ORD's

Computational Toxicology

Research Program

Implementation Plan

(FY 2006 – 2008)

April 2006

Table of Contents

Table of Contents ii

List of Figures iii

Acronyms iv

Executive Summary v

I. Introduction and Background 1

II. ORD Computational Toxicology Program 3

NCCT's Mission 4

STAR Centers 5

BOSC 6

III. Research Projects 6

IV. Summary of Outcomes from Research 9

V. Partnerships 9

VI. Conclusion 11

Table 1 -Identification of CTRP Projects 13

Table 2 -CTRP Output/Input Table 14

Table 3 -FTE's from ORD's Supporting MYPs 22

Appendix A 23

Appendix B 24

.

List of Figures

Figure 1 – Component of the Computational Toxicology Research Program within ORD 4

Figure 2 - NCCT Staffing Profile 5

Figure 3 -Alignment of the Five Research Track 7

Figure 4 -Interactions between LTGs 8

Figure 5 -Computational Toxicology Research Program Design..... 12

.

Acronyms

BOSC	Board of Science Counselors
CPCP	Categorization and Prioritization Community of Practice
CoP	Communities of Practice
CompTox	Computational Toxicology
CTISC	Computational Toxicology Implementation and Steering Committee
CTRP	Computational Toxicology Research Program
DNA	Deoxyribonucleic acid
DSSTox	Distributed Structure-Searchable Toxicity
EBRC	Environmental Bioinformatics Research Center
EPA	Environmental Protection Agency
FTE	Full Time Equivalents
FY	Fiscal Year
LTG	Long Term Goal
MOA	Memorandum of Understanding
MYP	Multi Year Plan
NCCT	National Center for Computational Toxicology
NCEA	National Center for Environmental Assessment
NCER	National Center for Environmental Research
NERL	National Exposure Research Laboratory
NHEERL	National Health and Environmental Effects Research Laboratory
NIH	National Institutes of Health
NIEHS	National Institute of Environmental Health Sciences
NRMRL	National Risk Management Research Laboratory
NICEATM	National Toxicology Program Interagency Center for Evaluation of Alternative Toxicological Methods
OPPTS	Office of Pesticides, Prevention, and Toxic Substances
ORD	Office of Research and Development
QSAR	Quantitative Structure Activity Relationship
RFA	Request for Application
RNA	Ribonucleic acid
SAB	Science Advisory Board
STAR	Science to Achieve Results
ToxCast	Chemical Prioritization Research Program

Executive Summary

This document presents the implementation plan for the Framework for the Computational Toxicology Research Program (CTRP), a strategy document developed by the Office of Research and Development (ORD) in 2003. Computational toxicology is defined as the integration of modern computing and information technology with molecular biology and chemistry to improve risk assessment and prioritization of data requirements of chemicals by the Agency. The program is intended to provide innovative solutions to a number of persistent and pervasive issues facing The Environmental Protection Agency (EPA) regulatory programs. The three objectives of the Framework have been translated into long term goals (LTGs) for the CTRP and the subsequent research has been aligned into five supporting tracks. The LTGs for the program are: (I) EPA risk assessors use improved methods and tools to better understand and describe linkages across the source to outcome paradigm; (II) EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation; and (III) EPA risk assessors and regulators use new models based on the latest science to reduce uncertainties in dose-response assessment, cross-species extrapolation, and quantitative risk assessment. The supporting research tracks are: (A) Development of Data for Advanced Biological Models; (B) Information Technologies Development and Application; (C) Prioritization Method Development and Application; (D) Providing Tools and System Models for Extrapolation across Dose, Life Stage, and Species; and (E) Advanced Computational Toxicology Approaches to Improve Cumulative Risk Predictions. A standing subcommittee of ORD's Board of Scientific Counselors has been established to provide guidance to the CTRP as it develops its research agenda.

The research supporting the CTRP flows from three areas. The first component is composed of the efforts of the National Center for Computational Toxicology (NCCT), an organizational entity created within ORD in 2005 to accelerate the development and use of computational tools in the EPA's regulatory operations. The staffing plan for the NCCT includes computational chemists, computational biologists and bioinformaticians who will work both on internal projects and through partnerships with scientists in other ORD Labs and Centers as well as external scientists to provide scientific expertise and leadership related to the application of mathematical and computational tools and models to high priority Agency needs. The projects located within the NCCT include those developing information databases for chemical toxicity (DSSTox), building toolboxes (ToxCast) for prioritizing chemicals for toxicology evaluation, providing computation models that extrapolate exposure and effects across life stages, developing multi-scale computational models of organ systems that will help provide understanding of how mechanisms of toxicity determine dose and time-responses, evaluating issues related to estimation of model parameters in computational models, and integrating diverse types of information needed to understand the impact of the environment in a broad sense on the health of individuals.

The second component is a suite of seven projects distributed across ORD that were funded in FY05 by a competitive process overseen by the Computational Toxicology Implementation and Steering Committee (CTISC), a cross EPA group formed to initiate the ORD program. These projects include efforts that bring systems level approaches to the study of chemical toxicity in small fish models and in amphibian metamorphosis, to the effects of diesel

particles on lung cells, and to factors associated with development of asthma in children. Other projects are designed to predict metabolites of chemicals, to use metagenomic approaches to assess microbial pollution sources, and to evaluate the use of toxicogenomic data in risk assessment of reproductive toxicants.

The final component of the CTRP is derived from the Science to Achieve Results (STAR) program of ORD's National Center for Environmental Research (NCER). The STAR program supported a grant solicitation for investigator directed research in this area of systems level understanding of the hypothalamic-pituitary-gonadal axis in FY04 that funded two projects involving small fish models and one on the female rat. In FY05, the STAR program supported the establishment of two Centers for Environmental Bioinformatics, one at the University of North Carolina in Chapel Hill and the other at the University of Medicine and Dentistry of New Jersey. These Centers were funded as cooperative agreements, which allows for close interactions between scientists within the University-based Centers and EPA scientists. With project periods of five years, they are intended to provide considerable momentum in the development and application of bioinformatics tools in the protection of human health and the environment.

