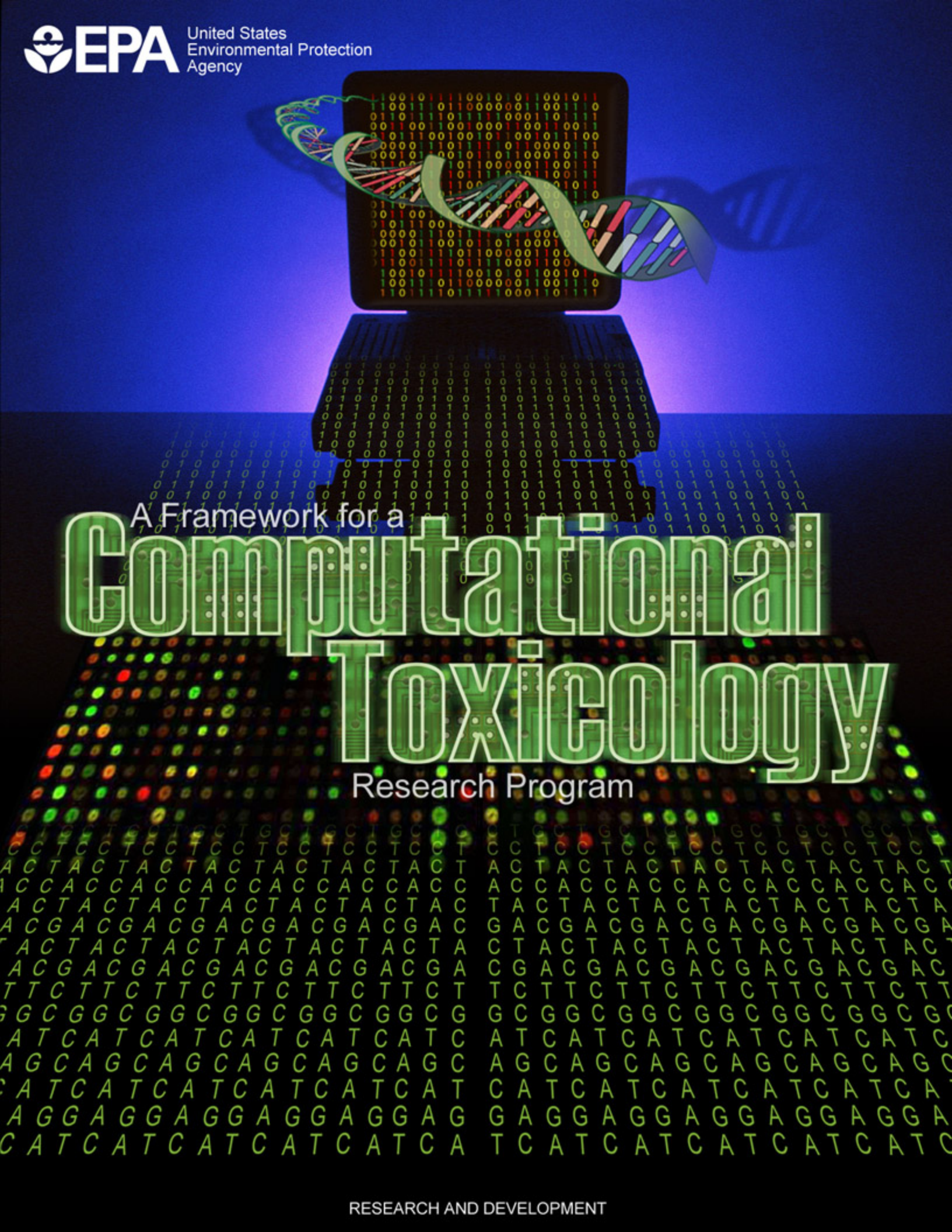
United States Environmental Protection Agency



A Framework for a Computational Toxicology Research Program

RESEARCH AND DEVELOPMENT

EPA/600/R-03/065
November 2003

A Framework for a Computational Toxicology Research Program in ORD

US Environmental Protection Agency
Office of Research and Development
Washington, DC 20460

DISCLAIMER

This document has been subjected to internal and external review for clearance. The Agency's Science Advisory Board provided comments which have been addressed in this draft. This report does not constitute an Agency position or policy concerning computational toxicology. Any mention of trade names does not constitute Agency endorsement.

FOREWORD

The 2003 *Framework for a Computational Toxicology Program in ORD* addresses research needed to apply novel technologies derived from computational chemistry, molecular biology, and systems biology-collectively known as computational toxicology-to improve the Agency's prioritization of data requirements and risk assessments. The *Framework* identifies three strategic objectives: (1) to improve linkages across the source-to-outcome continuum, (2) to develop approaches for prioritizing chemicals for subsequent screening and testing, and (3) to produce better methods and predictive models for quantitative risk assessment. The *Framework* provides a conceptual basis for building on current research in ORD to develop a research program on computational toxicology.

The *Framework* outlines the development of a multidisciplinary, integrated research program that will use computational approaches to link chemical transformation and metabolism, exposure indicators, dose metrics, toxicity pathways, systems biology, and modeling frameworks. The research program will improve the Agency's ability to screen and test for chemical hazards and will address uncertainties associated with dose-response assessment, cross-species extrapolation, and the assessment of chemical mixtures. The program also aims to predict aggregate and cumulative risk, protect susceptible subpopulations, and provide principles for the use of mechanistic data in human health-risk assessments, which are research needs identified in the ORD 2003 *Human Health Research Strategy*. The achievements of the computational toxicology research program will improve the Agency's understanding of the links between human activities, natural dynamics, ecological stressors, and ecosystem condition and, in doing so, will fulfill significant goals of the Agency's 2003 *Strategic Plan*.

The *Framework* is intended to identify the research needs and unique capabilities of ORD laboratories to support a more focused and integrated research program in the future. This document was reviewed by the Agency's Science Advisory Board (SAB) in September 2003 and was the focus of discussion at the ORD *Computational Toxicology Workshop: Framework, Partnerships and Program Development* held in Research Triangle Park, North Carolina, on September 29-30, 2003. Comments from the SAB and workshop participants were used to revise the *Framework* and will be used to guide the development of a multi-year plan that will outline research goals and measures for the next 5-10 years.

Paul Gilman

Assistant Administrator

TABLE OF CONTENTS

Page

List of Figures — vi

Authors — vii

Peer Review — viii

Acronyms — ix

Executive Summary — E-1

I. Introduction — 1
- A. The Computational Toxicology Research Program — 3
- B. Application of Computational Toxicology to Risk Assessment and Research — 4
- C. Overall Goal and Strategic Objectives — 5

II. Research Needs and Applications of Computational Toxicology to Goals — 7
- A. Improve Linkages in the Source-to-Outcome Paradigm — 9
  - 1. Chemical Transformation and Metabolism — 9
    - a. Chemical Transformation in Ecosystems — 9
    - b. Chemical Metabolism — 10
  - 2. Exposure Indicators — 10
  - 3. Dose Metrics — 12
  - 4. Characterization of Toxicity Pathways — 13
  - 5. Metabonomics — 15
  - 6. Systems Biology — 16
  - 7. Modeling Frameworks and Uncertainty Analysis — 17
- B. Provide Predictive Models for Hazard Identification — 18
  - 1. QSAR and Other Computational Approaches — 18
  - 2. Pollution Prevention Strategies — 19
  - 3. High Throughput Screening — 20
- C. Enhance Quantitative Risk Assessment — 20
  - 1. Applying Computational Toxicology in Quantitative Risk Assessment — 20
  - 2. Examples of Applications of Computational Toxicology to Quantitative Risk Assessment — 21
    - a. Dose-Response Assessment — 21
    - b. Cross-Species Extrapolation — 22
    - c. Chemical Mixtures — 24

Page

III. Current Activities — 25
A. Proof-of-Concept: Endocrine Disrupting Chemicals (EDCs) — 27
B. Internal Linkages — 29
1. Human Health Research — 30
2. Ecological Research — 31
C. External Linkages — 31
1. Chemical Industry Institute for Toxicology (CIIT) Centers for Human Health — 32
2. Department of Energy (DOE) — 32
3. National Institute of Environmental Health Sciences (NIEHS) — 33
4. Science to Achieve Results (STAR) — 33

IV. Next Steps — 35
A. Review of the Framework — 37
B. Priorities for Research on Computational Toxicology in ORD — 37
C. Process for a Research Program on Computational Toxicology — 38

Appendix A. Examples of Current ORD Projects Associated with Computational Toxicology — A1

LIST OF FIGURES

		Page
Figure 1	The Source-to-Outcome Continuum	5
Figure 2	An Example of a Toxicity Pathway	15

AUTHORS

Robert Kavlock	Chair, National Health and Environmental Effects Research Laboratory
Gerald T Ankley	National Health and Environmental Effects Research Laboratory
Jerry Blancato	National Exposure Research Laboratory
Tim Collette	National Exposure Research Laboratory
John “Jack” Fowle III	National Health and Environmental Effects Research Laboratory
Elaine Francis	National Center for Environmental Research
Earl Gray	National Health and Environmental Effects Research Laboratory
Karen Hammerstrom	National Center for Environmental Assessment
Lawrence Reiter	National Health and Environmental Effects Research Laboratory
Jeff Swartout	National Center for Environmental Assessment
Hugh Tilson	National Health and Environmental Effects Research Laboratory
Greg Toth	National Exposure Research Laboratory
Gilman Veith	National Health and Environmental Effects Research Laboratory
Eric Weber	National Exposure Research Laboratory
Doug Wolf	National Health and Environmental Effects Research Laboratory
Douglas Young	National Risk Management Research Laboratory

PEER REVIEW

Peer review is an important component of evaluating research documents in the Office of Research and Development. The peer review strategy history for this document follows:

ORD Executive Council: June 7, 2003

External Peer Review: Science Advisory Board/ Framework for a Computational Toxicology Program in ORD, September 12, 2003

External Peer Review Panel Members:

Science Advisory Board Members:

Dr. George W. Lucier, Panel Chair, Consulting Toxicologist, Private Consultant, Pittsboro, NC
Dr. Charles A. Pittinger, Principal, The Cadmus Group Inc., Cincinnati, OH

Consultants:

Dr. Melvin Anderson, Director, Department of Biomathematics and Physical Science, CIIT Centers for Health Research, Research Triangle Park, NC
Dr. John Balbus, Director, Environmental Health Program, Environmental Defense Fund, Washington, DC
Dr. Richard Becker, American Chemistry Council, Arlington, VA
Ms. Patricia Billig, Environmental Toxicologist, Waterstone Environmental Hydrology and Engineering, Boulder, CO
Dr. Stuart Cagen, Toxicology Advisor, Shell Chemical LP, Houston, TX
Mr. Harvey Clewell, Principal, ENVIRON Health Sciences Institute, Ruston, LA
Dr. Darrell Donahue, Associate Professor, Chemical and Biological Engineering, University of Maine, Orono, ME
Dr. B. Alex Merrick, Toxicologist, National Institute of Environmental Health Sciences, Research Triangle Park, NC
Dr. Clifford P. Weisel, Professor, Environmental and Community Medicine, EOHSI, Robert Wood Johnson Medical School-UMDNJ, Piscataway, NJ
Dr. Angela Wilson, Assistant Professor, Department of Chemistry, University of North Texas, Denton, TX
Dr. Andrew Worth, Scientific Officer, Joint Research Center, Institute for Health and Consumer Protection, European Chemicals Bureau, European Commission, Ispra, Italy

Science Advisory Board Staff:
Dr. James Rowe, Designated Federal Official, 1200 Pennsylvania Avenue, NW, Washington, DC

ACRONYMS

AR	Androgen Receptor
BBDR Models	Biologically Based, Dose-Response Models
CEBS	Chemical Effects in Biological System
CNS	Central Nervous System
CIIT	CIIT Centers for Health Research
DBPs	Disinfectant By-Products
DOE	Department of Energy
EDCs	Endocrine Disrupting Chemicals
EDSTAC	Endocrine Disruptor Screening and Testing Advisory Committee
ERDEM	Exposure Related Dose Estimating Model
FQPA	Food Quality Protection Act
HPG/T Axis	Hypothalamic-Pituitary-Gonadal/Thyroid Axis
HTPS	High Throughput Screening
IT	Information Technology
JGI	Joint Genome Institute
LC	Lethal Concentration
LD	Lethal Dose
MENTOR	Modeling Environment for Total Risk Studies
MOA	Mode or Mechanism of Action
MOU	Memo of Understanding
NCEA	National Center for Environmental Assessment
NCER	National Center for Environmental Research
NCT	National Center for Toxicogenomics
NERL	National Exposure Research Laboratory
NHEERL	National Health and Environmental Effects Research Laboratory
NIEHS	National Institute of Environmental Health Sciences
NMR	Nuclear Magnetic Resonance
NRMRL	National Risk Management Research Laboratory
OEI	Office of Environmental Information
ORD	Office of Research and Development
PBPK Models	Physiologically Based, Pharmacokinetic Models
PC	Personal Computer
PD	Pharmacodynamic
PFOS	Perfluorooctane Sulfonate
PK	Pharmacokinetic
QSAR	Quantitative Structure Activity Relationships
SAB	Science Advisory Board
SAR	Structure Activity Relationship
SHEDS	Stochastic Human Exposure and Dose Simulation Model
SNP	Single Nucleotide Polymorphism
STAR	Science to Achieve Results

EXECUTIVE SUMMARY

The mission of the U.S. Environmental Protection Agency (the Agency) is to safeguard public health and the environment from adverse effects that may be caused by exposure to pollutants in the air, water, soil, and food. Protecting human health and the environment carries with it the challenge of assessing possible hazardous effects for tens of thousands of chemicals. The large number of chemicals that the Agency must consider under many different regulations, together with the large cost of conducting test batteries, limits the full use of standard toxicity test methods to only a small number of chemicals. The Agency is also faced with reducing uncertainties associated with performing quantitative risk assessments on chemicals for which data have been submitted by the chemical industry.

