals.

The AOP approach ultimately focuses on health in an escalating/ increasing level of biological organizational approach. Integration of rodent and fish testing has high potential for use in this approach. NOAA has used an AOP Health Approach in research and monitoring for toxins from harmful algal blooms (HAB) using an integration of rodent and fish testing. For several years, research at NOAA on Marine Biotoxins has used fish seizure profiles to help identify different HAB toxins, which are further tested in rodent models. A combination of molecular based toxin quantification approaches has been developed which enables quantification of HAB toxins and neurotoxic effects. Ultimately, the mouse bioassay is the legally recognized assay for HAB toxin quantification and has been benchmarked against these newly developed bioassays. Health outcomes are the ultimate final benchmark in this integrated approach.

As fish and other ecotoxicological models are developed it will be important to look at models where there is good information on reproductive effects. The National Marine Fisheries Service (NMFS) in NOAA has fish population models developed for many critically managed commercial and recreationally important species including salmonids and other fish and shellfish species. The use of these NMFS fish reproduction models, with which reproduction and growth are easily modeled, has great potential for use in the development of AOP ecotoxicological approaches.

In developing IATA and AOP for invertebrates, one may wish to focus on compounds (e.g. chitinase inhibitors) that have toxicity that is specific to invertebrates, as suggested by one Panelist. This will provide a focus that may also allow comparisons between the responses of freshwater and marine invertebrates. In particular metabolomic responses will be very different between freshwater and marine species for some metabolites (e.g., osmolytes) but those of toxicological importance may have more in common. It is very important to discover those metabolites of toxicological relevance. Recent genomics and transcriptomic work at NOAA has focused on developing gene expression differences among different key marine species including shrimp, oysters and dolphins. Oyster genomics data allow determination of differences in genetic expression between oysters in industrial, urban, suburban and pristine areas. The important genes in this include several genes involved in calcium regulation (calmodulin 2), growth and gonadal development, energy metabolism and neurological disorders (GABA A Receptor). These gene responses may help shape the metabolomic end points developed in concert within this research (Borpujerdi, et. al, 2009; Schlock et. al. 2010; Morrison et. al, 2007; Viant et. al, 2009).

One invertebrate model that provides a useful example of an integrated approach is the marine copepod model based on *Amphiascus tenuiremis,* and developed by Dr. Tom Chandler at the University of South Carolina (Cary *et al*., 2004; Chandler, 2004; Chandler *et al*., 2004; and Volz and Chandler, 2004). This model allows for a multigenerational assessment using multiple life stage assessments (nauplii, copepodites and adults) to be made on multiple endpoints (growth, survival, development and reproduction). Since modeling will be the key to linking these new novel testing approaches at different levels of biological complexities with higher level effects in species, populations and communities, it would be useful to have animal models that can verify predictions for one or two key cornerstone ecotoxicological species. The Panel recommended the multi-level, multi-generational approach exemplified in this model, and future considerations of invertebrate models for the AOP Framework should use models that provide these linkages.

The use of field monitoring data with these AOP/ HTS approaches may continue to be useful in discerning exposure and possible pesticide-specific molecular responses. One thing that may be important to examine in terms of current body-burden measurements for exposure is the transition of exposure body-burden measurements from wet weight/dry weight or lipid based measurements to include measurements in plasma for fish or plasma equivalence in invertebrates. This exposure measurement may help us better relate exposures in HTS in rats, mice and other mammals with more ecotoxicologically focused species such as fish. The importance of this approach was presented in the case study with trioclosan.

As mentioned by one Panel member, to illustrate this (as a consideration and/or comparison to what could be potentially be applied to propiconazole), data in dolphins for triclosan from NOAA's Health and Environmental Risk Assessment (HERA) Project show that on a plasma basis the highest level of triclosan exposure in dolphins is an order of magnitude below what is considered a lowest level of exposure in humans and T4

levels in these dolphins are normal (Fair, 2009). The measurement of concentrations in plasma in dolphins facilitates a direct comparison of HTS results between mice/rats and marine mammals. These studies in the dolphin illustrate another important point. When field exposure measurements are related to a battery of health metrics (liver, thyroid enzymes as well as molecular end points), the highest correlations are often with conventional legacy pollutants such as PCBs and DDT/DDT metabolites that may have common modalities for what we would term a MOA event along the AOP. In this SAP the importance of impurities in single compounds or differences in different metabolites of a parent compounds in laboratory studies have been discussed. In laboratory studies these problems may be the basis of excluding data; however, in the field, the presence of multiple compounds is the norm. Thus, information on impurities may be the basis for exclusion of data in some data bases, but the excluded data may be of paramount importance in discerning field effects where multiple compounds may be the norm. In these circumstances, inclusion of impurities may have ecological relevance in interpreting field effects. Further, these results may be relevant in extrapolation to humans who may also have high body burdens of these same legacy pollutants. This means that human exposures will also be different from pure controlled exposures in laboratory studies.

On way to help resolve this problem of multiple exposures to compounds in the field is to advocate the use of controlled mesocosm studies to provide field data that are more readily related to laboratory data using the IATA and AOP approach. This may help to distinguish indirect effects in the field from direct toxicological effects. This approach is illustrated by a mesocosm study where the herbicide atrazine had indirect effects on bivalve mollusks by reducing food availability. Observed ecotoxicty was due to a marked reduction in the availability and quality of food. Ability to discern indirect versus direct toxicological effects is important in helping better understand and relate field and laboratory findings using this AOP/HTS approach (Lawton *et al*., 2006). The use of mesocosm studies of nanomaterials identified non selective particle feeders as the vulnerable population (Ferry *et al*., 2009), and focused research on those species. Combining these approaches of mesocsom testing and AOP/HTS approaches along with field effects data may enable the development of better predictive models of effects at higher trophic levels that can be linked to molecular HTS end points.

The AOP approach may be very useful for dealing with the problem of the toxicity of mixtures. One thing that could be done is to use these approaches to discern potencies for compounds with similar modes of action or mechanisms of action. Given the high through put nature of the new molecular based assays it may be a very useful approach to test mixtures that are determined to be high risk for additive effects.

Speaking specifically about the propiconazole AOP, unfortunately there is relatively little information about this pathway in non-mammalian taxa. Some studies have suggested that non-mammalian vertebrates do not have CAR receptors, only PXR-like receptors (e.g., Maglich *et al*. 2003). In addition, a number of studies with teleost fishes have indicated that phenobarbital and conazoles either do not induce CYP activity in the liver (e.g., Kleinow *et al*., 1990) or induce the "wrong" kind (e.g., CYP1 and not

the CYP2 type) (Hasselberg *et al*. 2005 and references therein) to qualify for inclusion in the proposed propiconazole MOA framework. On the other hand, in reptiles there is evidence that phenobarbitol can in fact induce CYP proteins of the right type (CYP2 type) (e.g., Ertl et al., 1999). Overall, the information available for non-mammalian taxa is insufficient to make generalizations and application of omic technologies to these taxa would be of great value to determine if the proposed MOA for propiconazole, or triclosan or any other xenobiotic for that matter, can be used in ecological risk assessment for these and related compounds. Lastly, the Panel advised it would be of interest to determine the quantitative relationships between receptor activation and CYP enzyme induction in non-mammalian species to assess the global applicability of the proposed propiconazole MOA.

Finally, these new approaches need to be phased in over time so that they can replace animal models as more confidence is gained. This phased approach will reduce the number of animals tested and replace them over time with HTS and other technologies. Providing incentives for industry to transition some testing for re-registration or for new registration of pesticides to use more HTS approaches in lieu of conventional animal testing protocols should be advocated by EPA OPP.

Case Study 2 Charge Question 1*: Use of the adverse outcome pathway concept to characterize the impact of triclosan on the thyroid hormone system promotes a research and risk focus on toxicodynamic and kinetic differences in thyroid homeostasis between (1) the study species (e.g.., rats) and the population of interest (e.g., pregnant women/fetus and young children) and (2) potential species similarities and/or differences in the MIEs. Please comment on scientific knowledge and types of data considered to be informative in selecting appropriate factors for interspecies differences.*

Panel Response

The Panel noted that the Agency's issue paper (pg. 60)correctly states "If the key molecular event(s) within an AOP are phylogenetically conserved across mammals and other taxa, then a mechanistic-based understanding of an AOP in mammals may be extrapolated to those other taxa." . In evaluating the triclosan case study some important questions need to be considered. Are there toxicokinetic and toxicodynamic differences in thyroid homeostasis between rats and humans (including pregnant women/fetus and young children)? As Bernstein noted (public comment letter, May 17, 2011) "This (the AOP-based IATA) approach would then necessitate a framework for determining variations within a species for a particular parameter, at either the molecular, cellular or whole animal level, before clarifying cross-species differences."

With regard to triclosan, it is assumed for now, though this needs to be tested, that the data showing a decrease in serum T4 in the rat are valid and not due to an artifact generated by interference in the radioimmunoassay. The lack of change in serum TSH in the rat and the failure to find similar effects of triclosan on serum T4 in humans is concerning; the human study used a different type of assay system. This technical issue

should be addressed before one can determine whether the triclosan case study provides proof of principle for establishing an AOP.