Within this plan, the research issue and relevance, experimental approach, progress to date and milestones over the next three years are articulated for each of the individual projects. Collectively the elements of the plan have been developed to provide a sequential series of short to medium term projects that will advance the utilization of computational tools in hazard and risk assessment. Particular progress is expected over the next few years in the area of prioritization tools, as the ToxCast program begins developing fingerprints of chemical activity based on the collection of broad spectrum, high throughput data for a large number of well characterized chemicals. Projects developing systems level understanding of biological functions will require longer investments, but the results in terms of improving the linkages in the source-to-outcome paradigm and subsequently in the extrapolation of information across dose, time, chemical and species will reduce some of the inherent uncertainties present in risk assessment as currently practiced.

I. Introduction and Background

This document lays out the rationale and short to medium term objectives (up to 3 years) of the ORD new research program in computational toxicology. The emerging field of computational toxicology applies mathematical and computer models and molecular biological and chemical approaches to explore both qualitative and quantitative relationships between sources of environmental pollutant exposure and adverse health outcomes. Recent technological advances make it possible to develop molecular profiles using genomic, proteomic, and metabolomic (the "omics") methods to identify the impacts chemicals may have on living organisms or the environment. With these tools, scientists can produce a more-detailed understanding of the hazards and risks of a much larger number of chemicals. The integration of modern computing with molecular biology and chemistry will allow scientists to better prioritize data, inform decision makers on chemical risk assessments, and understand a chemical's progression from the environment to the target tissue within an organism and ultimately to the key steps that trigger an adverse health effect.

"Computational toxicology: integration of modern computing and information technology with molecular biology"

Currently, risk estimates are most often based on gross outcomes of disease such as occurrence of cancer, a neurological disorder, or a visible birth defect. It has long been assumed that these disease outcomes were the result of numerous and crucial alterations at the molecular level inside living cells. Molecules of DNA, RNA, or endogenous proteins all have crucial chemical and physical arrangements and interactions. Alterations in these arrangements and/or interactions might be the first step in a cascade that can then lead to disease, morbidity, and mortality. Research in genomics, proteomics, metabolomics, and computational toxicology are expected to result in risk assessments based on specific changes at the molecular level, rather than just the number of tumors, deaths, overt clinical changes observed in test animals. In the future, assessments will be based on changes in molecular markers, such as the number of DNA molecules altered at a crucial site, the change in an allosteric membrane protein that acts as a receptor, or the change in a regulating protein inside the cell. The key is we will better understand how those changes lead directly to the types of adverse health effects that have been the traditional basis of EPA risk assessments and to use this understanding to reduce the uncertainties in the extrapolation of effects across dose, species and chemicals.

Over the past 15 years the explosive growth of information regarding the structure of many of the "molecules of life" has occurred. The sequencing of the human genome for example, has given us a wealth of information only dreamed of in the previous decades. It will soon be possible to accurately ascertain exactly where a xenobiotic chemical interacts with regions of the genome and with other endogenous molecules. Binding of a toxic chemical with a crucial portion of membrane proteins for example, may alter the structure and porosity of the cell membrane. Such a change might lead to a change in the membrane potential of the cell. If that cell is a neuron or cardiac cell critical physiologic changes might occur that could lead to overt disease.

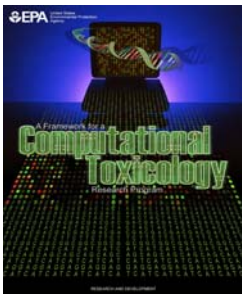
Advancements in the "omics" technologies coupled with the advances in analytic tools such as microarray techniques will enable us to predict changes and evaluate which changes can initiate, promote, or adjoin the cascade to disease. Computational systems biology techniques

will help us integrate the necessary quantitative information related to those changes. The concepts of computational chemistry and relating structure to activity are not new. For years chemists and physicists have been relating key chemical and physical properties to molecular and atomic structures. However, with advances in high performance computing and the application of high throughput biological screening assays, there is unprecedented opportunity to conduct this research on a much larger scale of operation.

From the sciences of "omics" technologies arise two possibilities: the first is the broad spectrum (or high content) interrogation of multiple molecular events within a cell or tissue (e.g. microarrays) and the second is the application of one or a few molecular targets in a high throughput screening mode against large numbers of chemicals. Both of these possibilities present opportunities and challenges. The opportunity is to use the ensuing outputs of these technologies to shift toxicology from a descriptive to a predictive science and therefore improve the ability of the EPA to assess hazard and characterize risk. The challenge lies in reducing vast amounts of data generated by these approaches into useful information and in determining their biological significance. The EPA's ORD has formulated its research program with this in mind and will approach the issues through several targeted research tracks. The main thrust of the research program will be to develop generic approaches that are adaptable to the more specific types of risk assessment activities that are undertaken in the other Laboratories, Centers, or Program Offices.