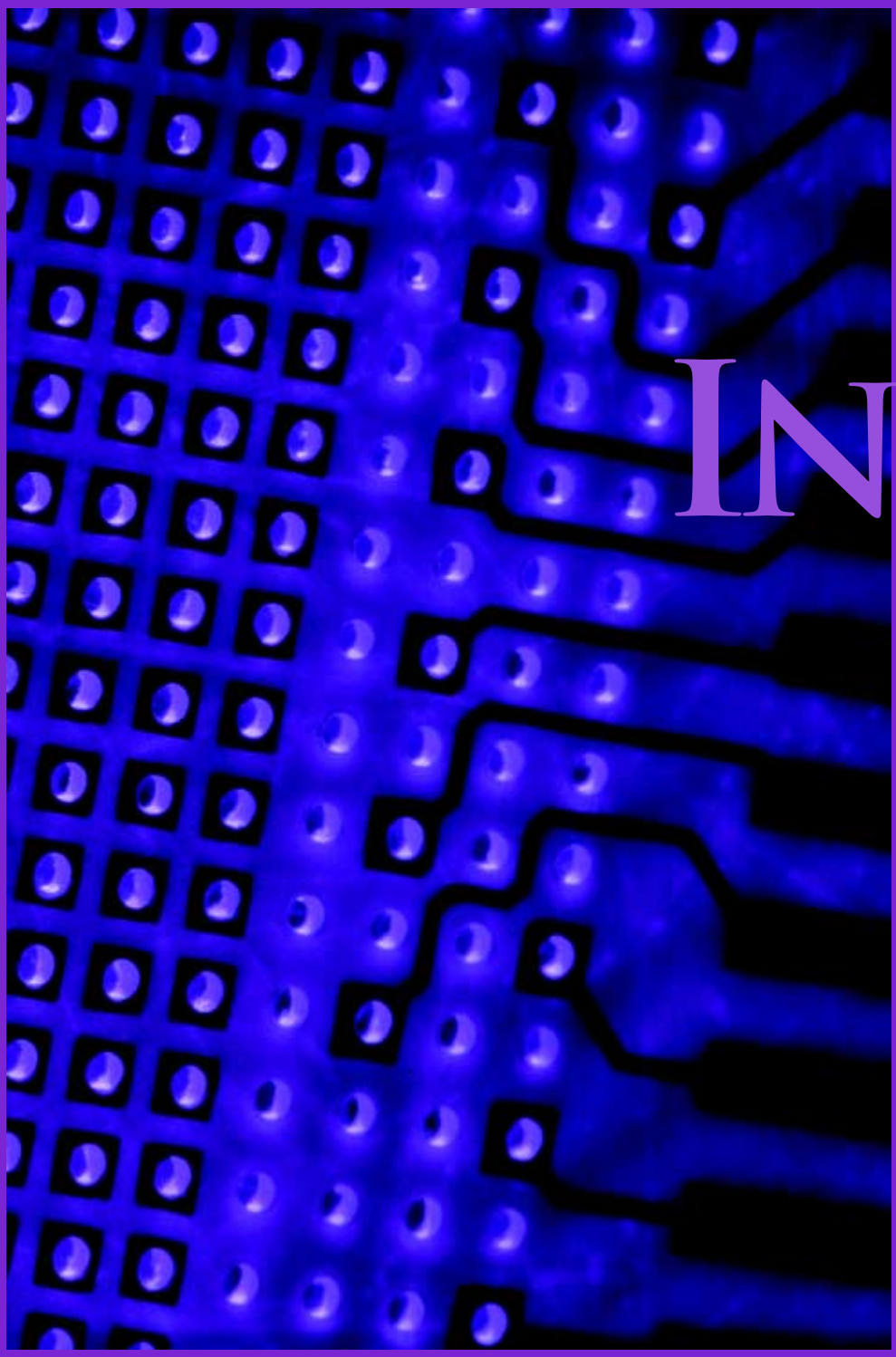
Over the last several years, there have been increased opportunities to utilize novel technologies derived from computational chemistry, molecular biology, and systems biology in toxicological risk assessment. These new areas have been referred to collectively as "Computational Toxicology," which is defined in this document as the application of mathematical and computer models and molecular biological approaches to improve the Agency's prioritization of data requirements and risk assessments. This document describes a framework for the development of a research program within the Agency's Office of Research and Development (ORD) to utilize computational toxicology to address the questions of "when and how" to test specific chemicals for hazard identification and to improve quantitative dose-response assessment.

In assessing risk associated with exposure to a chemical or other environmental stressor, there are a number of uncertainties associated with detecting and quantifying the presence of the chemical in the environment, the uptake and distribution of the chemical in the organism or human or environment, the presence of the active chemical at a systemic target site, and understanding the series of biological events that lead to the manifestation of an adverse outcome. The overall goal of ORD's Computational Toxicology Research Program is to use emerging technologies to improve quantitative risk assessment by reducing uncertainties in this source-to-outcome continuum. The three strategic objectives of the Computational Toxicology Research Program are (1) to improve linkage across the source-to-outcome continuum, (2) to develop approaches for prioritizing chemicals for subsequent screening and testing, and (3) to produce better methods and predictive models for quantitative risk assessment. The tools developed in the first objective will be critical for the research conducted in the remaining two objectives. The use of computational toxicological approaches is discussed for a number of links along the source-to-outcome continuum, including chemical transformation and metabolism, better exposure indicators, improved dose metrics, characterization of toxicity pathways, metabonomics, system biological approaches, modeling frameworks, and uncertainty analysis. Computational toxicological approaches are also needed to develop better predictive models for screening and testing including quantitative structure activity relationship (QSAR) models,

improved pollution prevention strategies, and approaches to high throughput screening. Computational toxicological approaches will also be used to address a number of research needs associated with dose-response assessment, cross-species extrapolation, and the assessment of the effects of chemical mixtures.

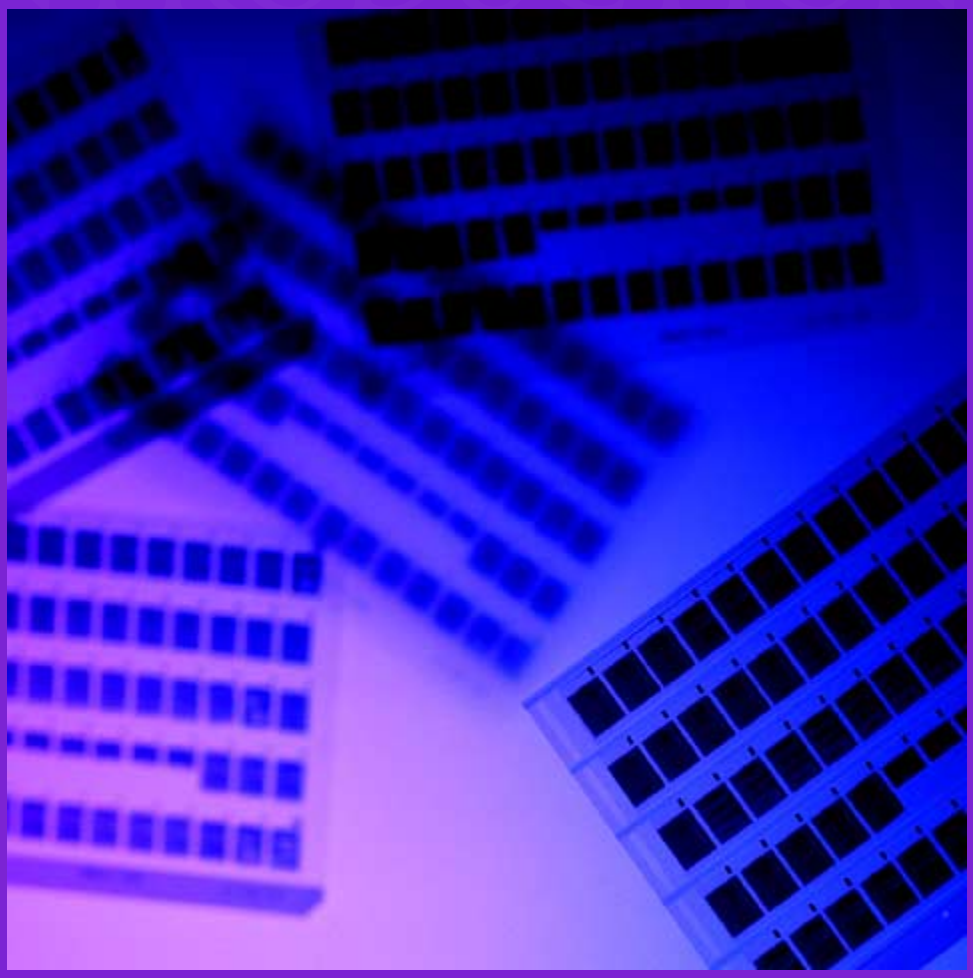
The research program at ORD currently uses many computational and biological approaches that fall under the general area of computational toxicology, and examples of such work are described in this document. Other research agencies such as the National Institute of Environmental Health Sciences and the Department of Energy have significantly greater capabilities for research on computational toxicology than the Agency. ORD has initiated discussions with these agencies in order to facilitate the development of a national approach to the use of computational procedures in toxicology.

This document is intended to identify the research needs of the Agency and the unique capabilities of ORD laboratories to provide the basis for a more focused and integrated research program in the future. To accomplish this, ORD has consulted the Agency's Science Advisory Board (SAB) on this framework and held a workshop with scientists from across ORD to discuss the content and intent of this document. Based on comments from the SAB and the workshop, ORD will develop an implementation plan to guide research on computational toxicology over the next 5-10 years.



INTRO

Introduction



I. Introduction

A. The Computational Toxicology Research Program

The overall objective of this document is to describe a framework for the development of a Computational Toxicology Research Program by the Environmental Protection Agency's (the Agency) Office of Research and Development (ORD). Computational toxicology involves the application of various mathematical and computer models to predict effects and understand the cascade of events (sometimes referred to as mode or mechanism of action) that result in an adverse response.

Computational Toxicology is the application of mathematical and computer models and molecular biological approaches to improve the Agency's prioritization of data requirements and risk assessments.

The Computational Toxicology Research Program is a technology-based, hypothesis-driven effort to increase the soundness of risk assessment decisions within the Agency. It is designed to increase the capacity to prioritize, screen, and evaluate chemicals by enhancing the ability of the Agency to predict chemicals' toxicities. Success will be measured by the ability to improve risk assessments by understanding the potential of chemicals to affect molecular and biochemical pathways of concern, i.e., their toxicity pathways.

Computational Toxicology Involves:

Computational chemistry, which refers to the physical-chemical mathematical modeling at the molecular level and includes such topics as quantum chemistry, force fields, molecular mechanics, molecular simulations, molecular modeling, molecular design, and cheminformatics;

Molecular biology, which allows for the characterization of genetic constituency and the application of wide coverage technologies such as genomics, proteomics, and metabonomics to provide the key indicators of cellular and organismal response to stressor input;

Computational biology or bioinformatics, which involves the development of molecular biology databases and the analysis of the data; and

Systems biology, which refers to the application of mathematical modeling and reasoning to the understanding of biological systems and the explanation of biological phenomena.

In the area of computational biology, recent advances in "omic" technologies make this a particularly appropriate time for such a program. Current research in this area focuses on sequencing whole genomes and understanding the complexity of cellular biology at the molecular level. The development of "omic" technologies has evolved into three scientific disciplines: genomics, which is defined as the study of genes and their function; proteomics, which is defined as the study of the full set of proteins encoded by a genome; and metabonomics, which is defined as the study of the total metabolite pool. The recent technological advances in these areas have led to the development of the field of toxicogenomics in which the effects that chemicals have on living organisms and/or the environment can be examined using genomic, proteomic, and metabonomic methods. Although the technology continues to change and improve, conducting these types of analyses is no longer a question of capability. The Agency

has traditionally used the term "mode of action" to refer to the key events and processes that lead to an adverse outcome and the term "mechanism of action" to refer to a more detailed understanding and description of events than is meant by mode of action. The potential provided by "omic" technologies is that molecular profiling at multiple levels of biological organization will lead to a depth of understanding that was not possible in the past. This document uses the term "toxicity pathway" to denote this deeper level of understanding which toxicogenomics may supply.

Parallel to efforts in computational biology, there have been major advances in computational speed and access to data. Less than a decade ago, describing the complexity of chemical behavior in biological systems was severely limited because realistic models presented combinatorial and other problems beyond the capabilities of most computers. In the field of bioinformatics, for example, major advances were made not from faster statistical analysis of data after its acquisition, but from the integration of computational and data acquisition technologies. It is now possible to consider how to evaluate the vast amounts of information generated by "omic" technologies using data-mining tools made possible by rapid advances in computational storage capacity and speed.

One area where computational toxicology has shown promise is in the discipline of physical organic chemistry known as Quantitative Structure Activity Relationships (QSAR). Application of QSAR has resulted in the development of novel predictive capabilities for representing chemical structures as a distribution of conformations and properties rather than discrete structures. Another promising area brought about by the joining of computer science, biology and medical programs is an emerging discipline known as systems biology. Systems biology has the potential to lead to the development of virtual biological systems via the development of computational models of a cell, organ, or an organism's function based upon an understanding of the component parts.

In developing this document and its subsequent implementation, it is recognized that the various technologies and tools that form the current state-of-the-science are in varying stages of maturity. How quickly the technologies are established and validated will affect the development of critical paths to solving particular problems and the timeframe in which these solutions are put into place.

B. Application of Computational Toxicology to Risk Assessment and Research

ORD's research programs support the Agency's regulatory decision making by providing scientific information for human health and ecological risk assessment. Risk assessment is the process used to evaluate the potential hazards of and exposures to environmental stressors to produce estimates of the probability that populations or individuals will be harmed by chemical exposure and to what degree. It is one component of the process by which the Agency and many other organizations recognize a potential risk and decide how to respond.