The choice of a case study that assumes a thyroid hormone-related AOP is very reasonable and powerful because of the deep and specific knowledge base for thyroid hormone physiology, pharmacology, toxicology, and metabolism. Consequently, human (as well as animal) thyroid biology provides a strong framework on which to map any proposed AOP. The triclosan case study project presented by the Agency demonstrated an interesting and novel set of data regarding alterations of hepatic enzyme gene expression.

The particular AOP in the triclosan case study (activation of CAR/PXR as the MIE, which results in increased activity of hepatic enzymes that catabolize T3 / T4, decreased T4 level, and altered TR-regulated processes that lead to adverse neurodevelopment) is not yet sufficiently well supported to use predictively, based on the USEPA issue paper reviewed by the Panel, though that is the ultimate goal of risk assessment. Consequently, better understanding of this AOP or substitution of a more plausible AOP will contribute to the iterative process of the IATA.

It is important in AOP studies to understand the healthy, undisrupted, physiological system to be studied in toxicological assessments. The thyroid system is well understood, and has been studied in a range of species. A formal description (in the form of a computational model consisting of a set of non-linear differential equations) of the human thyroid system has been published (Degon *et al*., 2008). Although this contains a lot of parameters, it may form a useful starting point for testing the impact of particular disruptions to normal functioning of the system caused by toxicants. A biologically based dose-response (BBDR) model was developed by McLanahan *et al*. (2008) for dietary iodide in adult rats. This model includes a number of submodels for dietary iodide, thyroid-stimulating hormone (TSH), and the thyroid hormones, T4 and T3. It may be possible to modify models such as those described above to cover other species and to use parts of the models to test specific hypotheses. Some useful measures of kinetic parameters are available in the literature for some aspects of the thyroid system in a range of species and stages of development. Some longitudinal studies by Soldin *et al*. (2010) of thyroxine kinetics in pregnant and non-pregnant women provide useful information about the large changes in the behavior of the thyroid system that occur in pregnancy.

Thrall *et al.* (2009) used New Zealand white rabbits as a model organism to develop a physiologically based pharmacokinetic model of the behavior of iodide in pregnant females and fetuses. Le Traon *et al*. (2008) developed a model to describe the pharmacokinetic behavior of total thyroxine in male and female beagle dogs. Leghait and co-workers (2009, 2010) developed pharmacokinetic models for the behavior of the thyroid system following disruption by fipronil in male and female Wistar rats and in sheep (Lacaune ewes and cross-bred rams). These studies provide not only some potentially useful data that could be used to check models, but also indicate the necessity of taking into account a range of pharmacokinetic processes (including absorption,

metabolism, and elimination) and factors such as binding to non-target tissues when making interspecific predictions. A genetic study of thyroid hormone transporters in humans provides indications of the range of polymorphisms in humans and the importance of these proteins in the maintenance of thyroid hormone levels. The sorts of information described above could be useful toward informing the design and interpretation of *in vitro* assays, and their relevance to the *in vivo* behavior of toxicants in the development of AOPs. The Agency could derive benefit from work to develop computational models that can be parameterized and used to predict the impacts of a range of toxic insults affecting the thyroid axis. This approach would also be worthwhile for other physiological systems that are disrupted by exposure to xenobiotics.

The potential for adverse outcomes due to thyroid disruption is very large given the critical role that this hormone plays in neurological development. Thyroid deficiency during fetal and neonatal life leads to severe condition of growth and mental deficits known as cretinism. The basic biology of the thyroid system in mammals is very well characterized; it is a complex system, and there are many potential sites of action for chemicals to disrupt thyroid function.

The vertebrate hypothalamo-pituitary-thyroid (HPT) axis is evolutionarily conserved; it formed in the earliest jawless vertebrates and was then maintained over more than 450 million years of evolution. However, the scientific knowledge clearly shows that there are derived features of the HPT axis that lead to variation in thyroid physiology. For example, thyroid hormone transport and clearance are different in rodents and humans (and also in non-mammalian species). This may in part be due to differences in serum thyroid hormone transport proteins, which influence the distribution of TH in the body and modulate TH clearance. There are three proteins that transport TH in mammals with varying affinities for thyroxine (ranked by affinity): thyroxine binding globulin (TBG)>transthyretin (TTR; a.k.a. prealbumin)> albumin. Unlike humans, rodents make TBG only during early stages of development and rely more on TTR. TBG becomes especially important in the human mother during gestation. Non-mammalian species do not have TBG, and instead rely on TTR and albumin for TH transport. The specificity of TTR for THs differs among species: while mammalian TTRs have high affinity for thyroxine (T4), most non-mammalian TTRs are high affinity binding proteins for 3,5,3'-triiodothyronine (T3). Some species have additional TH binding proteins; e.g., turtles have a dual function protein that binds vitamin D and T4 with high affinity.

The Charge Question goes to the heart of the rich knowledge of human, mammalian and non-mammalian vertebrate thyroid hormone pharmacology, etc. Intentional, extensive, and comprehensive use of such knowledge that is available from studies of the effects of pharmaceuticals, mutations, and disease processes would be an extremely valuable a priori exercise in cases such as a proposed thyroid hormone AOP. Briefly, determine what steps in the metabolism and mechanism of action of the target of interest (thyroid hormone, here) have been productive drug targets; what components of the system of interest are subject to mutations that produce target-related disease (ref. OMIM database); or, what nutritional interventions end up affecting the target of interest. Contrasting these known or knowable interventions that selectively affect the target, and

therefore must lie within any relevant AOP, would provide a knowledge-based screen against other pathways that may, in some general sense, intersect or overlap with the target, but not have any pathway-specific, and, most importantly, rate-limiting participation in the function of the target. Rate-limiting steps, generally identifiable as drug and/or mutation targets, ought to be central focal points and intrinsic checks for the quality of proposed AOPs. By analogy, in the world of cancer biology, if one proposes that a gene product is a critical component of a pathway to carcinogenesis, they will immediately be asked whether there is a relevant mutation of the encoding gene. These mutation and drug targets will not be exhaustive in all cases because of peculiarities of druggability or mutability, or because of AOPs that rely on multiple nodes. Nonetheless, any effort to simplify rather than limitlessly expand the search space for AOPs will require making occasional tradeoffs, and this seems like a reasonable and relatively unbiased way to bootstrap the process. The example presented by the Agency in the public presentation entitled "Building Libraries of Adverse Outcome Pathways" included in Slide 24 (Figure 3) a map that is based on drugs (chemicals), and is an excellent representative of maps that could be more complete (using not only chemical, but also genetic and nutrient data). In the case of the proposed AOP for thyroid hormone effects of triclosan, it is unclear whether any known thyroid-related drug or mutation targets are explicitly within the AOP. Consequently, it is not yet possible to have confidence that the AOP would explain a key change in thyroid function.

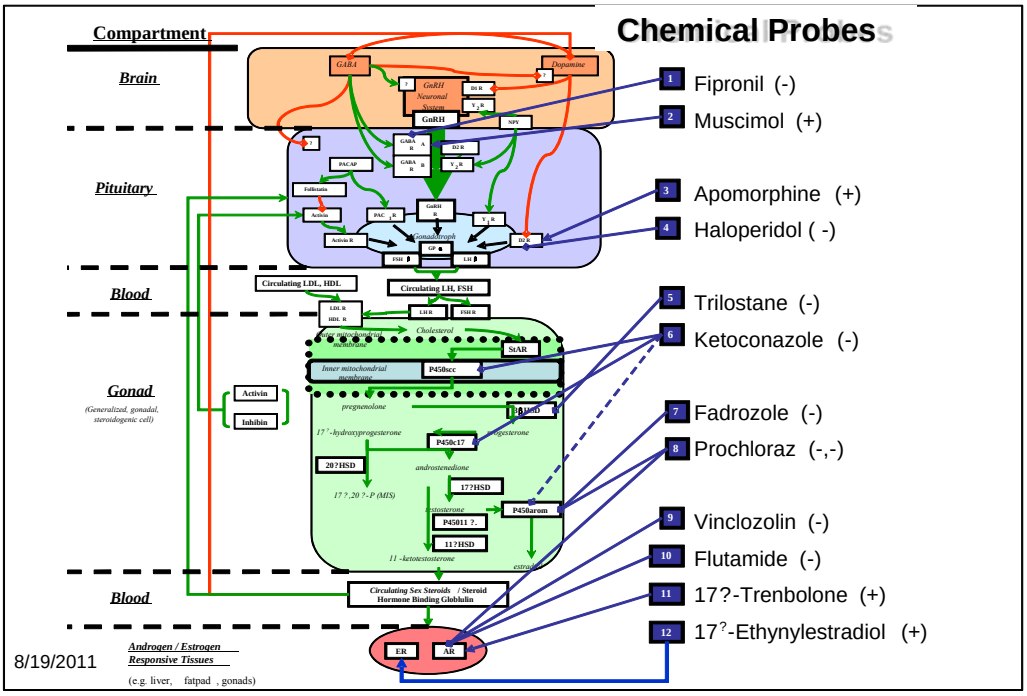

Figure 3. Slide number 24 from the EPA Public Presentation entitled "Building Libraries of Adverse Outcome Pathways"(Presented during the May 24-26, 2011 FIFRA SAP Meeting on

Integrated Approaches to Testing and Assessment Strategies: Use of New Computational and Molecular Tools)

The following scientific knowledge and types of data would be informative in accounting for interspecies differences and special human population (pregnant women/fetus, young children) differences in toxicodynamic and kinetic aspects of thyroid hormone disruption by triclosan. These items are not triclosan specific, but would be helpful in consideration of quantitative differences among species and for special human populations. For some of the items, the specific information for triclosan may be available in the literature, but it was not presented in the case study.