This challenge was given urgency in 2002, when Congress ordered a redirection of $4 million from available EPA funds, "*for the research, development and validation of non-animal alternative chemical screening and prioritization methods, such as rapid, non-animal screens and Quantitative Structure Activity Relationships (QSAR), for potential inclusion in EPA's current and future relevant chemical evaluation programs*". To fulfill this directive, the EPA embarked on development of a research program that: (1) was consistent with the Congressional mandate; (2) complemented and leveraged related on-going Agency sponsored efforts to consider alternative test methods; (3) further advanced the research to support the Agency's mission; and (4) would not duplicate the mission and programs in this area conducted by other agencies. An innovative program, entitled Computational Toxicology (CompTox), was initiated to target these goals and, in the process, significantly advance toxicology and risk assessment as currently practiced by the Agency and the broader environmental sciences community. As recommended by Congress, the proposed approach was developed in consultation with the Office of Pesticides, Prevention, and Toxic Substances (OPPTS). The research projects funded in FY02 and FY03 were largely devoted to proof-of-concept demonstrations that the approaches of computational toxicology could be adapted to the study of endocrine disruptors (chemicals which perturb the functioning of an endocrine system and which subsequently lead to adverse health effects in an individual or a population). Early successes of these efforts included refinement of estrogen receptor ligand binding data for use in development of quantitative structure-activity models, evaluation of two EPA-developed cells lines for the detection of activity of estrogens and androgens, and the development of an alternative test method for evaluating effects on steroidogenesis that avoids the use of animal tissues. As the current document is intended to be primarily forward-looking, these early projects will not be discussed further in this document (see the 2004 Activity Report for more details).

With increasing attention to and expectations for the CompTox program over the next several years, ORD developed *A Framework for a Computational Toxicology Research Program* in 2003, which provided strategic direction for the program. This document was the product of a cross-ORD team of scientists, and was endorsed by the Science Advisory Board (SAB). The Framework identified three objectives for computational toxicology in the EPA and these have been translated into three Long Term Goals (LTGs) for the research program described herein:



- I. Risk assessors will use improved methods and tools to better understand and describe the linkages of the source-to-outcome paradigm,

- II. EPA program offices will use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation and

- III. Assessors and regulators will use new and improved methods and models based on the latest science for enhanced dose-response assessment and quantitative risk assessment.

II. ORD Computational Toxicology Program

With issuance of the Framework, ORD began the process of implementation of a research program by establishing in FY04 the Computational Toxicology Implementation and Steering Committee (CTISC) to help guide the program and to communicate progress across the Agency. Membership on the committee consists of the chair, the Director of the NCCT, two representatives from each of ORD Laboratories and Centers, which are nominated by their Directors, appropriate ORD management representatives (e.g., Associate Laboratory Directors or National Program Managers from aligned programs), an ORD Regional Scientist, and representatives of the main client offices of the CTRP Program. Membership consists of a three year term, renewable once, (link to current membership). The recognition of the potential impact of computational approaches to toxicology has continued to grow over the past several years, and now research contributions flow from three distinct areas within ORD. The first component is composed of the efforts of the National Center for Computational Toxicology (NCCT), an organizational entity created within ORD in 2005 (see below) to accelerate the development and use of computational tools in the EPA's regulatory operations. The second component consists of projects funded in FY05 by a competitive process overseen by the CTISC. The final component of the CTRP is derived from the STAR program of ORD' NCER. Within the CTRP, the STAR program supports a small number of investigator lead grants and two larger Centers. The relationship of these components is depicted in Figure 1. The innermost component (grey oval) represents the efforts taking place within the NCCT. The location in the inner circle indicates the central role of the NCCT in providing leadership to the overall program. In the next outer circle (green oval) are the CTISC funded projects, which represent research being performed in other Laboratories and Centers. These efforts are largely within the National Health and Environmental Effects Research Laboratory (NHEERL) and the National Exposure Research Laboratory (NERL). Smaller efforts are contained within the National Risk Management Research Laboratory (NRMRL) and the National Center for Environmental Assessment

(NCEA). The third and final component of the CTRP is research supported by the EPA's NCER through the STAR program, funded through Request for Applications (RFAs). A particularly important effort of the STAR program was the funding of two academic centers to support the advancement of bioinformatics in environment health (see below). Greater detail on the projects in each of the three components is provided in Section III. It is important to recognize that considerable efforts in computational research and in "omic" research reside outside those directly or indirectly associated with the NCCT and CTISC. These efforts, represented by the yellow oval in Figure 1, are part of ORD's overall research in health, ecological, and risk assessment, and are contained within a variety of Multi-Year Plans (MYPs).

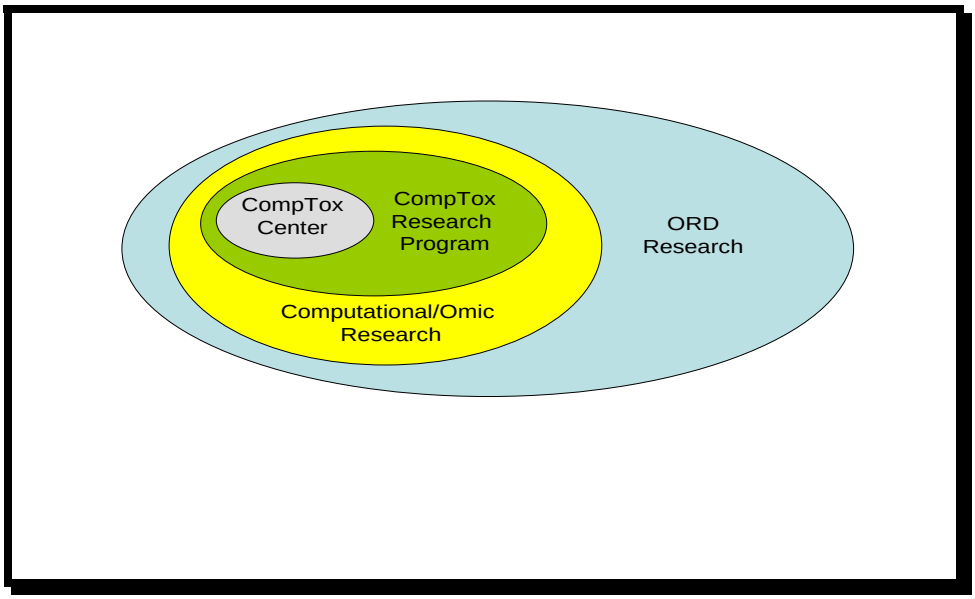


Figure 1 – Component of the Computational Toxicology Research Program within ORD

The National Center for Computational Toxicology: In October 2004, the EPA Science Advisor and Assistant Administrator for ORD, Dr. Paul Gilman announced the formation of the NCCT, which began official functions in February 2005. The announcement states:

> "The Center will advance the science needed to more quickly and efficiently evaluate the potential risk of chemicals to human health and the environment. The center will coordinate and implement EPA's research on computational toxicology to provide tools to conduct more rapid risk assessments and improve the identification of chemicals for testing that may be of greatest risk."