The Agency's risk assessments of chemicals rely primarily on laboratory testing on a chemical-by-chemical basis to obtain data about adverse effects and the quantitative relationship between dose level and likelihood of response. In human health risk assessment, these laboratory data are extrapolated to humans to estimate human risk. The large number of chemicals in commerce, coupled with the expense of laboratory testing, limits the application of extensive standard toxicity testing to relatively few chemicals. ORD will explore the feasibility of

using computational approaches to improve quantitative dose-response assessment and the development of sensitive and specific tests for hazard identification. Preliminary efforts in this area began in FY02 with a congressional reprogramming action which directed the Agency to explore the use of alternative methods to animal testing for hazard identification. ORD interpreted this as an opportunity to evaluate genomic and computational tools for screening purposes, and it initiated several research projects on endocrine disrupting chemicals (EDCs) as a "proof-of-concept" effort. Endocrine disruptors were selected because it was felt that a considerable amount of knowledge concerning mechanisms of actions and toxicity pathways existed for this class of environmental pollutants. This feasibility effort is described in greater detail in III.A.

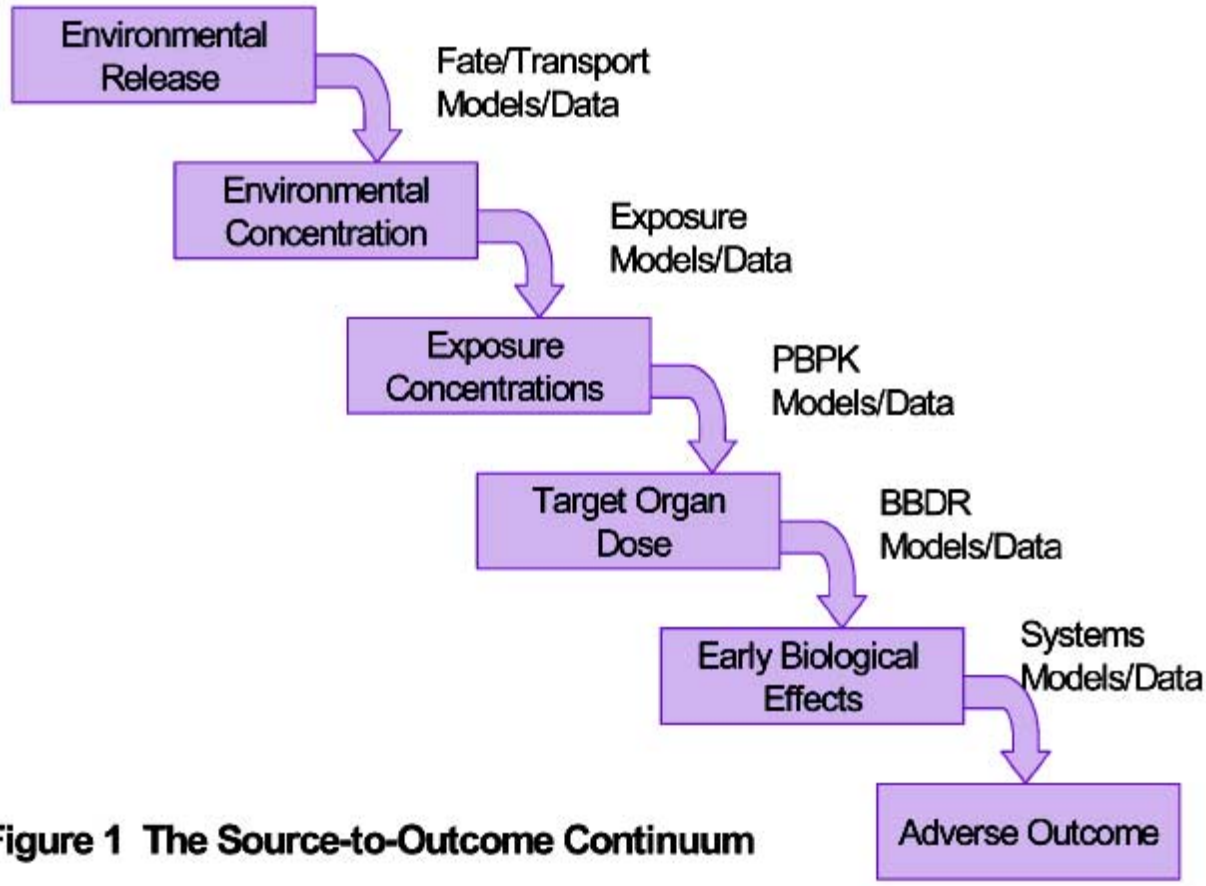


Figure 1 The Source-to-Outcome Continuum

C. Overall Goal and Strategic Objectives

It is useful to envision the risk assessment paradigm as a continuum of events leading from release in the environment to adverse effect. Figure 1 is a simplification of this concept showing points along the continuum where a measurement or an observation can be made. The arrows between the boxes represent a cascade of events that lead from one measurable event to the next. ORD's research program focuses on learning more about the processes that lead from exposure to adverse outcome in order to allow the Agency to perform better risk assessments.

Objectives of the Computational Toxicology Research Program

- Improve Linkages in Source-to-Outcome Continuum
- Provide Predictive Models for Hazard Identification
- Enhance Quantitative Risk Assessment

The overall goal of ORD's Computational Toxicology Research Program is to use the tools of modern chemistry, biology, and computing to provide the Agency with approaches to improve quantitative risk assessments and to reduce the uncertainties in the source-to-outcome continuum. To meet this goal, ORD has identified three strategic objectives for the Computational Toxicology Research Program. First, research is needed to develop improved linkages across the source-to-outcome continuum. Understanding those linkages will decrease uncertainties in assessing risk to human health and the environment. The tools developed in the first objective will be important for the second objective which is to conduct research to develop strategies for prioritizing chemicals for subsequent screening and testing. The current approach for screening and testing chemicals requires extensive resources. Therefore, an approach must be developed to determine which chemicals or classes of chemicals should be screened and tested first. Finally, research is needed to develop better methods and predictive models for quantitative risk assessment because current approaches take too long and are too costly. Benefits of this program include the identification of molecular indicators of exposure and toxicity that can be applied to other areas such as epidemiology, the harmonization of cancer and non-cancer risk assessments, and the integration of human and ecological risk assessments. The following sections of this document describe how ORD is currently using, and how it proposes to use, emerging technologies associated with computational toxicology to address the Agency's needs for approaches to screen and test more efficiently and to improve quantitative risk assessment.

Resea

Research Needs

II. Research Needs

and Applications of Computational Toxicology to Goals

A. Improve Linkages in the Source-to-Outcome Paradigm

1. Chemical Transformation and Metabolism

At several points along the source-to-outcome continuum, it is critical to accurately model the fate of chemical stressors to determine the level of exposure to an organism. It is also crucial to accurately model the metabolism of a chemical inside the target organism because it is often a metabolite of the original stressor that induces a biological event. In many cases, the reaction processes controlling the fate of chemicals outside of the organism are similar if not identical to those reaction processes controlling metabolism within the organism. For example, enzyme-mediated processes such as redox reactions and hydrolysis that result in the formation of reactive intermediates (i.e., radicals, carbenes) that react irreversibly with biological receptors (e.g., DNA) are often the rate-determining processes controlling the fate of these chemicals in natural aquatic ecosystems. Consequently, the process of developing and refining simulators for environmental transformation and metabolism will have many commonalities with understanding metabolism within organisms.

a. Chemical Transformation in Ecosystems

The state of chemical fate and transport modeling for exposure assessment has advanced significantly in recent years. For example, it is now possible to forecast many of the physicochemical properties that ultimately govern chemical transformation. Nonetheless, many unknowns and uncertainties remain, and ORD continues to conduct research aimed at reducing them. There are several key areas of uncertainty, however, that can be reduced greatly by informing and validating fate models with molecular indicators of exposure. The rapid advances in "exposure genomics" (see Section II.A.2) will provide early signs of chemical exposure based on changes in gene expression, which will lead to the development of a new array of molecular indicators that can guide chemical fate and metabolism studies. The integration of genomics and molecular indicators into chemical fate studies may improve linkages in the source-to-outcome continuum.

Chemical Fate Models

- Determine minimal concentrations at which biological events occur
- Identify biologically relevant chemical(s) in mixtures
- Identify crucial biotransformations in the environment

The application of molecular indicators to chemical fate studies is several-fold. For chemicals that trigger biological events of concern, molecular indicators can be used to determine the minimal concentration at which biological events occur. This approach will narrow the task of the exposure models to answering only the question of whether the toxicant (i.e., parent or reaction product) is above or below this minimal concentration (i.e., it will "bound" the model). Narrowing the model requirements can significantly increase certainty in the risk assessment process. Another area to address is the elucidation of the biologically relevant components

in chemical mixtures by measuring changes in gene expression in exposed organisms. This application of molecular indicators can focus exposure models on a much smaller subset of candidate chemicals, including those potentially linked to initiation of adverse effects. Finally, molecular exposure indicators can be used to advance our understanding and ability to model biotransformations of chemicals in ecosystems. Biotransformation is widely recognized as the largest uncertainty in exposure modeling, and accurate models to predict and describe biotransformation have eluded scientists because the universe of enzyme reactions is so large. Gene expression tools will be used to narrow down this universe only to those that are biologically relevant, and this will enable more accurate and meaningful prediction of significant chemical transformations.

B. Chemical Metabolism

Many toxic effects result from metabolic activation of parent chemicals that forms metabolites which are much more toxic than the parent. Moreover, many cross-species differences in toxic effects are the result of differences in detoxification. Consequently, an accurate computerized simulator of metabolism in the liver and other target tissues (e.g., kidney) is essential to meet the objectives of the program. The primary goal of this research is the development of a computational system that will predict and prioritize metabolic pathways for liver metabolism.

Metabolic Simulator

- Build libraries of relevant metabolic transformation pathways
- Develop high quality data metabolic maps
- Provide probability indices for substructural units biotransformations

The first step in the development of a metabolic simulator is to create a library of all known metabolic transformations which are nested according to the substructural elements being transformed. Algorithms will then be used to recognize the relevant substructural units in a chemical of concern that can undergo metabolism. The transformation products from each possible reaction are then stored as a list of first-level metabolites. Each metabolite, in turn, is subsequently submitted to the substructural matching routine in order to generate a set of second-level metabolites from each first-level metabolite. The process is continued until the metabolic map is completed.

This approach to simulating metabolism tends to identify many metabolic candidates that are ultimately improbable because of kinetic considerations. This problem can be overcome by associating a probability with each substructural transformation process in the library. Such transformation probabilities can be derived statistically from a library of high quality metabolic maps. Unfortunately, currently available metabolism libraries have significant gaps in relative rates for many important metabolic reactions. Identification of these gaps will direct the generation of new high quality data on metabolism. These data will be generated using traditional experimental methods and new advanced analytical techniques [e.g., wide-bore, high-resolution nuclear magnetic resonance (NMR)] for measuring metabolic rate constants and identifying metabolites in vivo and in vitro. Mechanism-based predictive methods can also be used to fill some important data gaps.

2. Exposure Indicators

Exposure assessment has historically been based on the use of chemical analysis data to generate exposure models. While the biological activity of chemicals has been recognized

as important for exposure assessments, the measurement of such activity has largely been limited to whole organism toxicity tests. Considerably less research has been done using *in vitro* tests to assess specific types of biological activity present in whole or fractionated environmental samples. Current *in vitro* capabilities are not sufficiently validated to address exposure assessment.

There is a need to develop cellular and molecular indicators of exposure that can be used to assess the vulnerability of humans and wildlife to single and multiple pathways of exposure to chemicals in the environment. However, correlation of such indicators will require a greater understanding of the linkage between the cellular and molecular indicators with specific cellular and tissue-level effects (e.g., reproductive or neurologic toxicity) or outcomes (e.g., fertility or neurological disease).

Exposure Indicators

- Few environmental stressors have specific or sensitive indicators
- Exposure indicators are poorly correlated with effects
- Molecular indicators could validate fate and transformation models
- Crucial for mixtures risk assessment
- Essential for integrated approach to risk assessment

The indeterminate condition of exposure indicator research stands to change remarkably with attempts to link molecular biological technologies with cellular or tissue effects and outcomes. The Computational Toxicology Research Program aims to develop a platform or sequence of approaches through which "the earliest recognizable signatures of exposure" (i.e., unique patterns of up- and down-regulation of genes) can be identified for scores of different stressors; become user-friendly procedures; are demonstrated in case studies; and are incorporated into Agency, State, and Regional studies supported by the Agency's Environmental Monitoring and Assessment Program and other programs. Acting on the tenet that "any response to" or "effect from" a stressor will involve changes in the expression of some genes, it is hypothesized that gene discovery and DNA microarray synthesis and use will provide a window on hundreds of changes that may or may not be linked to downstream cascades of activity responsible for adverse effects. Bioinformatic tools will be used to discriminate unique signatures and families of signatures indicative of stressors or groups of stressors. The scope of the Computational Toxicology Research Program is moving past the use of few genes in an organism, such as the ecotoxicology model fathead minnow (*Pimephales promelas*), to the use of hundreds of genes and gene homologues acquired by less direct alternate molecular methods. Ultimately, the scope of the approach will move to the level of 25,000 to 30,000 genes associated with the complete genome of selected organisms. The existence of hundreds of signal transduction pathways in cells of higher organisms, which will be elucidated through traditional biochemistry studies, will heighten the likelihood of unique exposure signatures for a great number of individual chemical stressors and families of stressors and will provide tools for future exposure studies. These studies will be both adverse effect-driven (initially, development of molecular indicators for early molecular events in sex steroid-mediated mechanisms of toxicity) and empirically based (clusters of activity identified from watershed or regional stream surveys). The development of molecular diagnostic indicators of exposure will present the opportunity for the simultaneous, near real-time measurement of biologically relevant exposures of organisms to multiple stressors in mixtures. Nanotechnological instrumentation and robotics offer the promise of extremely high throughput analysis of indicators that can allow for larger-scale exposure studies at the watershed and regional levels to be undertaken. Once discriminated, these molecular events can potentially be linked to toxicity pathways as described in Section II.A.4.

In addition to chemicals, humans and wildlife are exposed to other environmental stressors such as microorganisms. Genomic technologies offer unique opportunities to discriminate changes that occur within cells that define a particular microorganism's pathogenicity and to develop strategies for microbial source tracking. Exposure to microorganisms is also likely to be a factor in characterizing exposure of humans and wildlife to other stressors. Prediction of the outcome of chemical exposures stands to be significantly enhanced by an understanding of changes in cellular responses contributed to by both biotic and abiotic influences. This information will be especially important for the development of criteria for drinking water.