1. A scaling relationship for the dosage (exposure) metric for species and body size. The mg/kg/day metric may be inappropriate because the systemic clearance (CL) of a chemical from the body typically varies among species in proportion to body weight (BW) to the ¾ power; i.e., as BW increases, CL capacity per unit BW declines and the same mg/kg dosage in rat and human would produce a larger area under the systemic chemical concentration vs. time curve (AUC) in the human. The draft document reports the opposite effect for triclosan, and using a ¾ power scaling relationship would make the rat-human difference in exposure even larger. The more ideal internal exposure metric is the unbound (i.e., free) concentration at the sites of the chemical initiating event; e.g., AUC-free and Css-free for acute and chronic chemical administration. The administered dosage that creates the same internal concentration across species and subpopulations would be useful in this regard. Interspecies scaling of dosage should probably utilize body weight to the ¾ power as the default.

2. In pregnancy, some xenobiotic clearance pathways are enhanced; e.g., glomerular filtration rate is elevated by 50% and some drug metabolism enzymes are induced, others are unchanged or diminished, driven by altered endogenous steroid levels. Plasma albumin concentration declines and elevated free fraction of albumin-bound drugs ensues. All else equal, the dosage in pregnancy may produce a lower exposure at the site of initial biochemical reaction than it does in the non-pregnant adult. For triclosan, is CL in pregnancy elevated compared with in the non-pregnant state?

3. In the fetus, knowledge of placental transfer of the chemical would be useful, in particular regarding whether the concentration in the fetus is similar to that in the maternal system, or whether a lower or higher concentration is achieved during the exposure of the fetus via the mother.

4. In neonates, CL pathways are immature. Renal CL matures to adult levels exponentially with a half life of about 8 months and fully mature renal function occurs at about two years of age in humans. In addition, renal function at birth on a mL/min/kg basis, is equivalent to adult renal CL, and over the initial several years of life, renal clearance capacity exceeds the normal adult value by as much as a factor of two, reaching peak capacity at about 30 months of age (Hayton 2000). For hepatic CL, different xenobiotic metabolism enzymes show different ontological

patterns, with some developing slowly to adult values throughout childhood and others rising to adult values quickly (discussed further in response to Question #2).

5. Do the same key events and AOP apply in the particular species? Are the dose-response relationships the same for the key events? Thyroid hormone transport and clearance clearly vary among species. One MOA for many chemicals that leads to reduced serum [T4] is competition for binding to TTR. The diversity of structure and function of vertebrate TTRs (and other TH transport proteins) suggests that a chemical that disrupts TH binding to TTR in one species may not do so in another, and vice versa. For example, modeling of the structures of polychlorinated biphenyls (PCBs) has shown that some have T4-, and others T3-like characteristics. Limited empirical data supports this modeling. This suggests that compounds that disrupt TH binding to TTR in a mammal may not do so in a bird (or a reptile or amphibian or fish).

6. Liver metabolism is another point for regulation of thyroid physiology and likely to vary across species. Within mammals there are important differences in the degree to which different compounds enter pathways for glucuronidation and sulfation, two important enzymatic conjugation processes that target THs for clearance. The transcriptional regulation of these and other liver enzymes also varies across species, as was illustrated by the different functions (ligand independent and ligand dependent) of the constitutive androstane receptor (CAR) isoforms among mammals (Elcombe, 2010).

7. *In vitro* hepatocyte and metabolism enzyme activity studies might be useful as a guide to quantitative *in vivo* interspecies differences in metabolic CL.

8. For *in vitro* studies, the stability of the chemical in the system should be characterized. For example, for cell culture exposed for a period of time to a chemical, we should know not only the time = 0 concentration, but also at the end and perhaps at intermediate times. The area under the concentration time curve would then give a reliable measure of the exposure of the system to the toxicant.

9. Other components that regulate thyroid physiology and may vary among species include, but are not limited to, the expression and structure/function of the following: tissue monodeiodinases; organic anion, monocarboxylate and amino acid transporters; and cellular TH binding proteins. For example, mutations in MCT8 lead to severe neurological deficits in humans, but such mutations do not have similar effects in mice.

10. The most highly conserved feature of the HPT axis may be the nuclear receptors that mediate TH action, the thyroid hormone receptors (TRs). Constraint on the structure of the ligand binding pockets of the proteins has led to strong stabilizing selection that has maintained their structure and function. This means that ligands that interact with high affinity with the ligand binding domain of human TRα will also bind to a fish or a frog or a bird TRα. This appears to also hold for synthetic

TH analogs. These analogs have similar activities in mammals and frogs, despite an evolutionary distance of approximately 350 million years. However, it appears to date that relatively few synthetic chemicals interact directly with the TRs, probably owing to constraints on ligand binding. Nevertheless, for chemicals that may be found to have a MOA that involves binding to TRs in one species, one would predict that these chemicals would have a similar MOA in distantly related species.

11. A potential confounding factor in interpreting MOAs for chemical disruption of thyroid function that could apply to many or most species is the possibility for independent or concurrent activation or inhibition of other pathways that impinge on TH homeostasis. There is now strong evidence for cross regulation among nuclear receptors in these different pathways. An apparent effect on a predicted thyroid-dependent endpoint may actually be due to activation or inhibition of a different pathway that impinges on thyroid-dependent processes. One example is the synergistic activation of TH target genes by TH and stress hormones (glucocorticoids). The stress hormone essentially titrates the sensitivity of the target gene to being transactivated by thyroid hormone.

12. PBPK Models – The Agency has defined a MIE as "the initial point of chemical-biological interaction within the organism that starts (*i.e.*, initiates) the AOP"(US EPA, 2011). One of the best tools currently available for predicting chemical concentrations at a target site *in vivo* is a physiologically based pharmacokinetic (PBPK) model. Development and use of a PBPK model for triclosan across life stages and species will be a necessary component of IATA in order to incorporate apparent species differences in the kinetics of triclosan. For example, enterohepatic recirculation of triclosan in the rat results in prolonged terminal half-lives and a larger fraction of triclosan eliminated in the feces compared with humans. A PBPK model would allow for the Agency to explore and understand how different physiological variables such as altered glomerular filtration, metabolic differences, changes in relative tissue volumes, and changes in protein binding would affect the pharmacokinetics (PK) of triclosan and ultimately its effect on thyroid homeostasis. A pharmacodynamic model should also be developed and used to describe the thyroid homeostatic mechanism to couple with the PK of triclosan, as well as other chemicals that may interfere with the normal function of the thyroid system. The Agency has made progress with life stage PBPK models with the pyrethroids and chlorpyrifos, with the latter being linked to a pharmacodynamic model for acetylcholinesterase inhibition.

13. While a PBPK model is the best approach to account for differences between the study species (i.e., rat) and the population of interest (i.e., pregnant women/fetus and young children), it is often difficult to parameterize a PBPK model or to have the experimental data to validate the model. One alternative is the use of chemical specific adjustment factors (CSAFs) (IPCS, 2001). IPCS 2001 addresses the use of quantitative toxicokinetic and toxicodynamic data to address interspecies and inter-individual differences in dose/concentration-response assessment. Since

AUC data are available for triclosan in humans and in rats, it may be possible to derive a CSAF for the toxicokinetic portion of the interspecies uncertainty factor (UF). The use of data to derive an interspecies CSAF will result in a more predictive value when using rat data to predict effects in humans. It may also be possible to estimate the toxicokinetic portion of the interindividual variability between an average human and the populations of interest (i.e., pregnant women/fetus and young children). If the data are not available, the default intraspecies UF of 10 will adjust the point of departure (POD) applicable to an average human to a sensitive population.

14. The general notion of mining databases and literature on mutations (human and animal, natural and induced) and drugs (prescribed and non-prescribed legal compounds, and drugs of abuse) with explicit attention to identifying key nodes relevant to AOPs will be powerful, but not exhaustive. In particular, such "apical endpoints" as neurobehavioral and cognitive deficits are so poorly understood mechanistically that they are likely to be resistant to such *a priori* analyses, except in cases such as hypothyroidism, where the neurobehavioral deficits are fairly-well connected to simpler metabolic pathways and known key developmental stages.