NCCT's Mission

The Center's mission is to achieve the goals set forth by ORD and the Framework document by performing research that integrates modern computing and information technology with molecular biology to improve Agency prioritization of data requirements and risk assessment of chemicals. The Center's staffing profile is designed to provide expertise in systems biology, computational chemistry, and bioinformatics, as outlined in Figure 2. Recruitments are currently underway for two senior level investigators, one in the area of computational systems biology and the other in bioinformatics. These positions are expected to

be filled by the end of FY06 and with supporting positions will complete staffing of the NCCT, as currently envisioned.

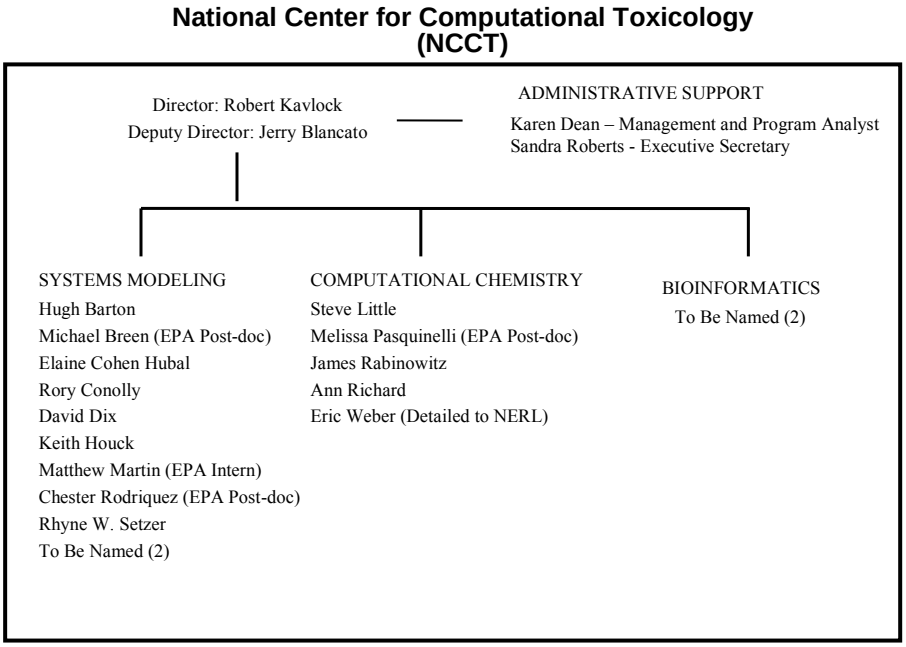


Figure 2 - NCCT Staffing Profile

The NCCT scientists will advance the field of computational toxicology by working across ORD to develop predictive models for screening and testing chemicals and by reducing the uncertainties associated with extrapolating predicted effects across chemicals, levels of biological organization and species. They will provide scientific expertise and leadership related to the application of mathematical and computational tools and models, conduct research to improve the predictive capabilities of the methods, models and measurements that constitute the input materials to the computational approaches. In addition, they will conduct and sponsor research to provide models for fate and transport of chemicals, environmental exposures to humans and wildlife, delivery of the chemical to the target site of toxicity, molecular and cellular pathways of toxicity, and ultimately systems-level understanding of biological processes and their perturbation. A key part of the process of advancing the science will involve developing partnerships with other government and private organizations so as to best leverage resources committed to the effort.

STAR Centers

STAR Environmental Bioinformatics Centers: One area of research need, implicit in many of the research projects contained with the CTRP, is bioinformatics. This rapidly emerging technology is crucial to the computational toxicology program, and whereas the NCCT will be adding a senior level bioinformaticist in FY06, there remains a large gap in ORD relative to the ability to analyze the high volumes of molecular data and to predict potential toxicity, modes of action, and ultimately risk. To help bridge this gap, NCER has supported the establishment of two STAR Environmental Bioinformatics Research Centers (EBRC). The Research Center for

Environmental Bioinformatics and Computational Toxicology at the University of Medicine & Dentistry of New Jersey (UMUDJ), Piscataway, NJ and The Carolina Environmental Bioinformatics Research Center at the University of North Carolina, Chapel Hill, NC will operate as cooperative agreements and help facilitate the application of bioinformatics tools and approaches to environmental health issues supported by the CTRP. It is anticipated that over the next 5 years the EBRCs will conduct research to improve the science of bioinformatics and assist in the analysis and interpretation of data relevant to the protection of human health and the environment. Descriptions and web links of the two Centers are provided in Appendix A Bioinformatics staff of the NCCT will serve as the technical liaisons with the EBRCs and help guide the interactions with other ORD investigators.

The STAR program is planning another Request for Applications in the area of computational toxicology in FY07, and the topic for that call is currently under discussion.

BOSC

Board of Scientific Counselors: To help guide the program, ORD established a standing panel of Board of Science Counselors (BOSC) to provide review and advice to the NCCT and CTRP. The panel first met in April 2005 to review the organization of the NCCT, the initial plans for implementation, and the progress of the early CTRP work. The panel commented very favorably on the early progress of the Center and the means outlined to achieve its goals (BOSC CompTox Review). The composition of the staff, plans for future hiring, the establishment of working partnerships, and the Center's strategic plan were especially highlighted. The panel made several main recommendations for consideration, two of which are addressed here. The first was to develop a formal implementation plan for the future. The second was to develop Communities of Practices (CoPs) within the EPA that can serve as a networking function for interested scientists. Three such CoPs have been organized and are discussed later in this document. A few other minor suggestions were also made, which were addressed in the formal ORD response to the review (ORD BOSC Response). A second site visit of the BOSC is being planned for June 19–20, 2006, which will entail an in-depth assessment of the progress NCCT has made in executing this implementation plan.