3. Dose Metrics

Qualitative and quantitative evaluations of the relationship between dose and response are key components of the quantitative risk assessment process. ORD has developed several modeling systems to assess exposure and dose. The Stochastic Human and Dose Simulation Model (SHEDS), for example, supports efforts to better understand exposure to multimedia and multi-pathway pollutants. The Exposure Dose Estimating Model (ERDEM) is a physiologically based pharmacokinetic (PK) modeling system that estimates relevant toxic doses within the body after actual and realistically simulated exposures. ERDEM can be applied to single or multiple chemical exposure scenarios. ORD also supports the development of the Modeling Environment for Total Risk Studies (MENTOR) with one of its university partners. MENTOR uses models that help to quantitatively account for pollutants along the entire source-to-outcome continuum. With new information and data from research on biological indicators, these models will be significantly enhanced. Increased knowledge of the biology of potentially toxic processes will enable ORD to enhance these models so it can better assess the effects of various and detailed exposure conditions such as co-exposure, varied patterns of exposure, and intermittent exposure.

Dose Metrics

- Dose is often inferred from stressor uptake
- Dose models stand to be enhanced with specific data on stressor interactions with molecules initiating toxicity pathways
- Genetic polymorphism data will reduce uncertainty stemming from assumptions of homogeneous populations
- Susceptibility indicators will be developed for input into exposure models

The choice of the chemical species and the actual dose metric used for the risk assessment process depends upon the particular mode or modes of action being assessed. Because the biological steps between the external exposure and various internal toxicologically relevant doses are often non-linear, PK models are often applied to estimate the effect of an exposure of interest to an observed adverse outcome. PBPK models depend on knowledge of anatomy, physiology, and biochemistry. Thus, in order for PBPK models to be used effectively, several pieces of key information are needed. Some of this information (e.g., body size, organ volumes, and blood flows) is known for several species including the human. Other important pieces of information, such as metabolic transformation rates, are chemical specific and may vary from species to species. The expense and time required to gain this information in laboratory studies has limited the use of this modeling technology, as has the lack of detailed knowledge concerning the molecular events that lead to toxicity (i.e., the toxicity pathway). The Computational Toxicology Research Program should enable broader use of PBPK models by overcoming these limitations by providing better indicators of the relevant doses and receptors within the target organism.

Through application of modern molecular tools, this research program will provide better definition of the most relevant dose metric for chemicals entering the body. For example, specific binding to a particular part of the DNA, RNA, receptors, or enzymes might be a much more relevant dose metric than simply the amount of chemical in a particular tissue. In addition, it might be possible to identify related biomarkers in easily obtained biological fluids.

It is also expected that developments in the science of genomics will greatly help define and characterize sensitive subpopulations, such as children and older individuals. It has long been recognized that not all members of the population are equally sensitive to the same environmental pollutant. While not all the factors leading to susceptibility or resistance are understood, advances in genomics and clinical medicine are showing that, for many adverse processes, a key component of the genome increases an individual's vulnerability to a specific clinical disease. Additionally, the field of metabonomics offers new opportunities to characterize variation in metabolic processes at the cellular and tissue levels. Integration of studies of stressor dose, transformation, and metabolic sequelae have the potential to provide the clearest perspective yet on relevant dose and its variation across organisms and populations.

Advances in molecular biology enable the characterization of the genetic variation (polymorphisms) within populations. Much of the current activity in molecular biology has been directed at identifying single nucleotide polymorphisms (SNPs) of key xenobiotic metabolizing enzymes. Once the metabolic significance of these SNPs is understood, they can be readily factored into PK models. In turn, these polymorphisms are beginning to be used to define the contribution of genetic variation to the overall level of variation in dose-response in populations. Ultimately, integrated research in genetics and genomics has the potential to elucidate specific altered molecular processes associated with genotypes representative of sensitive or vulnerable subpopulations. Dosimetry models such as PBPK and pharmacodynamic (PD) models can incorporate these data to reduce the uncertainties associated with assuming populations are homogeneous regarding toxic response to stressors. Indicators of susceptibility obtainable from body fluids may also be developed to provide simple methods to characterize population composition, thus refining the exposure characterization for human and ecological risk assessments.

In summary, technologies developed and applied as a result of this program have the potential to better define toxicologically relevant doses. Approaches developed by this program will also help provide the information needed to develop mechanistically based quantitative models to estimate the relevant doses and to realistically assess their impact.

4. Characterization of Toxicity Pathways

Computational toxicology techniques offer the potential to reduce uncertainties in both ecological and human health-risk assessment. However, use of key predictive toxicology tools/approaches, including PBPK and QSAR models and/or alterations in gene (or protein) expression profiles, is useful only in the context of a thorough understanding of toxicity pathways of concern (i.e., the mechanism or mode of action). Specifically, for these types of predictive methods to be useful, it is necessary to link adverse outcomes (e.g., reproductive or developmental changes or cancer) to initiating events, ideally through a cascade of biochemical and physiological changes that occur as a result of the initial interaction(s) of xenobiotics with biological molecules (e.g., receptor binding, enzyme inhibition). A particularly key aspect of this linkage is identification of the proximal (often initial) biological alteration associated with any particular toxicity pathway. For example, chemicals which bind to and activate specific nuclear

receptors elicit a relatively predictable suite of biochemical and physiological responses that are species/class-specific but culminate in very similar adverse reproductive and developmental effects across numerous vertebrate species. Identification of common initiating events, such as receptor activation, can enable the successful use of models or gene expression assays to deal with xenobiotics as classes of compounds rather than individual chemicals. Further, through understanding the cascade of events that occur as a result of receptor activation, in conjunction with accurate dosimetry predictions, it will become possible to predict adverse outcomes associated with exposure to, as yet, untested chemicals. For this to be feasible, an understanding of toxicity pathways based on discrete initiating events is needed.

Definition of toxicity pathways associated with discrete initiating events has a variety of direct benefits and implications germane to the risk assessment process. For example, the ability to associate endpoints to one another through a continuum of biological organization (i.e., across molecular, cellular, target organ, and apical endpoints) would be powerful, both for prospective and for diagnostic risk assessments. In the former case, it would be possible to better link responses at intermediate biological levels of organization to both the initiating event and the adverse outcome. In this way, it is expected that the biological effects of molecular level changes will be understood and, therefore, their use in risk assessment facilitated. In the case of diagnostic assessments, delineation of toxicity pathways would contribute directly to an understanding of the toxicological significance of alterations in markers of exposure based on changes in gene expression. In addition, as the key initiating events are identified, polymorphisms in the genes can be identified; and this information will provide insight into individual susceptibility on a dynamic level, just as how understanding SNPs in biotransformation genes provide insight into susceptibility at the kinetic level. From another perspective, knowledge of key initiating events relative to alterations in endpoints at higher levels of organization could enable a direct assessment of the technical validity of using mixture models based on similar versus dissimilar initiating events. In addition, identification of these events via alterations in gene expression could help in species extrapolation. Demonstration that toxic initiating events are similar across species would reduce uncertainty associated with extrapolation across species as knowledge of common initiating events for a chemical or class of chemicals would focus the challenge of extrapolation across species on comparative dosimetry. An underlying assumption of this approach is that there is a relatively finite number of key initiating events and that these can be understood and characterized using wide coverage molecular biological techniques combined with bioinformatic processing. One must always be aware, however, that species can display unique responses to the same perturbation; and while the initiating event may be identical across species, responses can diverge significantly such that different genes and tissues are affected in different species. Thus, the research focus may be on concordance of the initiating event rather than on the effect or site of response.

Understanding Toxicity Pathways

- Identification of discrete molecular initiating events
- Linking adverse outcomes to molecular alterations
- Elucidating linkages across biological levels of organization
- Biological basis for cross-species extrapolation
- Prediction of possible interactions for untested chemicals and mixtures

Approaches used in computational toxicology will significantly improve our ability to understand and predict how xenobiotics can interact with biological systems. Figure 2 illustrates components of a toxicity pathway for some commonly used adverse outcomes in human health and environmental risk assessment. This schematic demonstrates the linkages between biologically effective concentrations of a chemical at a receptor/ligand site that lead to

cellular and organ responses associated with an adverse effect at the individual level. These approaches can then be used for identifying and utilizing profiles of gene expression linked to cellular alterations and adverse effects or outcome.

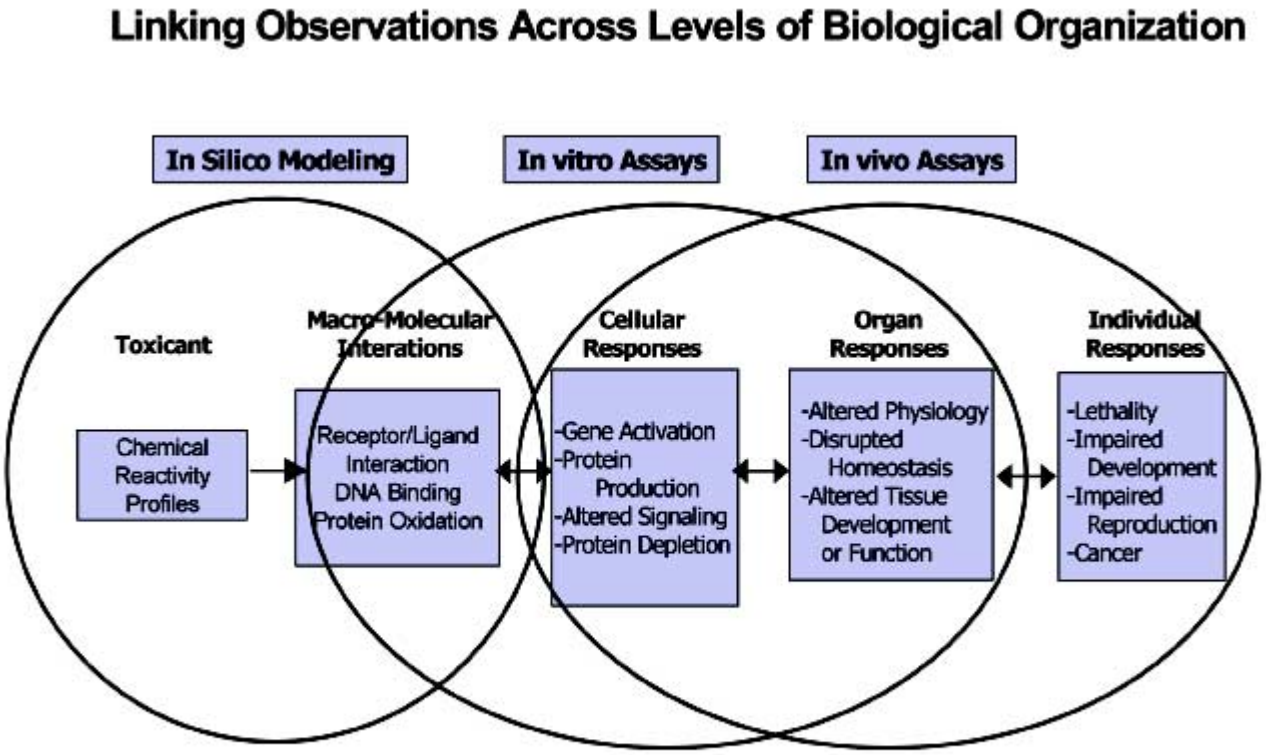


Figure 2 An Example of a Toxicity Pathway

5. Metabonomics

Genomics and proteomics allow for the measurement of response to chemicals on the genetic and cellular protein level, respectively; however, neither provide a complete description of metabolism and chemical toxicity. For example, in some instances a xenobiotic may elicit changes in gene and protein expression that are compensated for elsewhere and result in no net change to the organism (i.e., no change in endogenous metabolite profile). To fully understand xenobiotic metabolism and toxicity in the context of genomics and proteomics, it is crucial to understand the metabolic status of the whole organism. The use of metabonomics, the multi-parametric measurement of metabolites in living systems due to physiological stimuli or genetic modification, provides such a means by augmenting and complementing genomic and proteomic responses to xenobiotic exposure and by providing a connection between genomics and proteomics with tissue function. The ability to conduct metabonomic studies depends on the application of advanced analytical techniques such as high-resolution NMR spectroscopy and multi-variable statistical programs. ORD is in the process of purchasing a wide-bore 600 MHZ NMR for metabonomic analysis in support of the Computational Toxicology Research Program.

Metabonomics

- Elucidate changes in chemical-induced metabolic patterns for range of endogenous metabolites
- Generate NMR spectral profiles for chemicals
- Build models to evaluate effect of novel chemicals on endogenous metabolites

The application of metabonomics to toxicity testing involves the elucidation of changes in metabolic patterns associated with chemical toxicity based on the measurement of component profiles in biofluids (i.e., urine), cells, or tissues, and enables the generation of spectral profiles for a wide range of metabolites. NMR pattern-recognition technology associates target organ toxicity with NMR spectral patterns and enables the generation of spectral profiles for a wide range of endogenous metabolites. Metabolite profiles (e.g., endogenous metabolites such as creatine, lactate, glutathione) could provide a measure of the real outcome of potential changes as the result of xenobiotic exposure.

The application of metabonomics can provide mechanistic information that could have significant implications for the risk assessment process. Assuming that groups of compounds induce similar changes in gene, protein, and metabolite profiles, it should be possible to classify compounds based upon their profiles. Assessment of the profiles should also help scientists understand the key initiating events that are involved in activating critical toxicity pathways. By including a large number of compounds with known toxicity in databases, ORD can build predictive models for comparing and evaluating expression profiles of novel compounds. For example, association of a given toxic endpoint with a characteristic shift in cellular metabolites could provide a fingerprint that is characteristic of a specific mechanism of toxicity. Once diagnostic fingerprints are defined for different mechanisms, the metabolite pattern for a toxic chemical having an unknown mechanism could be compared to the database. This could provide a very powerful tool for categorizing toxicants according to a toxicity pathway.