Case Study 2 Charge Question 2: *Please comment on the factors most responsible for the variability of toxicokinetics and toxicodynamics of compounds affecting thyroid hormone homeostasis among humans (particularly susceptible subpopulations such as pregnant women and children).*

Panel Response

The Panel concluded that the main contributing factors to variability both in natural ranges of T4, T3, and TSH in human populations as well as responses to replacement therapy with T4 or combinations of T4 and T3 are currently poorly understood in both humans and animal models. It is interesting to note that determination of TSH levels is still the front-line assay of choice for assessing the function of the thyroid axis, rather than T4, implying that individuals develop their own set point of adequate T4 and ultimately tissue T3 levels. The underlying mechanism of action of environmental chemicals that affect T4 levels without significant effect on TSH or T3 levels is also poorly understood and warrants more investigation.

There has been some recent progress on the influence of specific gene polymorphisms on individual human variability in thyroid axis set points and responses to administered thyroid hormones (Panicker, Saravanan *et al*., 2009, Dayan and Panicker 2009), with one interesting linked candidate as the Type II deiodinase (D2). In addition, there are clear rodent strain differences in responses to a variety of exogenous toxic chemicals and the underlying genetic or even epigenetic basis, while presently unknown, is being actively investigated (Koturbash, Scherhag *et al*. 2011, Bradford, Lock *et al*. 2011, Liu, Ichihara *et al*. 2009). In the end, a great deal of individual variability may reside in liver or other extrathyroidal metabolism as well as differential tissue uptake, but yet again this is all still very much in question at the moment.

One Panel member noted that there are several physiological factors that may result in variability in the toxicokinetics of compounds that affect the thyroid homeostasis. There are numerous changes that occur during pregnancy including changes in protein binding, increases in the plasma volume, and increases in the percentage of body fat (Pavek *et al*., 2009), which would be expected to alter the toxicokinetics of compounds compared with men and non-pregnant women. In addition, the enzyme ontogeny and differences in relative volumes of tissues in children will impact their toxicokinetics compared with adults. Infants and children will also have a different exposure pattern than adults, both due to different diets in nursing infants and mouthing behavior in babies and toddlers. PBPK models would allow these differences to be taken into account for IATA.

The studies presented by Dr. Kevin Crofton on the plasma concentrations of triclosan and concomitant concentrations of T4 in pups exposed via milk from dosed dams provide an example of the difference between adults and children. Triclosan with its relatively high octanol water partition coefficient (log Kow=4.76) would have been expected to be present in high concentrations in the milk. However, decline of the plasma concentrations and rebound of T4 suggested that the milk did not contain high concentrations of triclosan. The Crofton laboratory did not measure actual milk concentrations in the dosed dams.

Pharmacokinetic models can become quite complex with the degree of complexity driven in part by the purpose for which the model is to be used. However, for assessment of exposure of the initial biological target to the chemical, a high degree of complexity may be unnecessary. For exposure assessment, the controlling PK parameters, in addition to dosage are bioavailability and the systemic clearance (CL) of chemical. Two cases are described in the following paragraphs that highlight situations where the parent chemical or its metabolite(s) participate in the MIE.

Case 1. Toxicant acts directly and not via toxic metabolite(s). Both acute and chronic exposure of sites of toxicity are controlled by 1) the dose (mg/kg single dose) or the dose rate (mg/kg/day chronic dose), 2) the fraction of the dose or dose rate that is absorbed, F, (hepatic first-pass elimination can be ignored here), and 3) the CL of the toxicant, which is usually the sum of renal and hepatic clearances.

Ignoring person-to-person variability in dose or dose rate (not part of toxicokinetic variability but a contributor to person-to-person variability in exposure), variability in F and in CL both contribute to toxicokinetic variability that would affect exposure to the chemical. For triclosan in particular, F-oral would be expected to approach 100%. It is a small molecule that is lipophilic, yet sufficiently water soluble to dissolve in the fluids of the GI tract lumen and be absorbed in adults, pregnant women and children. Its systemic CL is small and hepatic first-pass elimination via the oral route would not be expected, but in the event that first pass is significant as reported in the Agency issue paper (p. 74), this does not necessarily affect exposure. It is unlikely that there would be significant person-to-person variability in F-oral for triclosan among

various subpopulations. For other compounds affecting thyroid homeostasis, person-to-person variability would depend on the absorption characteristics of the chemical. For chemicals such as triclosan that are generally well absorbed via the oral route, minimal variability would be expected. For chemicals that were not well absorbed due to low water solubility or to low permeability of the GI epithelium, more variability would be expected, and it is generally not possible to predict whether there would be differences among subpopulations although this may well be the case.

For the percutaneous route, a relatively low F-pc would be expected, and reported during the SAP meeting as about 10%. Low bioavailability is generally associated with significant person-to-person and subpopulation variability.

Person-to-person variability in CL will lead to inversely proportional variation in the area under the plasma concentration-time profile (AUC acute dose) and the steady-state plasma concentration (Css chronic dose) of the toxicant. It is generally the case, and assumed here, that the metric of exposure of sites of toxic action is the AUC or the Css (Alternatively, the peak concentration may also be used as the exposure metric.). When the toxicant is plasma protein bound, the unbound (free) concentration is generally considered to be the toxicologically relevant concentration. With the exception described in the following paragraph, the AUC-free and the Css-free are controlled by the hepatic metabolism activity (intrinsic unbound clearance) and the unbound renal clearance. While CL values may also be controlled by organ blood flow and plasma protein binding of toxicant, it is only the intrinsic unbound clearance that controls AUC-free and Css-free. Therefore, it is only person-to-person variability in the activity of toxicant metabolism enzymes in the liver and the capacity of the kidneys to clear unbound toxicant (GFR, active tubular secretion capacity) that introduce person-to-person variability into the AUC-free and Css-free. This is the case for both enteral (oral) and parenteral routes of administration. Person-to-person variability in organ blood flow and plasma protein binding generally have no effect on and contribute no variability to the AUC-free and Css-free.

The exception to the above generalization is when the intrinsic hepatic clearance or the active tubular secretion clearance of the toxicant is highly effective at removing toxicant from the clearing organ blood flow. This is referred to as the high-E (E = extraction ratio) case, and for hepatic high-E toxicants administered parenterally (for enteral administration of hepatically cleared chemical, previous paragraph applies), the AUC-free and the Css-free are controlled by hepatic and renal blood flow and plasma protein binding, and person-to-person variability in flow and binding will produce variability in exposure. Variability in intrinsic clearance (metabolism activity, GFR, active tubular secretion activity) does not contribute to variability in exposure for the high-E case. For drugs used in medical practice and probably for chemicals in general, only a minority are classifiable as high-E. To assess whether a chemical is high- or low-E requires knowledge of the organ CL and the plasma flow (Q) to the organ, where $E = CL / Q$.

Case 2. Toxicity is a result of exposure to metabolite that is formed by metabolism. This case is more complicated. The toxicity may result from a metabolite that is highly reactive (free radical), which immediately reacts within the cell in which it is formed. Regeneration of the enzyme activity occurs via biosynthesis of new enzyme. The exposure metric for this case is the amount of reactive metabolite formed; MIE ensues as long as reactive metabolite is formed. Alternatively, the metabolite may generate the MIE at a site remote from its formation. In that case the exposure metric is the AUC or Css of the metabolite at the target site and these metrics are controlled by the fraction of parent chemical dosage converted to the active metabolite and the CL for the metabolite. Interspecies and subpopulation differences in the fraction of parent dosage converted to metabolite and the CL of the metabolite would lead to differences in the dosage of parent chemical that produced a particular response.

Drug development scientists have addressed the problem of CL prediction (reviewed in Hayton and Hu, 2008). For renal CL, highly reliable inter- and intra-species scaling relationships based upon GFR have been described. *In vitro* hepatocyte, liver slice and drug metabolism enzymes for interspecies prediction of CL have been described.

Potential MIEs. The hypothalamus-pituitary-thyroid axis that produces thyroid hormone is composed of a number of discrete steps at the molecular level, any one of which could be the site of the MIE that ultimately alters thyroid hormone activity. The following list of events involved in the production of thyroid hormone or control of its concentration, not necessarily all inclusive, illustrates potential sites of occurrence of the MIE that disrupts (increases or decreases) thyroid hormone function:

1. Biosynthesis or release of TRH or TSH.
2. TSH binding to its receptor (G-protein) on the thyrocyte membrane.
3. Iodide collection by thyrocytes via the Na^+ / I^- symporter (NIS).
4. Iodine delivery to thyroglobulin in colloid via pendrin (Cl^- / I^- exchanger).
5. Biosynthesis of thyroglobulin by thyrocytes.
6. Iodine fixation to thyroglobulin via thyroid peroxidase.
7. Endocytotic uptake of colloid (and thyroglobulin), a step in the formation of T3 / T4 within the thyrocyte.
8. T3 / T4 removal from thyroglobin, or disruption of the signal that calls for T3 / T4 removal.
9. Binding of T3 / T4 to proteins (albumin, thyroxine binding globulin, transthyretin).
10. Hepatic enzyme activity involved in the elimination of T3 / T4.

A toxicodynamic model and an AOP could be developed around each of the above steps and any one of them would be appropriate when that was where the MIE occurred. Specific factors confirmed to contribute to toxicodynamic variability among individuals or subpopulations are generally unknown.