III. Research Projects

As noted above, research projects within the CTRP are composed of projects developed internally within the NCCT, the efforts funded across ORD by the CTISC, and the research supported by ORD's STAR program (in the form of both individual research grants as well as larger Center grants). Research efforts supporting the three LTGs identified in the CompTox Framework have been grouped into five research tracks: (A) Development of Data for Advanced Biological Models; (B) Information Technologies Development and Application; (C) Prioritization Method Development and Application; (D) Providing Tools and System Models for Extrapolation across Dose, Life Stage, and Species; and (E) Advanced Computational Toxicology Approaches to Improve Cumulative Risk Predictions. The alignments of the five research tracks with the LTGs of the Framework are depicted in Figure 3. The component research projects are identified in Table 1, with hyperlinks to both summary information on their individual outputs and impacts (Table 2), as well to more descriptive information on the research issue and relevance, approach, impact and partnerships in Appendix B. For those projects

housed outside the NCCT, Table 3 provides a summary of the supporting information from other MYPs.

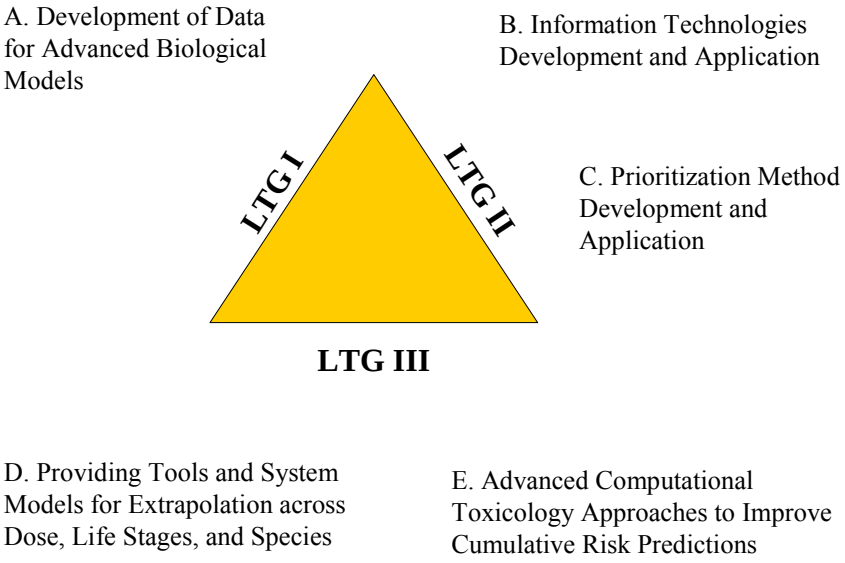


Figure 3 -Alignment of the Five Research Track

Track A contains seven projects that are targeted at collecting data to improve various aspects of the linkages within the source-to-outcome paradigm. Included is a project characterizing thyroid toxicity in an amphibian model, three projects applying "omic" technologies to the study of endocrine disruptors in small fish, one project examining gene expression networks in the rodent uterus exposed to estrogens, one project characterizing the properties of diesel particles that are toxic to human pulmonary cells, and one project seeking environmental influences on childhood asthma. These projects are distinguished in that they have a large data collection effort combined with a computational modeling component. Track B contains efforts to improve the management of information, both from the chemo-informatics and bio-informatics prospective, to provide the basis for development of better structure-activity models and to support interpretation of high throughput data being derived from Track C. Importantly, Track B is designed to migrate the maximum amount of information into the public domain so it can be used by others for a variety of purposes. Track C contains four projects intended to speed the development of data for hazard characterization, and includes components on predicting xenobiotic metabolism, molecular docking models for in-silico predictions of ligand-receptor interactions, high throughput data acquisition to fingerprint chemicals for potential hazard, and development of technologies to identify sources of fecal contaminants in aquatic environments. Track D contains seven projects that are primarily computational in nature, with efforts ranging from statistical considerations for model parameter fitting, to providing generic tools for cross species and cross life stage extrapolations, exploring model portability and linkage issues via use of software standards, modeling particular cell and organ processes, applying "omic" data in risk assessment, and developing tools and platforms for use of "omics" in regulatory decision processes. Finally, Track D contains two projects that are

exploring computational approaches to real-world type situations where exposure is to more than one form of stressor.

The outputs of the various projects within each LTG will also provide support for future activities in other LTGs through relationships depicted in Figure 4.