6. Systems Biology

Conventional molecular biology strives to examine key events at increasingly finer levels of detail. The Computational Toxicology Research Program, combined with the work being conducted by a number of outside organizations, will provide a wealth of information on the effects of toxicants by using genomic, proteomic, and metabonomic techniques. In order to be most useful, this information must be integrated into a coherent picture. Systems biology is a new field of science that uses computational methods to reconstruct an integrated physiologic and biochemical model of an organism's or cell's biology. The approach of systems biology is similar to developing a wiring diagram for a complicated electrical system or an engineering diagram of how a vehicle is put together and how the different parts interact and function together. Analogously, systems biology is targeted at studying how normal biological processes are governed and how alterations can lead to diseases or other unwanted outcomes.

Systems Biology

- Computational models that reconstruct a cell, organ or organism's function from component parts
- Allows validation and simulator experiments that build confidence in predictive ability of adverse effects

Understanding how a normal cell or organism works is key to understanding how toxicants can cause changes in the cell. For example, in developing a biologically based dose response (BBDR) model for the developmental effects of 5-fluorouracil, ORD researchers were able to describe the effect of this chemotherapeutic on thymidylate synthetase activity (its target enzyme), on subsequent nucleotide pool perturbations, on alterations in cell cycle times, and ultimately on the size of the fore-limb bud. However, the investigators could not describe why the fourth digit was the most affected because not enough was known about the normal biology of limb development to understand how the preceding events altered the developmental programming. In this example, a systems biology approach is necessary to understand the underlying biology of limb development in order to better understand how the outcome of concern actually resulted from the precedent biochemical and cellular events.

A systems biology approach will enable the integration of disparate data developed by biologists, computer scientists, chemists, engineers, mathematicians, and physicists to construct models of organismal function and how organisms respond to a toxic insult. A choice has to be made about the scale and level of detail for each systems biology model, and it is likely that models useful to the Agency will be built by integrating individual subcomponents into a larger system. Once these models are developed, hypotheses can be developed and tested through virtual simulations prior to designing targeted experiments to validate and inform the models. An integral part of the Computational Toxicology Research Program will be the use of relevant model organisms to expand our understanding of the regulation of biological processes and how toxicants can perturb these processes. In particular, cell signaling systems are receiving considerable attention in systems biology, and this might be a promising approach. An initial step to designing ORD's systems biology efforts might be to organize a workshop to help identify promising areas for research and development. To supplement its intramural efforts, ORD expects that the extramural grants [Science to Achieve Results (STAR)] program will be able to make contributions in filling the research needs presented by systems biology.

7. Modeling Frameworks and Uncertainty Analysis

The Computational Toxicology Research Program requires a modeling framework to develop a functional tool for prioritizing chemicals for subsequent screening/testing and enhancing quantitative risk assessment. Modeling frameworks are the software infrastructure required to facilitate modern environmental modeling. Modeling solutions to regulatory-based assessment needs require the development and application of science-based models and databases that span the source-to-outcome continuum. Modeling frameworks contain and manage the coordinated execution and data exchange of numerous science-based models. They also facilitate access to external data sources, model output data analysis, and user interfaces.

Establishment of a modeling framework (based on existing and proven technologies) that will standardize the format and interchange protocols for all information generated via computer simulation for the Computational Toxicology Research Program will be a critical undertaking. The technology will contain (or access via Internet) models for simulating the environmental fate and transport of chemicals (transformation simulators); human and ecological exposure; the fate and transport of chemicals within human and ecological receptors (metabolic simulators and PBPK); toxicity pathways (QSARs); and adverse outcomes (systems biology models). In support of this modeling, the technology will include database-connectivity tools for linkage to databases unique to the research program such as sequence databases, libraries of metabolic and toxic pathways, metabonomic profiles, and bio- and chemo-informatics.

Modeling Frameworks and Uncertainty Analysis

- Requires science-based models and databases
- Standardize format and interchange protocols for information generated by computer simulation
- Develop technology for linking required databases
- Develop uncertainty analysis methods
- Operating system managing numerous models
- Facilitate access to other databases and user interfaces

In addition to providing the necessary modeling technology, ORD will also need to target the development of uncertainty analysis methods. As the scope of science needed to answer the broad questions posed by modern regulatory initiatives expands dramatically, our ability to quantify the accuracy of our model-based estimates decreases. Consequently, a major focus of

emerging scientific inquiry is related to the characterization and quantification of uncertainty in "high order systems." Efforts will be undertaken to increase awareness of the full range of potential health effects, especially non-cancer endpoints, and to fully characterize variability and reduce overall uncertainty. A supercomputer (a cluster of 150 PCs) has already been configured to facilitate the distributed computing (parallel processing) necessary to execute sophisticated modeling simulations involving numerous models and databases. This hardware infrastructure is specifically designed to support the development of a wide range of uncertainty analysis methods.

B. Provide Predictive Models for Hazard Identification

As discussed previously, the Agency needs to develop predictive models for hazard identification. This section describes three areas where computational approaches are being considered or where such methods could prove fruitful.

1. QSAR and Other Computational Approaches

Like its physico-chemical properties, the biological activities of a chemical are the result of molecular interactions between the chemical and its immediate environment. When models for specific molecular interactions are developed, the activity of chemicals with respect to those interactions can be estimated directly from chemical structure. QSAR models could serve as important tools to screen untested chemicals for their potential to interact with hundreds of different environments using only the chemical structure and a virtual library of chemical and toxicological models. As such, QSARs could be used to optimize laboratory testing when the number of untested chemicals exceeds the resources available for testing. QSAR may also be used to provide estimates of missing data in lower tier risk assessment. For example, in the case of the initial screening of chemicals for their potential to disrupt the endocrine system, models of molecular interactions with critical receptors and enzymes could be used to develop a series of computational methods to classify each chemical based on its likelihood of binding to a receptor or inhibiting a crucial enzyme. The intent of screening using QSAR is to offer a list of chemicals most likely to test positive in standard toxicity screening assays. It is recognized that these models will have to undergo extensive validation in terms of the descriptors associated with structure, the chemical space that they are applicable to, and the quality of the data in the training sets.

QSAR and Other Computational Approaches

- Quantifying physico-chemical parameters to predict fate
- Identification of potential hazard in absence of empirical data
- Prioritizing large groups of chemical for later testing
- Framework for optimized use of "omic" data
- Estimate missing parameters for untested chemicals

Another application of QSAR is to estimate the toxicity of untested chemicals directly from chemical structure. Some chemical properties and reactions can be directly related to structural parameters and, for some of these, QSAR has been used as a cost-effective surrogate for routine laboratory experiments. For certain applications such as fate and effects modeling which require expensive laboratory measurements of chemical properties, the structure-property relationships are now sufficiently reliable that QSAR estimates rather than measured values are widely used. QSAR can also be useful in estimating potency within a class of chemicals relative to acute toxicity endpoints.

Emerging "omics" technologies offer excellent potential to generate information that will inform and improve the QSAR modeling process. Specifically, alterations in gene expression can be used to identify toxicity pathways and associated key molecular initiating events. Once initiating events are established for any given toxicity pathway/adverse outcome, predictive modeling can form the basis for dealing with large numbers of chemicals in a relatively rapid fashion. Specifically, there is little question that, given high-quality datasets, current QSAR modeling methods can effectively predict discrete biological phenomenon at the molecular level in terms of interactions of chemicals with different classes of lipids, nucleotides, or proteins. Therefore, the key is defining the biological phenomenon for which data should be collected. If molecular initiating events can be identified relative to specific adverse outcomes, appropriate *in vitro* or *in vivo* assay systems (e.g., receptor binding assays) can be identified and developed to serve as a basis for generating the data needed for robust models.

The relationship between the generation of genomics information and QSAR modeling can be depicted in a linear fashion. It is a process, however, that is iterative in nature with multiple potential points of "entry" in terms of genomics or modeling components. For example, if the initiating event through which a chemical elicits adverse effects is completely unknown, genomic approaches in which large numbers of expressed genes are assessed can be used in a "discovery" mode. This information may effectively identify the unknown chemical as similar to other previously tested chemicals for which there is an understanding of toxicity pathways and associated initiating events. Alternatively, genomic information may serve as the basis for defining previously unknown initiating events and may serve as the starting point for delineating new or alternate toxicity pathway(s). As key initiating events are identified, appropriate data "generation" assays can then be developed to provide data for QSAR and other computational models capable of predicting either acute or chronic toxicity.

2. Pollution Prevention Strategies

In support of pollution prevention strategies, ORD is developing methods to estimate the potential environmental impact of chemicals that are released into the environment. These methods are used to evaluate chemicals for potential harm to both humans and the environment in a life-cycle assessment framework. These chemical evaluations are performed over a wide range of environmental impact categories, including human health (acute, chronic, and carcinogenic indicators), aquatic health (acute and chronic indicators), terrestrial health, global warming, ozone depletion, smog formation, acid rain, eutrophication, and natural resource depletion. To this end, several pollution prevention tools have been developed for a variety of uses. Depending on the level of analysis desired, the complexity of the model for evaluating human and ecological health concerns will dictate the type of data required. These data may be collected from acute toxicity studies, detailed systems biology models, fate and transport models, exposure models, or chronic toxicity studies. Regardless of the level of sophistication in the models, the final impact indicators (e.g., a broad range of mid-point effects or final outcomes, such as human deaths, human illnesses, crop damage, water quality issues, air quality issues) could be used to compare a large number of chemicals. QSAR have also been developed by ORD to provide estimations for toxicity values for those chemicals which have no experimental toxicity data. For example, QSAR can be used to estimate endpoints, such as the 96-hr LC_{50} values for fathead minnows and the oral rat LD_{50} values. These estimated values can then be incorporated

Pollution Prevention Strategies

- Methods to estimate potential impact after release into environment
- Final indicators to compare large numbers of chemicals

into various environmental models and used in pollution prevention strategies. Such tools will greatly benefit from increased certainty in the fate and transport models (as discussed in Section II.A.1.a), and they will be improved by a better understanding of metabolic maps of chemicals (as discussed in Section II.A.1.b). These tools will also benefit from the enhancement of the quantitative risk assessment process in areas such as increased knowledge of dose-response assessments (as discussed in Section II.C.2) and cross-species extrapolations (as discussed in Section II.C.2).

3. High Throughput Screening

Applications of new molecular and other technological advances hold promise for the development of high throughput screens (HTPS). It has been suggested that HTPS be used as a rapid, efficient means to provide preliminary endocrine-effects data on chemicals considered in the endocrine disruptors screening and testing program. In view of the estimated 87,000 chemicals under consideration, it would be beneficial if rapid, HTPS systems could be developed to assist in the prioritization of chemicals for further testing. Because all processes are automated and can be programmed to run continuously in HTPS, large numbers of samples can be screened in a relatively short period of time using this technology. New approaches have the potential to make significant advances over existing EDC screens in terms of speed, high-throughput capability, sensitivity, reproducibility, and reduction in animal usage in a screening and testing program. HTPS will be a valuable tool to help elucidate and characterize toxicity pathways (see section II.A.4). Approaches under development could focus on classic ligand-steroid receptor-coregulator/cofactor interactions; non-genomic mechanisms of steroid hormone action; or mechanisms involving synthesis, metabolism or degradation of estrogens, androgens, and thyroid hormones. Furthermore, some HTPS approaches may be flexible and versatile enough to allow for screening to be carried out across vertebrate classes. This will help scientists address cross-species extrapolation issues (see section II.C.2.b).

High Throughput Screening

- Vast chemical inventory not tested
- Rapid, efficient means to provide preliminary data
- Recommended for Endocrine Disruptors

C. Enhance Quantitative Risk Assessment

Computational toxicology has the potential to enhance the Agency's current risk assessment methods and to contribute to the development of new methods that are consistent across endpoints and species. One aspect of the ORD program will be to develop broadly applicable risk assessment methods that take advantage of new technologies and the data generated by the Computational Toxicology Research Program.

1. Applying Computational Toxicology in Quantitative Risk Assessment

ORD has ongoing research in two areas that can be related to risk assessment: computational chemistry (i.e., QSAR) and mathematical biology (i.e., PBPK/BBDR modeling). The Agency's Program and Regional Offices are responsible for hundreds of site-specific risk assessments which are essential to inform Agency priority-setting. Many of the site-specific environmental issues involve chemicals about which there are insufficient data. The Agency is currently exploring the use of QSAR to estimate toxicity benchmarks for such chemicals, as well as the application of PBPK and BBDR models to risk assessment. It is likely that a systems

biology approach to dose-response modeling will facilitate the integration of PK and PD models. In addition, QSAR can be used to estimate parameters of PBPK models and cellular response, and such models are being developed to link QSAR to systemic (whole-organism) dynamics. Approaches are also being explored to determine how the use of mechanistic information can lead to human health risk assessment approaches that are consistent across all endpoints and replace the assumptions that some endpoints are non-threshold effects (cancer) and that other endpoints are threshold effects (non-cancer). The Agency is also exploring data-based approaches to adjusting default uncertainty factors in current non-cancer risk assessments. As a first step, PBPK data and models are being used to adjust default uncertainty factors used in inter- and intraspecies extrapolation and for developing an assessment of target organ dose for use in dose-response assessment of chemical mixtures. Similar efforts for the PD uncertainties would be the next step.

In the field of computational biology, the Agency is just beginning to consider application of genomic/proteomic data to risk assessments. The Agency's Interim Genomics Policy outlines the current state of the application of genomics data (which is defined to include proteomics and transciptomics) in risk assessment. Genomics data may be considered in Agency decision-making; but they are insufficient, standing alone, to inform decisions about environmental risk. It is essential, therefore, that the Agency consider how the information generated by this new technology will be utilized in human health and ecological risk assessment. While there are many issues that need to be considered before routinely adopting and/or replacing existing data requirements with expression data for Agency risk assessments, an overarching goal of the Computational Toxicology Research Program is to address these questions in order to develop ways to apply computational toxicology in quantitative risk assessment.

2. Examples of Applications of Computational Toxicology to Quantitative Risk Assessment

There are many potential applications of computational toxicology in quantitative risk assessment. The following sections discuss three areas: dose-response assessment, interspecies extrapolation, and toxicity of chemical mixtures.

a. Dose-Response Assessment

Genomics/proteomics technologies have important implications for health and ecological risk assessment, and there is a need to develop methods that use these data to improve quantitative dose-response assessment. One important area is the use of emerging technologies to determine the shape of the dose-response curve in the low dose range based upon *in vivo* and *in vitro* data that can be shown to be correlated with low dose adverse effects. Studies using emerging technologies may also result in the identification of useful (simple, sensitive, and relevant) biomarkers of effect so that they can be used in dose response studies, not just hazard identification, to more accurately diagnose effects in the low dose range and for compounds with weaker potencies.