For example, the triclosan case study describes an AOP that identifies activation of CAR/PXR as the MIE in the triclosan-associated decrease in T3 and T4. This

initiating event produces the following down-stream pathway events: increased activity of hepatic enzymes that catabolize T3 / T4, decreased T4 level, and altered TR-regulated processes that lead to adverse neurodevelopment. For each of these steps, definitive information on inter-individual and subpopulation variability is sparse. Some information is available on the ontogeny of hepatic xenobiotic metabolizing enzymes (, Stevens *et al*., 2003). CYP 3A isozymes are high in the fetus (3A7) which becomes negligible by 1 yr after birth. 3A4 and 3A5 increase to adult levels slowly over prepubertal years. All ages show substantial inter-subject variability. Prenatal SULT2A1 is negligible and higher during the first year of life that after age 1 year (Duanmu, 2006). SULT1E1 is high prenatally compared with the first year of life, which is higher than after age 1 year. Hepatic CYP2C9 activity begins to appear at gestational age 22-24 wk (Koukouritaki, 2004). It increases to adult levels over the first 6 month of life. CYP2C19 appears around gestational age 10 weeks, increases during gestation and reaches adult levels (pmol/mg protein) by age 6 month. SULT and 2C isozymes show large interindividual activity, with high / low ratios of 10 – 50 for various ages pre- and postnatal. While large, this variability is similar to that observed in adults.

Case Study 2 Charge Question 3*: Benchmark dose analysis involves selection of a benchmark response (BMR). The consequences of a perturbation of a key event depend on the magnitude, timing, and duration of the perturbation, which in turn depends on dose. Subtle perturbations of key events may be damped out and have little effect on downstream biology due to the operation of homeostatic processes. Please comment on the current understanding of the magnitude of perturbation of thyroid hormone levels in humans that may be required to lead to adverse outcomes.*

Panel Response

One of the first steps in performing benchmark dose analysis is the selection of a BMR. If there is an accepted level of change in the endpoint that is considered to be biologically significant by the scientific community, then that amount of change is chosen for determination of the POD. A BMR defines the demarcation between non-adverse and adverse changes in toxicological effect parameters for continuous data. The Agency has chosen a 20% decrease in T4 levels as the BMR for the triclosan case study. It would have been helpful if the Agency had provided information on the problem formulation behind the decision to use a 20% decrease in T4 levels for BMD modeling so the Panel could determine the applicability of this specific BMR value for the specific problem formulation (i.e., screening procedure, sensitivity analysis or traditional risk assessment). The question is whether a 20% decrease in T4 levels would lead to adverse neurodevelopmental effects.

When considering an adverse outcome pathway, the endogenous regulatory pathway needs to be known well enough to allow discernment of perturbations that would lead to adverse outcomes. The thyroid system can be ideal for investigating perturbations possibly leading to adverse outcomes, because of the body of knowledge concerning its physiological regulation and response to drugs, mutations and malnutritions. However, in the case of triclosan, the experimental data does not seem to

fit the typical pattern expected from perturbations on the thyroid system, since TSH levels are unchanged despite the decrease in T4 following exposure to triclosan. Typically, if thyroid hormone levels are low, the level of TSH ultimately has to be either abnormally high or abnormally low. Unless the pathology (triclosan toxicity) maps correctly on the physiology (thyroid hormone feedback loop), then something is missing and that missing information leads to the process of revising and refining the AOP.

However, one panel member noted that increases in TSH do not always accompany decreases in T4; citing polychlorinated biphenyls (PCBs) and TCDD as examples of chemicals that have been associated with neurodevelopmental effects in humans (Patandin *et al*., 1999) and observed to decrease T4 in the absence of an increase in TSH. For example, in 7-day gavage studies in rats, two PCB mixtures and 4 individual PCB congeners were shown to dramatically decrease free and total serum T4 levels in a dose-responsive manner, accompanied by less dramatic, but statistically significant reductions in free serum T3 levels, in the absence of significant effects on serum TSH levels (Martin and Klassen, 2010). TCDD was also included in the studies of Martin and Klassen (2010). TCDD decreased both total and free serum T4, while having no effect on T3 or TSH levels. A recent study in pregnant women of the association between levels of PCBs and T4 and TSH levels reported similar findings. Specifically, Chevrier *et al* (2008) analyzed the relationship between serum PCB levels and thyroid hormone levels in 334 pregnant women living in the Salinas Valley of California, and reported a negative association between serum PCB levels and free and total serum T4 levels. No association was observed in this study between serum PCB levels and TSH levels.

A condition in pregnant women known as hypothyroxinemia provides another example where discernable elevations in TSH do not always accompany decrements in T4. Hypothyroxinemia is a common condition in pregnant women, in which low free T4 concentrations are associated with TSH concentrations in the normal range. Studies of women with hypothyroxinemia may also be informative with respect to the possible neurodevelopmental implications of low T4 during early gestation. Kooistra *et al*. (2006) assessed newborn development at 3 weeks of age using the Neonatal Behavioral Assessment Scale (NBAS), comparing 108 neonates born to women with low maternal free T4 concentrations measured at 12 weeks gestation (< 10th percentile among 1353 pregnant women with measured thyroid parameters) to 96 neonates born to women with free T4 concentrations between the 50th and 90th percentiles. The authors reported that infants born to women with low maternal free T4 at 12 weeks gestation had significantly lower scores on the NBAS orientation index than infants born to women with free T4 concentrations in the 50th to 90th percentile range at 12 weeks gestation. In a regression analysis, no association was observed between neonatal orientation scores and maternal free T4 concentrations later in gestation, or maternal TSH concentrations.

The Charge Question specifically asked for feedback on the "understanding of the magnitude of perturbation of thyroid hormone levels that lead to adverse outcomes." There is unlikely to be a single correct answer (from human and animal physiology and pharmacology) to this question. Thyroid hormone levels in humans are never interpreted on their own. Many factors determine whether a perturbation of thyroid hormone level

(serum, tissue, etc.) is associated with adverse outcomes. The document addresses some, but not all, of these other variables. Free hormone concentrations are critical to know, but even knowing this is not sufficient because peripheral and central deiodination, transport, etc. determine biological activity. The most critical barometer of outcome is typically thought to be the state of the feedback loop that controls endogenous thyroid function (*i.e.,* TSH/TRH). This loop is thought to requires that changes in free thyroid hormone will be accompanied by rational changes in TSH (and/or hypothalamic TRH). However, as discussed above, this is not always the case (e.g., the rat studies of Martin and Klassen 2010 showing that PCBs and TCDD decrease T4 in the absence of an increase in TSH). Other reliable barometers of thyroid pathway outcomes are metabolism (O2 consumption), glandular morphology, and growth hormone secretion. Neurobehavioral (neurodevelopmental) consequences are harder to measure, in general. However, cognitive and developmental deficits downstream of thyroid deficiency are probably among the best understood.

The Agency issue paper states: *"Knowledge of the AOP can provide useful information in the dose–response assessment for a substance. If the quantitative relationships are understood for precursor events or key events within a causal path leading to a disease or adverse outcome, the key event(s) can serve as the basis of the dose response analysis."* The Agency should be careful about using precursor key events as the basis for a POD for a risk assessment unless the quantitative relationships are understood between precursor events or key events within a causal path leading to a disease or adverse outcome. Decreased T4 level blood levels are a precursor key event that may or may not be related to pathological changes (i.e., the possibility exists that subtle perturbations of key events may be damped out and have little effect on downstream biology leading to pathology due to the operation of homeostatic processes). Dose response data on pathological key events (abnormal perturbation in neuronal function or development leading to decreased neurodevelopmental effects) is not available for triclosan. Targeted animal testing should be conducted to obtain dose response data on abnormal perturbation in neuronal function/development leading to decreased neurodevelopmental effects. At this time and with the level of data presented in the triclosan case study, the selection of a 20% decrease in T4 levels as the BMR (with the implication that such a decrease in T4 levels alone, in the absence of other documented changes in thyroid hormones results in adverse neurodevelopmental effects) has not been adequately justified as being scientifically sound and protective of human health.. However, with targeted animal testing, it may be possible to determine the magnitude of perturbation of T4 alone or in combination with other thyroid hormones that would lead to adverse neurodevelopmental effects. This would be valuable information for this particular AOP and could potentially be applied to other chemicals that operate through a similar AOP.

An example where a precursor key event has been used as the basis for selecting a POD for use in a risk assessment is the development by the California Environmental Protection Agency's Office of Environmental Health Hazard Assessment of a Public Health Goal for Perchlorate in drinking water, based on the inhibitory effect of perchlorate on the uptake of iodide by the thyroid gland (OEHHA, 2004), Specifically,

perchlorate competes with iodide for uptake into the thyroid gland by the Na^+/I^- symporter (NIS). The Public Health Goal for perchlorate was based on a benchmark dose analysis of data from a human study measuring the inhibitory effect of perchlorate on iodide uptake by the thyroid gland. A benchmark dose response rate of 5% (i.e., a 5% decrease of mean iodine uptake by the thyroid gland) was selected in this analysis.