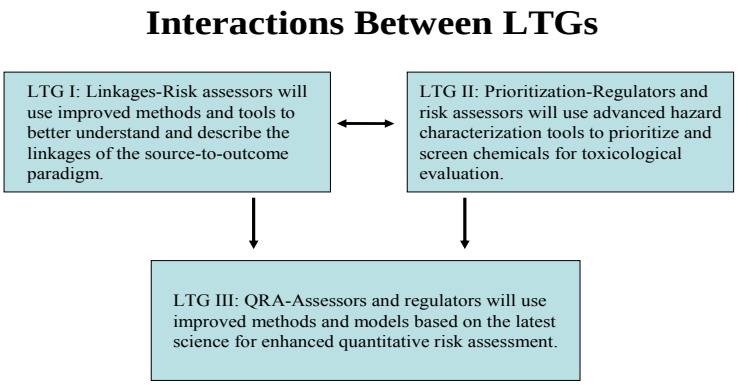


Figure 4 -Interactions between LTGs

For example, in projects under LTG II, involving the use of advanced hazard characterization tools by EPA Program Offices, early efforts are being directed towards providing new high throughput techniques for prioritization and for identifying those environmental stressors that should undergo further testing (see IIC-3). However, information gleaned from such high throughput technologies are expected to prove informative for advancing modeling efforts (e.g. Projects IIID-5). Preliminary information about mode-of-action can be extracted from these studies and the models can be exercised using a variety of scenarios derived from these studies. The models can help identify mode-of-action scenarios that would contribute most to potential hazard, and risk models could then target on those pathways. Additional molecular studies could then be conducted, concentrating on those scenarios identified as having the greatest potential impact. This next phase of information would be used to further refine the models and reduce uncertainty. Such iterative approaches between controlled experiments and advanced modeling techniques, while very useful, have proven to be very expensive and time consuming. The cost and required time could be reduced using some of the approaches developed under this LTG. The systems biology project (IIID-5) will serve as a culmination project whereby information from a variety of sources (high throughput, in-vitro, in-vivo, and in-silico) is used to formulate a risk assessment. In this project, models can be used to estimate the probability of adverse effects and the variability and uncertainty associated with those estimates from the available knowledge and data. The models can also be used as testing tools to estimate how much uncertainty could be reduced with more experimental data and also to identify the type of needed data. Such linkage between the LTGs will help advance the overall program. As noted, LTG II will make significant contributions to LTG III and will also feed information to LTG I. Similarly, as toxicity pathways are characterized at multiple levels of biological organization in LTG I, screening tools can be added to projects in LTG II that will enable larger numbers of chemicals to be characterized. Finally, furthering understanding of toxicity pathways

from projects in LTG I will advance computational work in LTG III, by examining how various toxicity pathways interact with one another to modulate activity.

IV. Summary of Outcomes from Research

The implementation plan described herein was developed to provide a sequential series of short to mid term projects that will advance utilization of computational tools in hazard and risk assessment. In particular, over the next 1 to 3 years we expect to make significant progress in developing and applying prioritization and screening tools. For example, there will be targeted application of our current understanding of selected toxic endpoints so that predictive models can be developed leading to more efficient testing paradigms and reduction in uncertainties in inter-species extrapolation. These models will be populated with data at the molecular and biochemical molecular levels and will provide a basis to interpret interspecies homology and comparative toxicity. Proof-of-concept of ToxCast, in particular, will provide a number of EPA program offices with an extremely useful tool to improve the efficiency and effectiveness of hazard identification and risk assessment methodologies. New and innovative ways will be developed to assimilate, evaluate, and use the myriad of data assorted with molecular and chemical information. The further development and application of DSSTox will be a key component of these efforts. Computational chemistry will provide in-silico models for predicting complex interactions of environmental chemicals with important biochemical receptors, which can then lead to adverse effects. Early examples of modeled target interactions include estrogenic, androgenic receptors, and acetyl cholinesterase.

Progress will be made at developing computational models and modeling systems that represent comprehensive descriptions of the underlying biology of adverse impacts caused by exposure to environmental agents. Computational models describing the relationship between diesel exhaust particle composition and its genotoxic and inflammogenic properties are examples. Other projects will develop methods to measure and then describe the mechanisms of childhood asthma resulting from environmental exposures. The whole-systems biology modeling approach will develop a range of models, from those describing pharmacodynamic connections between exposure and effects to those describing complex endogenous pathways and the perturbations in such pathways resulting from environmental exposures. Also, ways to incorporate and use “omics” information in these models will be explored. Finally, attempts will be made at formulating models of common, but complex, disease processes which are then exacerbated by exposures to exogenous substances and stressors. A detailed summary of outputs and impacts resulting from the projects listed in the previous Table 1 can be found in Table 2.

V. Partnerships

Given the broad nature of the challenges facing computational toxicology, the CTRP must engage collaborative partners both across ORD and also with outside organizations in order to be successful. Internally, development of these linkages began at the management level with the initiation of a Memorandum of Understanding (MOA) with the NHEERL and the NERL. This MOA provides the NCCT with access to administrative support functions such as oversight of extramural actions, formal quality assurance procedures, and planning for information management. Access to these and other functions allow for more efficient and effective use of FTEs within the Center. Management staff of the NCCT has a regularly scheduled monthly

meeting with counterparts in NERL and NHEERL to ensure smooth operations and to strategize on common scientific directions.

With the successful initiation of a cross-ORD research program supporting the framework and the creation of the NCCT in February 2005, the charter of the CTISC was redefined to include the following functions:

- Serving in an oversight function over the computational toxicology projects funded by the activities of the committee and coordinate, as appropriate, with relevant MYPs associated with the projects.
- Establishing partnerships external to the EPA to help support the Framework as implemented in the various Laboratories and Centers of ORD.
- Advising NCER on the formulation of ideas for new RFAs, providing suggestions for scientific peer reviewers, serving on relevancy reviews as appropriate, and keeping informed of the research progress of the STAR grants program.
- Acting in a consultancy role for activities of the NCCT.
- Providing a forum for communication of computational toxicology activities across ORD and the Agency.

The ORD's multi-year planning process provides another opportunity for linkage between the CTRP and related research efforts. Each of the MYPs is led by a National Program Director with support from staff of the relevant Laboratories and Centers. The NCCT participates actively in four of the MYP teams. Three contain similar research activities for screening and prioritizing chemicals, i.e. Endocrine Disrupting Chemicals (EDCs, LTG III), Safe Pesticides/Safe Products (SP2, LTG I), and Drinking Water Research Program (DW, LTG II) whereas the fourth has a major focus on the incorporation of biologically based mode-of-action information into quantitative risk assessment (the Human Health Research Strategy and the Human Health MYP, LTG II). The Director of the NCCT meets at least quarterly with the National Program Directors for these MYPs, and the NCCT is attempting to allocate at least 10% of its available extramural resources to computational toxicology-related projects in other Laboratories and Centers. Preference in deciding which projects to support will be given to those efforts that have the active participation of a member of the NCCT.