Research from the Computational Toxicology Research Program may also lead to the identification of biological effects that could be used as the adverse effect for risk assessment. For example, if fetal testis endocrine function is altered such that steroid hormone production and insl3 gene expression are reduced by 50% or greater, it is possible that this change could be considered an adverse effect. Chemicals producing such effects might then be regulated on this information alone. The assumption of such an approach is that reductions in insl3 are always

associated with developmental malformations. Chemicals could then be regulated on the basis of fetal endocrine effects alone. The Agency has used endocrine data in this manner on occasion, but the approach would be strengthened by including genomic or proteomic information to support the hypothesis that a specific pathway had been sufficiently disrupted by a chemical such that adverse effects would definitely result later in life. Hopefully, genomic information will prove useful in assessing population level effects in ecosystems, as well as in predicting risk to individuals in human risk assessment.

Computational Toxicology and Quantitative Risk Assessment

- Defining the shape of the dose response at low exposures using molecular indicators of response
- Developing biomarkers for use in analysis of low dose responses
- Validating the interpretation of molecular indicators of response
- Defining the relevance of modes of action for risk assessment
- Constructing BBDR models of high priority outcomes
- Assessing population level effects in ecosystems
- Cross-species extrapolation
- Assessing toxicity of chemical mixtures
- Integrated human and ecological risk assessment

Another application of emerging technologies from the Computational Toxicology Research Program would be a determination of the relevance of the mechanism or mode of action of low dose adverse effects to humans and other species using *in vitro* and *in vivo* approaches. Such research could address the critical pathway initially involved in chemically induced adverse effects and how well the mechanism is conserved between mammals and other vertebrates.

Although scientists both inside and outside of the Agency have been proposing the application of BBDR models in risk assessment to reduce uncertainties in the process for several years, this goal has not been realized in spite of several long-term research efforts due to the complexity of the models. One question that could be addressed by the Computational Toxicology Research Program is how quantitative mechanistic genomic data could be used to develop BBDR models that produce realistic values for risk assessment.

B. Cross -Species Extrapolation

One of the major challenges in regulatory toxicology is the prediction of toxicity of a chemical(s) or classes of chemicals across species. In its risk assessments, the Agency often predicts possible effects in humans from studies in rodents and other mammalian test species; while in ecotoxicology, extrapolations to literally thousands of other species are typically based upon results of assays with a handful of surrogate test organisms. In addition, it is generally assumed that adverse effects seen in vertebrate wildlife are relevant to other species, including humans. For example, when fish display evidence of exposure to estrogens or androgens in the environment or frogs exhibit limb malformations, there are immediate concerns about potential effects on humans and other wildlife. The concept of interspecies extrapolation is based upon the knowledge that all species arise from common evolutionary ancestors and that there is a great deal of homology among animal species with regard to basic biological pathways. However, the ability to extrapolate from species to species is not trivial for two reasons. First, exposures to chemicals vary as a function of an animal's physiology and its environment. For example, humans are not exposed to estrogenic or androgenic materials in effluents to the same

Cross-Species Extrapolation

- Understand differences in exposure, uptake, and metabolism across target species
- Define toxicity pathways for common chemicals in model species

degree as fish living in or adjacent to an effluent. In a more phylogenetically similar comparison, the Florida panther is at greater risk for exposure to (and effects of) bioaccumulating contaminants than are other populations of panthers because it feeds at a higher trophic level than other populations not because of genetic differences.

Another aspect of differential dosimetry among animals is the metabolism of xenobiotics. Pathways of metabolism can differ significantly across species, even in closely related animals. In fact, in many cases, failure to metabolically activate a chemical, or conversely to rapidly detoxify it, is a primary basis for the lack of response of an animal, not species-specific differences in gene expression and subsequent toxicity. For example, the anti-androgenic action of vinclozolin or the estrogenic activity of methoxychlor requires metabolism to active metabolites. Testing either of these chemicals in a species that does not produce "active" metabolites could lead to the mistaken assumption that they would not affect endocrine function. This component of species extrapolation, i.e., comparative dosimetry, can be addressed using BBDR models that focus both and on concentrations of parent chemicals and metabolites in target tissues. This type of modeling is an integral part of the Computational Toxicology Research Program.

The second challenge to species extrapolation involves how an animal actually responds to a given dose of chemical(s) of concern. Specifically, although there is remarkable similarity in basic biology among animals, there are also significant species-specific differences in genes, proteins, biochemistry, and physiology. These differences lead to uncertainties in interspecies extrapolation. Consequently, toxicologists are generally more comfortable extrapolating among closely related species than among those that have been separated phylogenetically for a longer period. To address this aspect of species extrapolation, definition of toxicity pathways from a comparative perspective is critical; as such, the Computational Toxicology Research Program will focus on extrapolation. Specifically, characterization of toxicity pathways in well-defined animal models serves as the basis for identifying key control points (e.g., receptor-mediated signaling) that, quite possibly, would be conserved across species. This type of conservation would be expected for many receptor-based processes (as illustrated by estrogen- and androgen-controlled pathways). When control points are known, state-of-the-art molecular biology (e.g., genomic) techniques can be used to assess the degree to which extrapolation of effects associated with chemicals with specific mechanisms of action can be supported. For example, once it has been demonstrated that receptor activation by xenobiotics is key to eliciting toxicity, it will be possible to focus on comparative receptor binding studies across species as a basis for extrapolation. The critical point here is that, if pathways can be reliably defined in representative model species, it is not imperative to know a comparable amount of toxicological information in untested species, only whether the latter possess key components of pathways of concern and how these components are affected by xenobiotics.

C. Chemical Mixtures

The Agency's guidance on risk assessment of mixtures indicates that multi-chemical exposures are ubiquitous, including air and soil pollution from municipal incinerators, leakage from hazardous waste facilities and uncontrolled waste sites, and drinking water containing chemical substances formed during disinfection. The guidance also indicates that exposure scenarios are very diverse. Two approaches are generally used to assess exposure to and effects from mixtures, i.e., the study of whole mixtures or of individual mixture components. The Food Quality Protection Act (FQPA) specifies that cumulative effects be addressed for exposure to multiple chemicals acting by the same mechanism. However, a number of uncertainties exist regarding the level of mechanistic similarity among chemicals. Considerable complexity arises even when examining interactions among chemicals that have "estrogenic" activity. While whole mixture studies offer environmental relevance, it has been recommended that whole mixture screening go forward only after screening had been conducted for a number of individual chemicals. Mixture research over decades is poised to take advantage of unifying principles from shared chemical and biological mechanistic research. New tools developed in the "omics" and computational fields show huge potential for employing molecular profiling to understand the complexity of mixture exposure and elucidating the mechanisms underlying biotransformation, uptake, distribution, and response. Technological advances now enable study of the joint and interactive properties of mixtures and definition of those characteristics that are sufficiently similar to allow extrapolation of data from one mixture to another. They have the potential to allow identification of emergent properties of real-world mixtures of xenobiotics at environmental exposure concentrations rather than from defined mixtures at high (toxic) concentration. Hypotheses can be tested to identify mixture classes (by activity or structure) that are amenable to component approaches. Mixtures assessment guidelines stand to become much more uniform, relevant, and easily applicable to risk assessments of real world exposure scenarios.

Chemical Mixtures

- FQPA specifies that cumulative effects must be addressed for exposure to multiple chemicals acting by a similar mechanism
- Lack of information regarding mechanism or mode of action for risk assessment of mixtures

CURRENT

Current Activities

III. Current Activities

The previous section described several areas in which a program on computational toxicology could provide methods and models that would lead to more efficient ways to assess chemicals for screening and testing, as well as improve quantitative risk assessment. Section III demonstrates that ORD currently possesses the capability to utilize approaches necessary for the ultimate development of a Computational Toxicology Research Program in the future. The work described below is divided into three categories, including ORD's initial research to demonstrate the feasibility of the computational toxicology concept, examples of on-going research, and linkages to external research groups.

A. Proof-of -Concept: Endocrine Disrupting Chemicals (EDCs)

Initiated by an FY02 congressional mandate to explore alternatives to the use of animals in toxicological testing, ORD began a research effort to explore the use of emerging technologies and computational approaches to better prioritize chemicals for screening and testing. Because much is known about how EDCs interact with biological systems to cause adverse health, it was decided to focus the program on this class of chemicals and to conduct several proof-of-concept experiments to determine the feasibility of using computational toxicology approaches to meet an immediate Agency need. Understanding the key biological pathways affected by endocrine disrupting chemicals affords the opportunity to design approaches that are more efficient in terms of resource utilization. It also allows the Agency to extrapolate findings from a smaller set of chemicals to the broader chemical universe using the tools of computational chemistry. Thus, projects using *in silico*, *in vitro,* and *in vivo* approaches could facilitate the prioritization of chemicals for screening, reduce the need for some *in vivo* assays, and provide *in vivo* assays that have a greater breadth of coverage of endocrine alterations and/or provide better predictiveness of potential adverse health outcomes. The overall program has short-term, intermediate, and long-term goals. The effort in this proof-of-concept activity is directed at developing better tools and assays to monitor selected aspects of endocrine disruption. Success will be measured against the recommendations put forth by the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) and will provide us confidence that the approaches are applicable to other pathways of toxicity where the underlying biology is not so well understood at the present time. Descriptions of the activities follow.

Proof-of-Concept Studies with EDCs

Refine existing QSAR models
Develop *In vitro* models
Characterize toxicity pathways

Receptor Binding Models - The EDSTAC-recommended approach to screening chemicals for endocrine activity proposed the use of QSAR models of receptor binding to help prioritize chemicals for further screening. EDSTAC also recommended that the Agency undertake a study to evaluate the *a priori* predictions of available QSAR models for estrogen receptor interaction by obtaining data on competitive binding affinity for the estrogen receptor from a single laboratory using a standardized protocol on approximately 300 chemicals. This demonstration exercise found the models needed additional work to improve their sensitivity, specificity, and predictive capabilities before they could be used in a regulatory context. The lack

of complete agreement between binding information and QSAR models appears to be related to inconsistent receptor-binding information in the training sets used to initially construct the QSAR models. The values to derive the model were derived from multiple laboratories and were based on IC_{50} values, which potentially can introduce errors because non-competitive binding might be present. Therefore, under the proof-of-concept effort, ORD is taking 70 of the 300 chemicals that showed some evidence of receptor interaction and generating K_i values for each of them. This will provide an unbiased and unequivocal measure of receptor binding. Following acquisition of these data, the QSAR models will be rederived using the new training set and their predictability will again be assessed. This effort will demonstrate how international criteria for transparency, domain coverage, and model acceptance evaluations can be met while reducing the number of chemical tests needed as compared to random chemical testing approaches. Through this iterative process, ORD will test the hypothesis that given a robust data set of high quality, reliable QSAR models can be developed and used in hazard identification. Assuming success in this effort, the research will progress to accomplish similar goals for the androgen receptor (AR).

In vitro models - Knowing the toxicity pathway of concern in whole animals should allow development of simpler *in vitro* systems that can provide quick and inexpensive evaluation of the potential for chemicals to interact with that pathway. EDSTAC raised concerns about the need to evaluate the effects of chemicals on the steroidogenic pathways in Tier I screening approaches. EDSTAC addressed this data gap by recommending a combination of studies in pubertal rodents and enzyme inhibition studies on either placenta or minced testes preparations from male rats. Both of these approaches are limited in their ability to study the synthesis of the key steroids (estradiol and testosterone) from cholesterol, including the fact that they are more targeted at detecting inhibition of synthesis rather than enhanced synthesis. A human cell line (H295R) has been identified that maintains ability to synthesize estrogen from cholesterol and that addresses many of the limitations of the EDSTAC approach. Under the proof-of-concept effort, ORD is developing standard operating procedures for use of the H295R cell line to evaluate each step of steroidogenesis at the genomic, proteomic, and metabonomic level. ORD will compare the results from a set of chemicals with the EDSTAC-recommended assays to determine if this cell line affords more powerful, yet easier to achieve, answers to the potential to alter steroidogenesis.

Toxicity Pathway Characterizations - These studies are integral to the proof-of-concept effort and will consist of studying thyroid gland functioning in ecologically relevant species and by examining the integrating function of the vertebrate hypothalamic-pituitary axis in responding to the presence of EDCs.

The thyroid gland was selected as a target for the effort because it represents an endocrine organ whose function is disrupted, not by direct xenobiotic interaction with the thyroid receptor, but by interactions elsewhere in the endocrine loop (e.g., iodine uptake processes, hormone synthesis, hormone modification, and hormone metabolism). Given the fact that thyroid function can be perturbed at different points in the thyroid pathway, research is being directed at developing a suite of endpoints that could ascertain which toxicity pathway is initiated by a specific chemical. Each chemical interaction with the thyroid signaling axis would be considered a separate toxicity pathway given that the chemical structural requirements for interaction with each site are likely unique. This could facilitate the development of QSAR models for each toxicity pathway. Therefore, it is important to realize that there may be multiple pathways for any given toxicological response. It would also be advantageous to maintain the multiple toxicity pathway distinction when assessing aggregate versus cumulative chemical risk. A common mechanism of action cannot be assumed unless it is shown that mixtures of chemicals across these different toxicity pathways exhibit additive behavior when equal toxic units are

combined for joint toxicity assessments. Success here would demonstrate that a more global focus on endocrine function using genomic approaches and critical life stages can provide more meaningful information than the EDSTAC-recommended amphibian metamorphosis test, which is neither highly sensitive nor diagnostic of disturbances in the function of the thyroid.