As discussed in the response to Case Study 2 Charge Question 2, perchlorate's inhibition of iodide uptake by thyrocytes via the NIS is one of at least 10 possible MIEs that can disrupt thyroid hormone function. Another possible MEI leading to disruption of thyroid hormone function is enhancement of hepatic enzyme activity involved in the elimination of T3 / T4—an effect attributed to triclosan in this case study. Thus, while the early events in the AOPs for perchlorate (inhibition of NIS and decreased iodide uptake, decreased T4 synthesis) and triclosan (CAR/PXR activation, increased catabolism of T3/ T4 by hepatic enzymes) differ, the latter events (decreased T4 levels, altered TR-regulated processes, adverse neurodevelopment) are identical. Unlike triclosan, however, perchlorate's inhibition of iodide uptake in the thyroid to reduce T4 synthesis results in an increase in TSH concentrations systemically, responding in a manner consistent with a prototypical disruptor of thyroid homeostasis. There is also consistency in this response pattern across species (USEPA, 2002).

Another example of a precursor key event used as the basis for selecting a POD is the use of cholinesterase inhibition to protect against organophosphate-induced neurotoxicity (see for example US EPA 2006 Organophosphorus Cumulative Risk Assessment – 2006 Update).

The BMDL guidelines (USEPA 2000) suggest that all relevant adverse health effects should be evaluated if their LOAELs are near to each other (i.e., within a factor of 10). Different adverse health effects may be more (or less) relevant to humans (Figure 4; Boobis *et al*. 2006). When an adverse response is not relevant to humans, then another adverse response that occurs at higher doses should be chosen as the critical effect and the AOP for that adverse effect be investigated to see if it is relevant to humans.

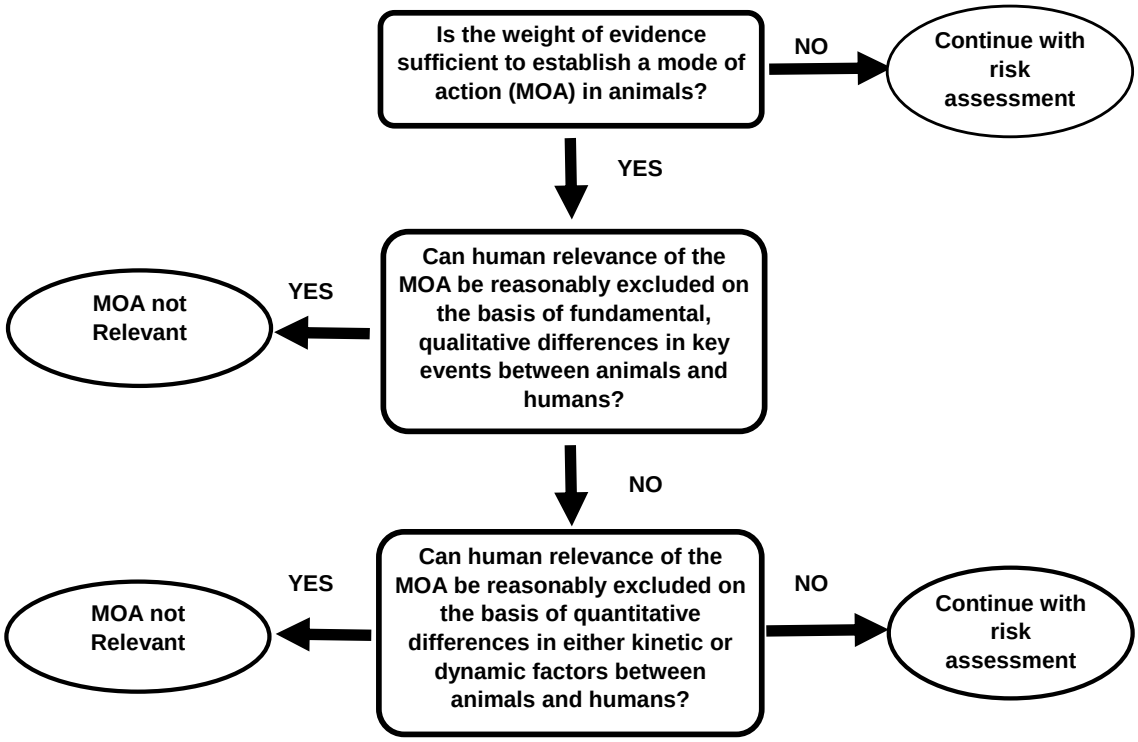

Figure 4. Human relevance framework (Boobis *et al*. 2006)

In the triclosan case study, the Agency stated "The 10% BMR is selected as a default approach and to provide a lower end of biological changes". However, if the BMR is not known for continuous data, the default BMR is one standard deviation (BMR_{1SD}) not the 10% BMR response. USEPA (2000) recommends the BMD_{1SD} and the $BMDL_{1SD}$ should be presented (USEPA 2000) even if a different BMR is chosen. However, a BMD_{1SD} from control mean corresponds to an approximately 10% excess risk for individuals below the 2nd percentile or above the 95th percentile of the control distribution for normally distributed effects.

There are advantages to conducting BMD modeling on early key events. There are MOA studies that have conducted BMD modeling for an early key event and calculated the BMD or BMDL for the initial key event (Schlosser *et al*. 2003). The BMD increases as the key step progresses. The amplification of the response along the AOP pathway can be quantified. This is a very powerful tool to understand MOA, and should be encouraged.

References

Allmyr, M., Adolfson-Erici, M., McLachlan, M.S., & Sandborgh-Englund, G. (2006) Triclosan in plasma and milk from Swedish nursing mothers and their exposure via personal care products. *Science of the Total Environment* 372, 87-93.

Ankley, G. T. Miller, D. H Jensen, K. M. Villeneuve, D. L. & Martinovic, D. (2008). Relationship of plasma sex steroid concentrations in female fathead minnows to reproductive success and population status. *Aquatic Toxicology,* 88 (1), 69-74

Bakal, C. & N. Perrimon. (2010). Realizing the promise of RNAi high throughput screening. *Developmental Cell,* 18 (4), 506-507.

Bearden, D. W. (2011). *Use of Metabolomics in Dolphin and Human Urine Analysis and Comparisons,* NIST, Hollings Marine Laboratory, 230 Ft. Johnson Road, Charleston, SC 29412: Personal Communication.

Bradford, B., Lock, E.,Kosyk, O., Kim, S. ,1, Uehara, T. , Harbourt, D., DeSimone, M., Threadgill DW, Tryndyak V, Pogribny IP, Bieyle, L., Koop, D., & Rusyn, I. (2011). Interstrain differences in the liver effects of trichloroethylene in a multistrain panel of inbred mice. *Toxicological Sciences,* 120 (1), 206-217.

Brenan, C. J., Roberts, D., & Hurley, J. (2009). Nanoliter high-throughput PCR for DNA and RNA profiling. *Methods in Molecular Biology,* 496**,** 161-174.

Blackburn, K., Bjerke, D., Daston, G., Felter, S., Mahony, C., Naciff, J., Robison, S., & Wu, S. (2011). Case studies to test: A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. *Regulatory Toxicology and Pharmacology,* 60, 120-135

Boobis, A.R., Doe, J.E., Heinrich-Hirsch, B., Meek, M.E., Munn, S., Ruchirawat, M., Schlatter, J., Seed, J. & Vickers, C. (2006). IPCS framework for analyzing the relevance of the non-cancer mode of action for humans. *Critical in Reviews, Toxicology*. 38, 87-96.

Borpujerdi, A. B., Vizcaino,M,. Meyers, A., Pollock,E. C . Huynh, S. L,. Schock, T. B, Morris, P. J., & Bearden, D. W. (2009). NMR-based microbial metabolomics and the temperature-dependent coral pathogen. *Vibrio coralliilyticus***.** *Environmental Science and Technology, 43,* 7658–7664.

California Environmental Protection Agency. Pesticide and Epidemiology and Toxicology Section, OEHHA (2004). *Public Health Goals for Chemicals in Drinking Water: Perchlorate*. Available online at: http://www.oehha.ca.gov/water/phg/pdf/finalperchlorate31204.pdf.

Cary, T.L., Chandler, G.T., Volz, D.C., Walse, S.S. & Ferry, J.L. (2004). Fipronil induces male infertility in the meiobenthic harpacticoid copepod, *Amphiascus tenuiremis*. *Environmental Science and Technology,* 39, 522-528.

Chandler, G.T. (2004). Standard guide for conducting renewal microplate-based life-cycle toxicity tests with a marine meiobenthic copepod. In: (S.J. Bailey *et al*., eds.) *ASTM E 2317-04, American Society for Testing and Materials,* W. Conshohocken, PA, USA, 15 pp.

Chandler, G.T., Cary, T.L., Volz, D.C., Walse, S.S., Ferry, J.L., & S.L. Klosterhaus. (2004). Fipronil effects on copepod development, fertility, and reproduction: A rapid life-cycle assay in 96-well microplate format. *Environmental Toxicology and Chemistry,* 23, 117-124.