On the scientific level, the NCCT has initiated three Communities of Practice (CoP) in the areas of Chemoinformatics, Biological Modeling, and Categorization and Prioritization that are intended to unite practitioners in the designated fields. Prior to implementation of these CoPs, the concept of 'adjunct' appointments with researchers outside the Center to create a collaborative environment was considered, but there was no clear idea of how such appointments would function. The concept of the CoPs was suggested by the BOSC in April 2005, and has since been adopted as a primary means of communication and integration of activities across ORD, the EPA, and outside entities. These efforts will serve to enhance communication and coordination, develop common standards, promote consistency, evaluate and provide guidance on best practices, recommend research priorities, and provide training to interested parties. Copies of the charters for these groups are provided here (Chemo; Modeling; and CPCP).

Extending further outside the confines of ORD and the EPA, the NCCT is also developing strong linkages with the National Toxicology Program of the NIEHS. The

commonalities between *A Framework for Computational Toxicology* and the A National Toxicology Program for the 21st Century: A Roadmap for the Future and the co-location on the RTP federal research campus provided a strong impetus for these interactions. Beginning in mid 2005, the two groups began an exchange of work-in-progress seminars that led to establishment of a number of specific events, including efforts that draw on the complementary strengths of the two organizations. These include a joint seminar program, collaborating on statistical analysis of complex dose-response patterns, sharing of workgroup and CoP memberships, developing a mathematical construct for model-building based on information from varying levels of biological information, meta analysis of PBPK model development across chemicals, participating in a class study on perfluornated compounds, and the offering of technical training courses to parties outside the organizations. Working relationships are also being developed with the National Toxicology Programs Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM), which is currently exploring high throughput hazard identification tools. These efforts extend to opportunities for collaboration and linkages with NIH through the NIH Roadmap, *Accelerating Medical Discoveries to Improve Health* are also being sought. This especially concerns the theme, "New Pathways to Discovery: Molecular Libraries and Molecular Imaging and Building Blocks Biological Pathways and Networks." Discussions with Dr. Chris Austin, Senior Advisor to Director of Translational Research of the National Human Genome Research Institute, have led to consideration for inclusion of the EPA-relevant chemicals and biological assays within the Molecular Libraries Initiative.

VI. Conclusion

Ultimately the CTRP and the NCCT will provide outcomes that will help reduce or prevent the risk to humans and the environment from environmental stressors. As such, the program is aligned with the overarching goals of the Agency. Figure 5 depicts the design of the program for achieving these goals. Shown is the progression of the research program over the next several years. We list the resources and inputs that will help define the program and the activities that are discussed in this plan. Next, we highlight expected research outputs, means for their transfer, and example clients and beneficiaries of these outputs. Finally, summarize the intent or expected outcomes and how those outcomes will improve risk assessment and management procedures and policies. This will clearly take several years to come to fruition. Hence this plan, while outlines a multifaceted process, has concentrated mostly on the first years where the activities are commenced and specific outputs are produced. It is hoped that the outputs will help achieve short-term outcomes and will greatly contribute to the outcomes expected in later years.

This plan should be considered a "living document" that will be updated and modified on a yearly basis. These annual updates will reflect and document progress, any necessary change in direction, and added objectives as the program matures.