The last area covered included in research on toxicity pathways is more exploratory in nature; i.e., it involves exploring the possibility that, through the use of genomic and proteomic evaluations, a single *in vivo* test for endocrine-disruptor activity can be developed that will meet the essential requirements for an EDC screening assay. This activity is based on the rationale that the central nervous system (CNS) contains all of the relevant receptor and enzymatic target sites of interest. Taking advantage of the fact that there is a wealth of information available concerning the physiological regulation of the thyroid and adrenal and gonadal axes by the hypothalamus and the pituitary glands, it may be possible to empirically test the extent to which changes in these endocrine systems are sensed and responded to by the CNS. It is anticipated that the development of a genomic response profile following exposure to EDCs of known action will provide the means to identify the target pathways that lead to altered reproductive/thyroid/adrenal function. This approach, if successful, would be superior to the current proposed male and female pubertal assays because it could predict all relevant target pathways whereas the current assays can not necessarily identify specific CNS target sites. Furthermore, this approach has the potential to advance beyond screening and may incorporate elements important to establishing efficient Tier 2 tests meant to characterize dose response relationships for endpoints of utility in risk assessment.

It is expected that the lessons learned from these EDC proof-of-concept projects will be immediately transferable to future studies on the potential of non-endocrine disrupting chemicals to affect other biological systems and signaling pathways. It is already apparent that expansion beyond the evaluation of EDCs to other toxicity pathways will encompass several key steps: (1) elucidation of pathways of chemical toxicity, from initiating event to adverse outcome in individuals or populations; (2) identification of key assays indicative of toxicity pathways that provide a means to efficiently, wisely, and with minimal animal testing, extrapolate across chemicals and species; and (3) the application of an iterative QSAR strategic test design, where models are developed and strategically improved, with minimal testing, until criteria for regulatory acceptance are met. This strategy will maximize the likelihood that predictive models developed under this approach will have a solid foundation in scientific principles and will provide the Agency with defensible approaches to priority setting.

While not currently underway, a likely second proof-of-concept project would be one that transcends all aspects of the source-to-outcome continuum and would therefore be able to demonstrate the advantages/long term or far reaching benefits of a computational toxicology approach throughout the entire risk assessment spectrum. Experience gained in the EDC proof-of-concept research will be invaluable in helping to formulate such a project.

B. Internal Linkages

There are several on-going research projects in ORD's core and problem-driven research program that are supportive of, or which will benefit from, efforts in computational toxicology. ORD's core research program aims to provide broad, fundamental scientific information to improve understanding of human health issues. In particular, research associated with ORD's Human Health Research Strategy will support major components of the Computational Toxicology Research Program. For example, research to develop a common approach for

the use of mechanistic data in cancer and non-cancer risk assessment will provide significant information concerning toxicity pathways for high priority environmental chemicals. Likewise, efforts to harmonize approaches between human and ecological risk assessment will also provide a conceptual basis for identifying linkages in the source-to-dose paradigm. It is expected that advances in computational toxicology will lead to reductions in uncertainty associated with extrapolation across species and will aid in protecting susceptible subpopulations, such as children and older individuals.

In order to illustrate the complementary nature of the various research programs within ORD, we have chosen to describe projects that have a high regulatory impact (i.e., cancer, reproductive, pulmonary, and neurotoxicity) and that are associated with specific environmentally relevant contaminants [e.g., particulate matter, arsenic, disinfectant by-products (DBPs), EDCs] or risk assessment issues. As examples, the two following sections describe on-going research on the human health effects of DBPs and the assessment of an aquatic species in ecological risk assessment. Appendix A provides additional examples of ORD projects that support the Computational Toxicology Research Program.

1. Human Health Research

Under the Information Collection Rule, DBP occurrence data for major water systems across the country are becoming available. Data are being collected on DBPs that have been predicted to have an adverse health effect but which have little or no previous quantitative occurrence information. In addition to this drinking water, utility-based information, data on individual water use and biomarker development are being developed to increase the precision of exposure assessment. These data will help to identify classes of DBPs and candidate model compounds that are present at the greatest concentration. Because DBP concentrations at the tap can vary widely and be dramatically different from reported averages from water utilities, a series of measurements on drinking water at different points in the distribution system is needed for better informed exposure models. Current exposure modeling from ingestion of disinfected drinking water reflects only the general characteristics of the source water and disinfection process. To provide more refined exposure assessments, improved exposure models are currently being developed using more quantitative data on DBP component analysis and to more accurately predict of the concentration of the most toxic DBPs at the tap. This work will then be combined with response models to develop a computational approach for predicting potential adverse effects based on the predicted and actual characteristics of the mixture of DBPs.

Examples of Chemicals Under Study by ORD

- Arsenic
- Disinfection By-Products
- Endocrine Disruptors
- Particulate Matter
- Pesticides

Computational chemistry methods and molecular modeling approaches are being employed to compute a variety of structural, electronic, and reactivity characteristics of DBPs; their postulated metabolites; and/or adducts for input and consideration in the development of structure-activity models. Efforts currently focus on differences based on the bromine content within a class of DBPs. The central issue, in this case, is the role of bromination in determining and modulating biological activity within these classes. Toxicity information, in conjunction with known principles of the organic chemistry of halogen-facilitated reactions, will be applied to understand the possible range of mechanism-based reactivity and how such reactivities are modulated by chlorination and bromination. These modulations include DNA adduct formation

and carcinogenicity. Computational chemistry and SAR modeling will be used to prioritize chemicals for testing and aid in the preliminary hazard assessment of DBPs for which little or no toxicity data are available. These approaches can also be used to generate mechanism-based SAR hypotheses pertaining to classes of halogenated drinking water contaminants and to help guide the design of experimental studies to most productively address areas of greatest uncertainty.

2. Ecological Research

ORD has made a significant commitment to genomic and proteomic research on model aquatic vertebrates. For example, ongoing work in this area has focused on the use of genomics and proteomics to delineate toxicity pathways associated with disruption of key elements of the thyroid system in the amphibian model, *Xenopus tropicalis*. Other work in this area is focused on small fish models, including the sheepshead minnow, medaka, and the fathead minnow. A primary goal of this research is elucidation of toxicity pathways induced by perturbation of processes controlled by the hypothalamic-pituitary-gonadal/thyroid (HPG/T) axes. The fathead minnow, *Pimephales promelas*, is an important model for this research because (1) it is the Agency standard for teleost aquatic toxicity testing; (2) the fathead minnow 21-day reproductive toxicity test is the priority Agency-recognized screen for endocrine disruptors in teleosts; (3) significant progress has been made in gene discovery and initial cDNA microarray synthesis via core microarray facilities; (4) high-throughput sequencing of cDNA libraries by the Department of Energy (DOE) is likely; and (5) a multi-laboratory effort is enabling integration of exposure and effects research. Ultimately, identification of initiating events or other critical pathway elements will enable the development of QSAR models, initially for estrogens and androgens in fish, and subsequently for other targets on the HPG/T axes. As research focused on toxicity pathways associated with these axes develops, classes of chemicals that operate via other modes/mechanisms of action also will be considered. An important group of chemicals currently under consideration in this regard is the polyfluorinated surfactants.

Besides the primary objective of elucidating toxicity pathways, this collaborative research aims to significantly support other goals of the Computational Toxicology Research Program. For example, the discovery of genes specifically induced/repressed by hormone agonists and antagonists will enable the development of specific diagnostic molecular indicators of exposure and the linkage of PB/PK models to mechanistically-based QSAR models (see section I-C). In addition, genomic, proteomic, and metabonomic research will provide the basis for comparison of toxicity pathways across species. For example, the present collaborations within ORD have already produced highly significant data on the conservation and function of the estrogen receptor across genera.

C. External Linkages

ORD is actively seeking to establish links with groups outside the Agency that have expertise and capabilities that could complement and augment the emerging program on computational toxicology. Some of these linkages have already been initiated, and they are listed below. Other partnerships, such as with other governments or private consortiums, will likely develop over time as the program matures and specific research needs in the intramural program are identified.

1. Chemical Industry Institute for Toxicology (CIIT) Centers for Health Research

ORD and CIIT have agreed to a memorandum of understanding (MOU) to advance the state-of-the-science of computational toxicology. Among the goals of the agreement is the utilization of the complementary capabilities and expertise of ORD and CIIT in developing and applying computational toxicology approaches to human health risk assessment. Research activities common to both organizations focus on the use of computational methods and molecular biology toward the characterization of risks of environmental contaminants. CIIT focuses on the development of PBPK and PD while ORD contributes to the joint effort by studying relevant toxicant-induced effects in various target organs (i.e., research on toxicity pathways). Improving the risk assessment process has been an emphasis of both groups for many years, and approaches involved in computational toxicology are compatible with the goals of both organizations. This collaboration will foster exchanges of information, training opportunities, and technologies. Initial research focuses on examining the toxicity pathways for the effects of dibutyl phthalate in the developing testes.

2. Department of Energy (DOE)

In the fall of 2002, ORD scientists visited the Sandia and Pacific Northwest National Laboratories of DOE. These visits led to the development of 23 possible project areas thought to be suitable for collaboration. These were subsequently narrowed down to five projects worthy of further exploration. The high priority projects of mutual interest to DOE and the Agency include a range of areas such as the following: (1) research to develop computational screening techniques beginning with traditional QSAR screening of chemicals to predict toxicity of EDCs; (2) the development of new techniques to specifically identify molecular structures associated with toxicity; and (3) the use of innovative and proteomic techniques to identify, characterize, and classify sensitive subpopulations based on biological factors to reduce uncertainty in risk assessment. Additionally, both DOE and the Agency are interested in the characterization of atmospheric pollutants to better understand the processes influencing human risks to atmospheric pollutants, to identify and characterize the casual agents associated with these effects, and to better understand the systems biology and mechanisms associated with the exposure-to-dose-to-effects parts of the continuum. An MOU and an Interagency Agreement have also been developed between DOE and ORD to provide high performance computing consulting and/or access to DOE non-classified computing equipment.

In January 2003, ORD scientists also visited the U.S. DOE Joint Genome Institute (JGI), which resulted in JGI's agreement to sequence the genome of selected species currently being used for testing and screening of environmental pollutants (i.e., *Pimephales promelas*, *Xenopus tropicalis* and *Daphnia pulex*). Sequencing the genome of these test species could lead to molecular-based models for predictive toxicology. Another visit to DOE laboratories was made in December 2003.

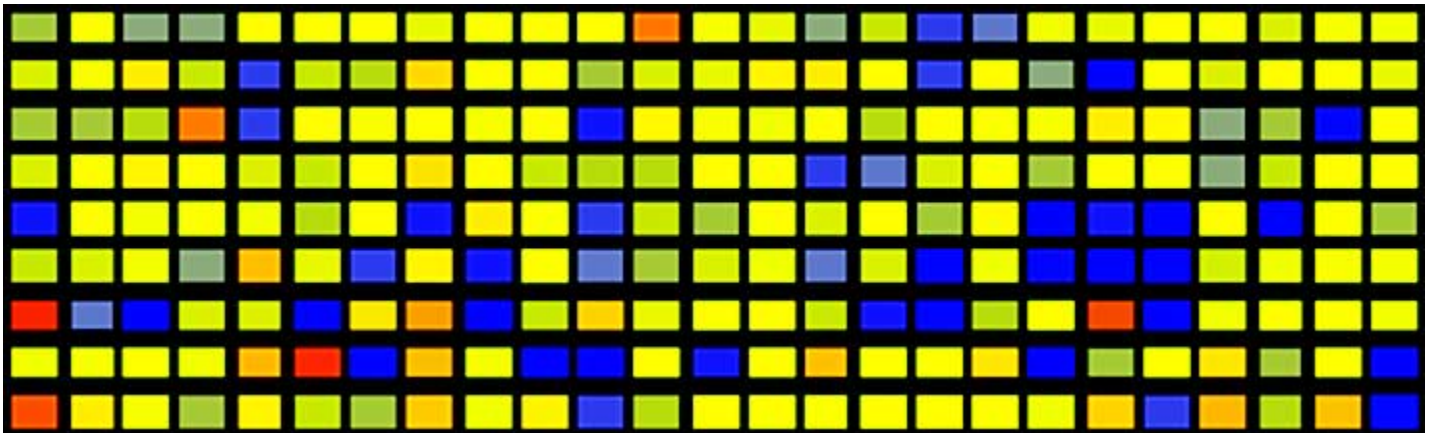
Other collaborations with genomics and bioinformatics experts at the University of Cincinnati Children's Hospital Research Center have begun to yield synthesis of complementary DNA microarrays with fathead minnow gene sequences and expressed sequence tags. These researchers represent the conduits for synthesis of large-scale microarrays using sequences gleaned from the JGI collaborations.

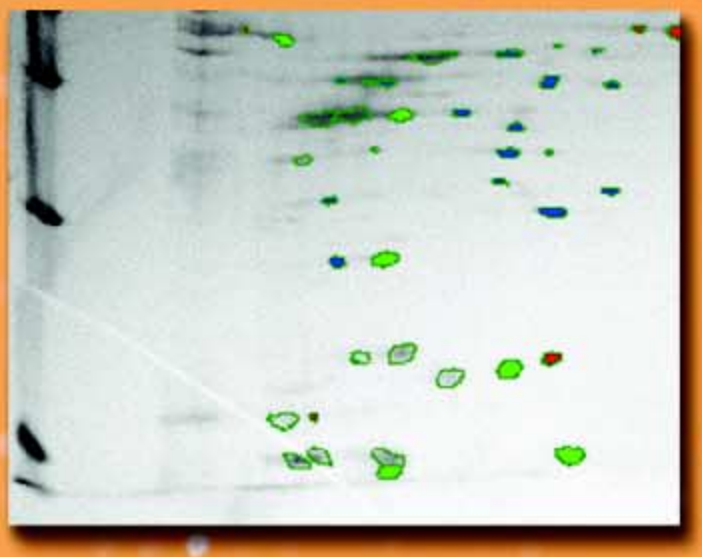
3. National Institute of Environmental Health Sciences (NIEHS)

The National Center for Toxicogenomics (NCT) at NIEHS was formed to facilitate development of gene expression and proteomic methodology, to create a public database relating environmental stressors to biological responses, to collect information relating environmental stressors to biological responses, to develop improved use of computational mathematics in understanding responses to environmental stressors, and to identify biomarkers of disease or exposure to enhance environmental health. Many of these objectives overlap with current core and problem-driven research at ORD. Preliminary discussion has been held with NCT concerning the Computational Toxicology Research Program at ORD. One possibility for collaboration is that the genomic, proteomic, and metabonomic information from ORD's Computational Toxicology Research Program will be made available to the NCT's Chemical Effects in Biological Systems (CEBS) database. CEBS is a relational and descriptive compendia of toxicologically important genes, groups of genes, and mutants and their functional phenotypes. This platform allows searching for information about the biological effects of chemicals and other agents and their toxicity pathways based on information from the literature and contributions from intramural and extramural sources. The database utilizes standardized procedures, protocols, data formats, and assessment methods to ensure that data meet a uniform level of quality. It is expected that ORD will help build a publicly accessible toxicological database that will be capable of predictive toxicology and that will serve as a major resource for researchers to pose and test mechanistic hypotheses.