Chandler, G. T. Cary, T. L. Bejarano, A. C. Pender, J., & Ferry, J. L. (2004) Population Consequences of Fipronil and Degradates to Copepods at Field Concentrations: An Integration of Life Cycle Testing with Leslie Matrix Population Modeling. *Environmental Science and Technology,* 38(23), 6407-6414.

Chevrier J., Eskenazi B., Holland N., Bradman A., & Barr D.B. (2008). Effects of exposure to polychlorinated biphenyls and organochlorine pesticides on thyroid function during pregnancy. *American Journal of Epidemiology,* 168, 298-310.

Cui, X., D. Ji, D. A. Fisher, Y. Wu, Briner, D. M. & Weinstein, E. J. (2011). Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nature Biotechnology,* 29(1), 64-67.

Dail M.B., Shack L.A., Chambers J.E., & Burgess S.C., (2008). Global liver proteomics of rats exposed for 5 days to phenobarbital identifies changes associated with cancer and with CYP metabolism. *Toxicological Sciences*, 106, 556-569.

Dayan, C. M. & Panicker, V., (2009). Novel insights into thyroid hormones from the study of common genetic variation. *Nature Reviews Endocrinology,* 5(4), 211-218.

Degon, M., Chipkin, S.R., Hollot, C.V., Zoeller, R.T. & Chait, Y., (2008) A computational model of the human thyroid. *Mathematical Biosciences,* 212, 22–53

Duanmu Z., Weckle A., Koukouritaki, S. B., Hines, R. N. Falany, J. L., Falany, C., N., Kocarek, T., & Runge-Morris, M. (2006). Developmental Expression of Aryl, Estrogen, and Hydroxysteroid Sulfotransferases in Pre- and Postnatal Human Liver *Journal of Pharmacology and Experimental Therapeutics,* 316,1310-16.

Elcombe, C.R. (2010), Human constitutive androstane receptor (CAR) and pregnane X receptor (PXR) support the hypertrophic but not the hyperplastic response to the murine nongenotoxic hepatocarcinogens phenobarbital and chlordane *in vivo*. *Toxicological Sciences,* 116, 452-466

Ertl, R.P., Bandiera, S.M., Buhler, D.R., Stegeman, J.J. & Winston, G.W. (1999). Immunocytochemical analysis of liver microsomal cytochromes P450 in the American alligator, *Alligator mississippiensis*. *Toxicology and Applied Pharmacology,* 157, 157-165.

Fair, P.A., Lee, H.B, Adams, J., Darling, C., Pacepavicius, G., Alaee, M., Bossart, G.D., Henry, N.D., & Muir, D. (2009). Occurrence of triclosan in plasma of wild Atlantic bottlenose dolphins (*Tursiops truncatus*) and in their environment. *Environmental Pollution* 157, 2248–2254

Ferry, J. L., Craig, P., Hexel, C., Sisco, P., Frey, R., Pennington, P., Fulton, M., Scott, G., Decho, A., Kashiwada, S., Murphy C.J. & Shaw, T.J. (2009). Transfer of gold nanoparticles from the water column to the estuarine food web. *Nature Nannotechnology,* DOI: 10.1038/NNANO.2009.157.

Gilberta, ME, McLanahan ED, Hedge, J, Crofton, KM, Fisher, JW, Valentín-Blasinie, L, & Blount, BC (2011) Marginal iodide deficiency and thyroid function: Dose–response analysis for quantitative pharmacokinetic modelling., *Toxicology,* 283, 41–48.

Gohlke, J. M., R. Thomas, Y. Zhang, M. C. Rosenstein, A. P. Davis, C. Murphy, K. G. Becker, C. J. Mattingly and C. J. Portier (2009). "Genetic and environmental pathways to complex diseases." BMC Systems Biology 3: 46.

Hasselberg, L., Grosvik, B.E., Goksoyr, A., & Celander, M.C. (2005). Interactions between xenoestrogens and ketoconazole on hepatic CYP1A & CYP3A, in juvenile Atlantic cod (*Gadus morhua*). *Comparative Hepatology,* 4, 2.

Hayton, W.L (2000). Maturation and growth of renal function: dosing renally cleared drugs in children. *The American Association of Pharmaceutical Sciences Journal,* 2(1), 3.

Hayton, W.L. & Hu, T.M. (2008). Allometric Scaling. Chapter 29 in *Preclinical Development Handbook: ADME and Biopharmaceutical Properties*, John Wiley and Sons. Hoboken, NJ.

Henry, N. D. & Fair, P.A. (2011). Comparison of *In Vitro* Cytotoxicity, Estrogenicity, and Anti-Estrogenicity of Triclosan, Perfluorooctane Sulfonate (PFOS), and Perfluorooctanoic Acid (PFOA). *Environmental Toxicology and Chemistry,* Submitted and In Review.

Herold, M. J., van den Brandt, J., Seibler, J., & Reichardt, H.M. (2008). Inducible and reversible gene silencing by stable integration of an shRNA-encoding lentivirus in transgenic rats. *Proceedings of the National Academy of Sciences of the United States of America*, 105.47, 18507-18512.

International Programme on Chemical Safety. (2001). *Guidance Document for the Use of Data in Development of Chemical-Specific Adjustment Factors (CSAFs) for*

Interspecies Differences and Human Variability in Dose/Concentration-Response Assessment. WHO/IPCS/01.4. Geneva: World Health Organization.

Kleinow, K.M., Haasch, M.L., Williams, D.E. & Lech, J.J. (1990). A comparison of hepatic P450 induction in rat and trout (*Oncorhynchus mykiss*): delineation of the site of resistance of fish to phenobarbital-type inducers. *Comparative Biochemistry and Physiology,* 96C, 259-270.

Kooistra, L., Crawford, S., van Baar, A.L., Brouwers, E.P., & Pop, V.J. (2006). Neonatal effects of maternal hypthyroxinemia during early pregnancy. *Pediatrics,* 117, 161-167.

Koturbash, I., Scherhag, A., Sorrentino, J., Sexton, K, Bodna, W., Swenberg, J. & Beland, F. (2011). Epigenetic mechanisms of mouse inter-strain variability in genotoxicity of the environmental toxicant 1,3-butadiene. *Toxicological Sciences,* 122(2), 448-456.

Koukouritaki, S.B., Manro, J.R., Marsh, S.A., Stevens, J.C., Rettie, A.E., McCarver, D.G., & Hines, R.N. (2004). *Journal of Pharmacology and Experimental Therapeutics,* 308, 965-974.

Lawton, J.C., Pennington, P.L., Chung, K.W., & Scott, G.I. (2006). Toxicity of atrazine to the juvenile hard clam. *Mercenaria mercenaria*. *Ecotoxicology and Environmental Safety,* 65(3), 388-394.

Le Traon, G., Burgaud, S. & Horspool, L.J. (2008). Pharmacokinetics of total thyroxine in dogs after administration of an oral solution of levothyroxine sodium. *Journal of Veterinary Pharmacology and Therapeutics*. 31, 95–101.

Lee, P. J., Gaige, T.A., Ghorashian, N., & Hung, P.J. (2007). Microfluidic tissue model for live cell screening. *Biotechnology Progress,* 23(4), 946-951.

Leghait, J., Gayrard, V., Picard-Hagen, N., Camp, M., Perdu, E., Toutain, P.L., & Viguié C. (2009). Fipronil-induced disruption of thyroid function in rats is mediated by increased total and free thyroxine clearances concomitantly to increased activity of hepatic enzymes. *Toxicology,* 255, 38–44.

Leghait, J., Gayrard, V., Toutain, P.L., Picard-Hagen, N., & Viguié, C. (2010) Is the mechanisms of fipronil-induced thyroid disruption specific of the rat: Re-evaluation of fipronil thyroid toxicity in sheep? *Toxicology Letters*, 194**,** 51–57.

Liu, F., Ichihara, S., Mohideen, S.S., Sai, U., Kitoh, J., & Ichihara, G. (2009). Comparative study on susceptibility to 1-bromopropane in three mice strains. *Toxicological Sciences,* 112(1), 100-110.

Mahle, D.A., Anderson, P.E., DelRaso, N.J., Raymer, M.L., Neuforth, A.E. & Reo, N.V. (2011). A generalized model for metabolomic analyses: application to dose and time dependent toxicity. *Metabolomics,* 7, 206-216.

Maglich, J.M., Caravella, J.A., Lambert, M.H., Willson, T.M., Moore, J.T. & Ramamurthy, L. (2003). The first completed genome sequence from a teleost fish (*Fugu rubripes*) adds significant diversity to the nuclear receptor superfamily. *Nucleic Acids Research,* 31, 4051-4058.

Martin, L. & Klassen, C.D. (2010). Differential effects of polychlorinated biphenyl congeners on serum thyroid hormone levels in rats. *Toxicological Sciences,* 117(1):36-44.

McDonnell, D. P. & Wardell, S.E. (2010). The molecular mechanisms underlying the *pharmacological* actions of ER modulators: implications for new drug discovery in breast cancer. *Current Opinion in Pharmacology,* 10(6), 620-628.