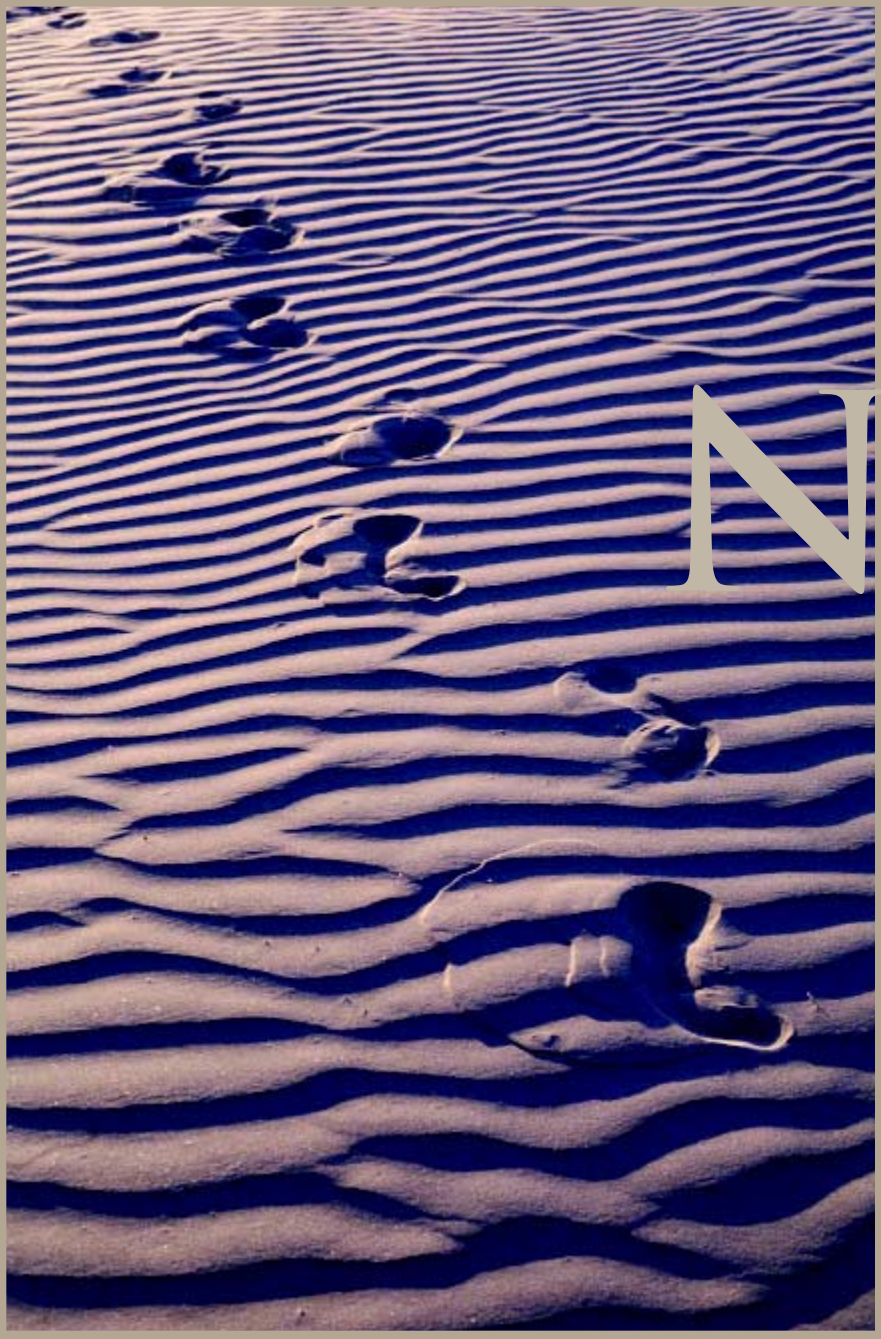
4. Science to Achieve Results (STAR)

Through its external grants program, ORD is developing a series of requests for assistance that will lead to the support of scientists in academic and not-for-profit institutions in research areas that complement the in-house research activities. Most of the support will be in the form of grants; but in some cases where ORD scientists would like to work more closely with the awardees, cooperative agreements may be awarded. The selection of both awarded grants and cooperative agreements will be through a competitive process. All awardees will be asked to participate in periodic progress reviews in which intramural and extramural scientists are brought together to share and review data.





$$\mathrm{Var}\,U \approx \left(\frac{\partial \gamma(p_1, t)}{\partial p_1}\right)^2$$



NEXT

Next Steps

IV. Next Steps

A. Review of the Framework

The next step in the development of a Computational Toxicology Research Program in ORD was to obtain external peer review of the *Framework for a Computational Toxicology Research Program in ORD.* On September 12, 2003, the *Framework* was reviewed by the Computational Toxicology Framework Consultation Panel of the Agency's Science Advisory Board. The current document includes several revisions based on the results of that consultation. In addition, ORD hosted a workshop on September 29-30, 2003, to discuss how ORD will use the data and experience from other research organizations in the development of a research program. A summary of the results of that workshop can be found on the Agency's computational toxicology Web site (http://www.epa.gov/comptox). In FY04, ORD will issue Requests for Proposals to identify themes and intramural and extramural research approaches in the area of computational toxicology based on the priorities and process described in the following two sections. Eventually, an ORD Multi-Year Plan on Computational Toxicology will be developed to describe the critical milestones that this area of research will accomplish over the next five to eight years.

B. Priorities for Research on Computational Toxicology in ORD

The following criteria will be used to set priorities for research:

1. Risk-Based Planning: Research should address an element in the source-to-outcome continuum, and it should be designed to improve quantitative risk assessment or to facilitate the development of screening or testing strategies for chemicals.
2. Utilizes New Technology: Research that employs emerging proteomic, genomic, and metabonomic methods.
3. Hypothesis-Driven: Research that tests hypotheses.
4. Scientific Excellence: The quality of the planned science must be able to withstand rigorous peer review.
5. Programmatic Relevance: Research that addresses a key Agency-related mandate concerning protection of human health and/or the environment.
6. Other Sources of Data: Research that involves partnerships and collaborations with other organizations outside of the Agency.
7. Capabilities and Capacities: Research that can be completed within a reasonable period of time using available material and human resources.
8. Sequence of Research: The critical path of the research dictates those efforts which must be started before other elements are initiated. This again refers to the relative state of maturity of different technologies included in the Framework.
9. Balance: There should be a balance of short-term versus long-term results.

C. Process for a Research Program on Computational Toxicology

The development of an ORD program on computational toxicology will require coordination and communication between ORD managers and investigators to ensure that the research is relevant, timely, and defensible. Engagement of personnel from Agency regulatory offices will be an important component of the implementation process in order to ensure that the activities are targeted towards where they can have the greatest impact. Efforts will be made to optimize existing intramural research by seeking out complementary extramural research efforts. A Web site has been established (www.epa.gov/comptox) that will serve as a repository of information and progress in the program. This will serve as an important tool for communication both within and outside the Agency.

To be successful, it is apparent that increased emphasis on bioinformatics and database management on a scale not generally available within ORD is critically needed. In this regard, ORD is collaborating with the Office of Environmental Information (OEI) to develop a plan for upgrading the Agency's information technology (IT) capabilities and making those resources available to Agency scientists and risk assessors. Through this plan, the Agency envisions establishing a state-of-the-science IT framework and architecture network that would liken the Agency's computing capabilities and applications to those of other federal agencies (e.g., DOE, Centers for Disease Control and Prevention) and non-governmental organizations (e.g., the University of Chicago Statistical Center). This enhanced IT infrastructure will enable Agency scientists to utilize sophisticated computational techniques to develop virtual biological systems (e.g., simulated liver metabolism models, virtual neuroendocrine system) and models to predict toxicological effects (e.g., QSAR approaches for endocrine disruptors).

Following final release of the *Framework*, the technical writing team that authored this document has been replaced by the ORD Computational Toxicology Implementation Steering Committee. The mandate for this group is to provide oversight of the process to implement a research program on computational toxicology in ORD. This includes the development of an ORD Implementation Plan for computational toxicology to guide and coordinate research in this area for the next five to eight years. Scientist-to-scientist meetings involving intramural and extramural researchers will also be sponsored by the Steering Committee to help integrate research with intramural and extramural groups working in this area. A first step in this process occurred on September 29-30, 2003, at a workshop entitled "Computational Toxicology: Framework, Partnerships and Program Development" at which nearly 200 participants heard detailed descriptions of the *Framework* and began discussions to help identify research areas for future emphasis. These discussions centered around two approaches: needs based on an existing regulatory driver and needs that would be most compatible with the application of emerging computational tools and approaches. The purpose of these discussions was to take maximum advantage of emerging technologies that would be of immediate relevance to Agency risk assessors and risk managers. Additional workshops are being planned that will provide a greater depth of discussion for various topics in the *Framework*, e.g., use of computational approaches for susceptible subpopulations, models addressing cumulative risk, or approaches for prioritization of chemicals for Agency data requirements. Periodic review of on-going research by the Steering Committee will be crucial to demonstrating progress and designing future research directions. Systematic communication with the Agency's Program and Regional Offices through the planning process will be important for providing them with an understanding of the nature and extent to which the applications of computational toxicological approaches meet the Agency's data requirement needs and improve quantitative risk assessment.

APPENDIX A

Examples of Current ORD Projects Associated with Computational Toxicology	
A. Improve Linkages in the Source-to-Outcome Continuum	
	1. Chemical Transformation and Metabolism Fate models-Core research to elucidate and model the behavior of organic contaminants in natural and impacted ecosystems and in complex biological systems (NERL) Fate models-Problem-driven research to develop the methods, tools, and databases to forecast the fate of pesticides and toxic chemicals during the drinking water process (NERL)
	2. Development of Diagnostic/Prognostic Molecular Indicators Developmental Biomarkers-core research that focuses on identifying molecular markers of developmental toxicity related to growth and maturation of organ systems (NHEERL) Virulence Potential-core research to develop molecular methods to measure virulence of microbial pathogens (NERL)
	3. Dose Metrics Cumulative Risk for Drinking Water Contaminants-problem-driven research to use QSAR, mixtures toxicity approaches and PBPK modeling to develop a cumulative risk method based on doses in target tissues (NCEA) PBPK Models in Fish-core research to use genomic data to validate output of PBPK models (NERL)
	4. Characterization of Toxicity Pathways Cell Signaling-core research to determine role of signal transduction pathways in toxicity pathways for high priority environmental chemicals (NHEERL) Framework for Defining Model and Mechanism of Action for Cancer and Noncancer Endpoints-research to explore using genomics and proteomics as an approach for understanding mechanisms and the implications for risk assessment (NCEA)
	5. Systems Biology None
B. Provide Predictive Models for Screening and Testing	
	1. QSAR Approaches Application of QSAR and modeled exposure estimates in risk-based chemical ranking model (NCEA) Perfluorooctane sulfonate (PFOS)-problem-driven research on mode or mechanism of action of PFOS, a breakdown product of several widespread and persistent chemicals in the environment (NHEERL)
	2. Pollution Prevention Strategies Endocrine Disrupting Chemicals Replacement Program - problem-driven research to develop a software tool that will allow users to quickly identify possible replacements for chemicals that are known or projected to have endocrine disrupting potential (NRMRL) Pollution Prevention Tools - core-driven research to incorporate results of other computational toxicology projects into software tools used in applying pollution prevention strategies NRMRL)
	3. High Throughput Screening Endocrine Disruptors-problem-driven research designed to identify endocrine-mediated effects using rapid high throughput protein fingerprinting techniques (NHEERL)

C. Enhance Quantitative Risk Assessment	
Examples of Current ORD Projects Associated with Computational Toxicology	
	1. Dose-Response Assessment Acute-to-Chronic Estimate-problem-driven research to develop regression and accelerated life testing models to predict long-term chronic toxicity from short-term acute responses and determine uncertainty at low concentrations (NHEERL) Develop and Apply Unified Modeling Procedure to Cancer and Noncancer Risk Assessment to support biologically based dose response modeling (NCEA)
	2. Cross Species Extrapolation Cross-species Extrapolation in Birds-problem-driven research to develop PBPK models to extrapolate reproductive and neurological effects of metals among bird species (NHEERL) Value-of-Information Approach to Motivate Uncertainty Factors with Mechanistic Data: Chlorine Human Health Risk Case Study: Experimental and computational efforts aimed at obtaining and integrating mechanism of action data to develop a biologically based risk assessment for chlorine. The results will be generalized to develop a formal framework for departing from default uncertainty factors based on PK/PD data (NCEA)
	3. Chemical Mixtures Chemical Mixtures-core research to apply genomic analyses of exposure of fathead minnows to binary and tertiary chemical mixtures (NERL) Interactions and Mechanism of Pesticide Mixtures: PBPK/BBDR modeling for immunotoxicity risk assessment of chemical mixtures (NCEA)

United States
Environmental Protection
Agency

Office of Research
and Development (8101R)
Washington, DC 20460

Official Business
Penalty for Private Use
$300

EPA 600/R-03/065
November 2003
www.epa.gov

Please make all necessary changes on the below label, detach or copy, and return to the address in the upper left-hand corner.

If you do not wish to receive these reports CHECK HERE ☐ ; detach, or copy this cover, and return to the address in the upper left-hand corner.

PRESORTED STANDARD
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Recycled/Recyclable
Printed with vegetable-based ink on paper that contains a minimum of 50% post-consumer fiber content processed chlorine free