McLanahan, E.D., Andersen, M.E., & Fisher, J.W. (2008). A Biologically Based Dose-Response Model for Dietary Iodide and the Hypothalamic-Pituitary-Thyroid Axis in the Adult Rat: Evaluation of Iodide Deficiency. *Toxicological Sciences*, 102, 241–253.

Meek, S., Buehr, M., Sutherland, L., Thomson, A., Mullins, J. J., Smith, A. J., & Burdon, T. (2010). Efficient gene targeting by homologous recombination in rat embryonic stem cells. *PLoS One*, 5(12): e14225.

Mischerikow, N. & Heck, A.J. (2011). Targeted large-scale analysis of protein acetylation. *Proteomics,* 11(4), 571-589.

Morrison, N., Bearden, D., Bundy J. G., Collette, T., Currie, F., Davey, M. P., Haigh, N. S., Hancock D., Jones, O. A. H., Rochfort, S., Susanna-Assunta, S., Stys, D., Teng, Q., Field, D., & Viant, M. R. (2007). Standard reporting requirements for biological samples in metabolomics experiments: environmental context. *Metabolomics,* 3, 203–210.

National Research Council. (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington, DC. USA.

Organization for Economic Co-operation and Development. (2011). Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories Series on Testing and Assessment No. 138 ENV/JM/MONO(2011)8 Available at: http://www.oecd.org/document/24/0,3343,en_2649_34377_47858904_1_1_1_1,00.html.

Panicker, V. P., Saravanan, P., Vaidya, B., Evans, J., Hattersley, A.T., Frayling, T.M., & Dayan, C.M. (2009). Common variation in the DIO2 gene predicts baseline psychological well-being and response to combination thyroxine plus triiodothyronine therapy in hypothyroid patients. *Journal of Clinical Endocrinology & Metabolism,* 94(5), 1623-1629.

Patandin, S., Lanting, C. I., Mulder, P. G., Boersma, E. R., Sauer, P. J., & Weisglas-Kuperus, N. (1999). Effects of environmental exposure to polychlorinated biphenyls and dioxins on cognitive abilities in Dutch children at 42 months of age. *Journal of Pediatrics,* 134,33–41

Paul, P., van den Hoorn, T., Jongsma, M.L.M., Bakker, M.J., Hengeveld, R., Janssen, L., Cresswell, P., Egan, D.A., van Ham, M., ten Brinke, A., Ovaa, H., Beijersbergen, R.L., Kuijl, C. & Neefjes, J. (2011). A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. *Cell,* 145(2), 268-283.

Pavek, P., Ceckova, M., & Staud, F. (2009). Variation of drug kinetics in pregnancy. *Current Drug Metabolism*, 10, 520–529.

Pless, O., Kowenz-Leutz, E., Dittmar, G., & Leutz, A. (2011). A differential proteome screening system for post-translational modification-dependent transcription factor interactions. *Nature Protocols*, 6(3): 359-364.

Ross, J., Plummer, S.M., Rode, A., Scheer, N., Bower, C.C., Vogel, O., Henderson, C.J., Wolf, C.R., Schlosser, P.M., Lilly, P.D., Conolly, R.B., Janszen, D.B., & Kimbell, J.S. (2003). Benchmark Dose Risk Assessment for Formaldehyde Using Airflow Modeling and a Single-Compartment, DNA-Protein Cross-Link Dosimetry Model to Estimate Human Equivalent Doses. *Risk Analysis,* 23, 473-487.

Schock, T.B., Stancyk, D.A., Thibodeaux, L., Burnett, K.G., Burnett, L.E., Boroujerdi, A.F.B. & Bearden, D.W. (2010). Metabolomic analysis of Atlantic blue crab, *Callinectes sapidus*, hemolymph following oxidative stress. *Metabolomics,* 6, 250–262.

Schlosser, P. M., Lilly, P. D., Conolly, R. B., Janszen, D. B., & Kimbell, J. S. (2003). Benchmark dose risk assessment for formaldehyde using airflow modeling and a single-compartment, DNA-protein cross-Link Dosimetry model to estimate human equivalent doses. *Risk Analysis,* 23, 473-487.

Seyfried, N. T., Xu, P., Duong, D.M., Cheng, D., Hanfelt, J., & Peng, J. (2008). Systematic approach for validating the ubiquitinated proteome. *Analytical Chemistry, 80*(11), 4161-4169.

Shi, Y., Xu, P. & Qin, J. (2011). Ubiquitinated proteome: ready for global? *Molecular Cellular Proteomics,* 10(5), R110 006882.

Soldin, O.P., Soldin, S.J., Vinks, A.A., Younis, I., & Landy, H.J. (2010). Longitudinal comparison of thyroxine pharmacokinetics between pregnant and nonpregnant women: A stable isotope study. *Therapeutic Drug Monitoring*, 32, 767-773.

Spurgeon, S. L., Jones, R.C. & Ramakrishnan,, R. (2008). High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS One,* 3(2): e1662.

Stevens, J.C., Hines, R.N., Gu, C., Koukouritaki, S.B., Manro, J.R., Tandler, P.J., & Zaya, M.J. (2003). Developmental expression of the major human hepatic CYP3A enzymes. *The Journal of Pharmacology and Experimental Therapeutics,*307(2), 573-82.

Thrall, K.D., Sasser, L.B., Creim, J.A., Gargas, M.L., Kinzell, J.H. & Corley, R.A. (2009). Studies supporting the development of a physiologically based pharmacokinetic (PBPK) model for methyl iodide: pharmacokinetics of sodium iodide (NaI) in pregnant rabbits. *Inhalation Toxicology*, *21,* 519-523.

U.S. EPA. (2002). *Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization* (External Review Draft). Office of Research and Development, Washington, D.C. NCEA-1-0503.

U.S. EPA. (2011). Drinking Water: Regulatory Determination on Perchlorate (40 CFR Part 141). Federal Register. Vol. 76, No. 29 . Proposed Rules. EPA–HQ–OW–2008–0692.

U.S. EPA (2000). *Draft Benchmark Dose Technical Guidance Document EPA* EPA/630/R-00/001, Available online at: http://www.epa.gov/ncea/pdfs/bmds/BMDExternal_10_13_2000.pdf.

U.S. EPA.(2009a) FIFRA Scientific Advisory Panel Meeting on *The Use of Structure Activity Relationships of Estrogen Binding Affinity to Support Prioritization of Pesticide Inert Ingredients and Antimicrobial Pesticides for Screening and Testing,* August 25-27, 2009

US EPA. (2009b) Science Advisory Board meeting report on n Aquatic Life Water Quality Criteria for Contaminants of Emerging Concern. Available at: http://yosemite.epa.gov/sab/sabproduct.nsf/WebReportsbyYearBOARD!OpenView&Start=1&Count=800&Expand=3#3

U.S. EPA. (2011). *Integrated Approaches to Testing and Assessment Strategy: Use of New Computational and Molecular Tools.* Issue Paper: FIFRA Scientific Advisory Panel, May 24 - 26, 2011.

Viant, M.R., Bearden, D.W., Bundy, J. G., Burton, I.W., Collette, T.W., Ekman, D.R., Ezernieks, V., Karakach, T.K., Lin, C.Y., Rochfort, S., de Ropp, J.S., Teng, Q., Tjeerdema, R. S., Walter, J. A., & Wu, H. (2009). International NMR-Based Environmental Metabolomics Intercomparison Exercise. *Environmental Science & Technology,* 43 (1), 219-225.

Van der Deure, W.M., Peeters, R.P., & Visser, T.J. (2007). Genetic variation in thyroid hormone transporters. *Best Practice & Research Clinical Endocrinology & Metabolism*, 21, 339–350.

Volz, D.C., & Chandler, G.T. (2004). An enzyme-linked immunosorbent assay for lipovitellin quantification in copepods: a screening tool for endocrine and reproductive toxicity. *Environmental Toxicology and Chemistry,* 23, 298-305.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics,* 10(1), 57-63.

Wu, S., Blackburn, K., Amburgey, J., Jaworska, J. & Federle, T. (2010). A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. *Regulatory Toxicology and Pharmacology,* 56, 67-81.

Yang, F., Waters, K.M., Miller, J.H., Gritsenko, M.A., Zhao, R., Du, X., Livesay, E.A., Purvine, S.O., Monroe, M.e., Wang, Y., Camp, D.G., Smith, R.D., & Stenoien, D.L. (2010). Phosphoproteomics profiling of human skin fibroblast cells reveals pathways and proteins affected by low doses of ionizing radiation. *PLoS One,* 5(11): e14152.

Zhang, Z., Alpert, D., Francis, R., Chatterjee, B., Yu, Q., Tansey, T., Sabol, S.L., Cui, C., Bai, Y., Koriabine, M., Yoshinaga, Y., Cheng, J.F., Chen, F., Martin, J., Schackwitz, W., Gunn, T.M., Kramer, K.L., De Jong, P.J., Pennacchio, L.A., & Lo, C.W. (2009). Massively parallel sequencing identifies the gene Megf8 with ENU-induced mutation causing heterotaxy. *Proceedings of the National Academy of Sciences.* U.S.A. 106(9), 3219-3224.