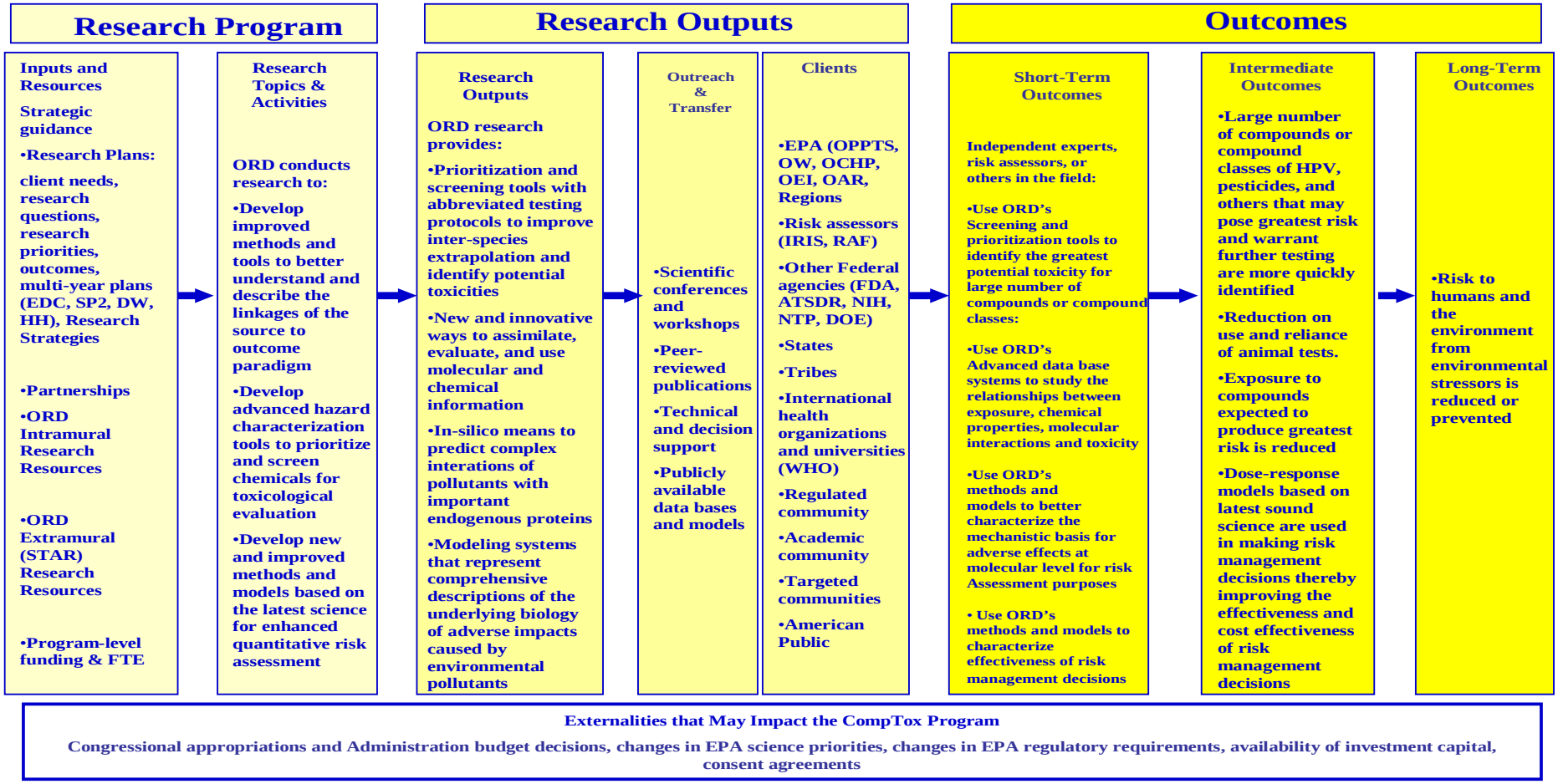


Figure 5 -Computational Toxicology Research Program Design

Table 1- **Organization of Projects in the Computational Toxicology Program** *(The Roman numeral indicates the Long Term Goal (LTG), the A-E letter indicates the Research Track and the Arabic numeral identifies the project number. Projects shown in* ***blue*** *are projects largely external to the NCCT; projects in* ***red*** *are primarily conducted within the NCCT; and projects in* ***green*** *are part of the STAR Program. Note: the STAR Environmental Bioinformatics Centers are not listed here, as they are cross cutting activities).*		
Long Term Goal	**Research Track**	**Project Title**
I. EPA risk assessors use improved methods and tools to better understand and describe linkages across the source-to-outcome paradigm	A. Development of Data for Advanced Biological Models	IA-1 Linkage of Exposure and Effects Using Genomics, Proteomics, and Metabonomics in Small Fish Models IA-2 Systems Biology Modeling of Fathead Minnow Response to Endocrine Disruptors IA-3 Chemically Induced Changes in Gene Expression Patterns along the HPG Axis at Different Organization Levels Using a Small Animal Model (Medaka) IA-4 A Systems Approach to Characterizing and Predicting Thyroid Toxicity Using an Amphibian Model IA-5 Estrogen Elicited Gene Expression Network Elucidation in the Rat Uterus IA-6 Risk Assessment of the inflammogenic and mutagenic effects of diesel exhaust particles: A systems biology approach IA-7 Mechanistic Indicators of Childhood Asthma (MICA)
II. EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation	B. Information Technologies Development and Application	IIB-1 Integrated Chemical Information Technologies Applied to Toxicology
	C. Prioritization Method Development and Application	IIC-1 Simulating Metabolism of Xenobiotic Chemicals as a Predictor of Toxicity IIC-2 Modeling Molecular Targets for Toxicity, a Computational Approach to Understanding Key Steps in the Mechanisms for Toxicity and a Tool for Prioritizing Bioassay Requirements IIC-3 ToxCast, a Tool for Categorization and Prioritization of Chemical Hazard Based on Multi-Dimensional Information Domains IIC-4 Development of microbial metagenomic markers for environmental monitoring and risk assessment
III. EPA assessors and regulators will use new models based on the latest science to reduce uncertainties in dose-response assessment, cross-species extrapolation, and quantitative risk assessment.	D. Providing Tools and System Models for Extrapolation Across Dose, Life Stage and Species	IIID-1 Statistical Methodology for Estimating Parameters in PBPK/PD Models IIID-2 Modeling Toxicokinetics for Cross-Species Extrapolation of Developmental Effects IIID-3 Development of a portable software language for physiologically-based pharmacokinetic (PBPK) models IIID-4 Systems Modeling of Prostate Regulation and Response to Antiandrogen IIID-5 Systems Biology Model Development and Application. IIID-6 Use of Toxicogenomics Data in Risk Assessment: Case Study for a Chemical in the Androgen-Mediated Male Reproductive Development Toxicity Pathway IIID-7 Developing Computational Tools for Application of Toxicogenomics to Environmental Regulations and Risk Assessment
	E. Advanced Computational Toxicology Approaches to Improve Cumulative Risk Predictions	IIIE-1 Dose-Time-Response Modeling for Evaluating Cumulative Risk of N-Methyl Carbamate Pesticides IIIE-2 Application of Visual Analytic Tools to Evaluate Complex Relationships between Environmental Factors and Health Outcomes

Table 1 -Identification of CTRP Projects

Table 2 -CTRP Output/Input Table I. EPA risk assessors use improved methods and tools to better understand and describe linkages across the source to outcome paradigm				
A. Development of Data for Advanced Biological Models				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IA-1 Linkage of Exposure and Effects Using Genomics, Proteomics, and Metabonomics in Small Fish Models	Development of a conceptual model for the HPG axis in small fish models as a basis for focused hypothesis testing of potential endocrine disrupting chemicals.	Preliminary results on the effects of model chemicals on the fecundity of fathead minnows and single gene and protein expression.	Gene and protein expression as the basis for extrapolation of the effects of endocrine disrupting chemicals across small fish models.	OPPTS - Development and validation of screening test methods for chemicals impacts the HPG in humans and wildlife species and for the incorporation of mechanism-specific data into OPP's probabilistic risk assessments.
IA-2 Systems Biology Modeling of Fathead Minnow Response to Endocrine Disruptors	To determine and compare gene and protein expression profiles and physiological and reproductive endpoints for adult FHM exposed to a model estrogen 17 alpha-ethinylestradiol (EE2), androgen (17β-trenbolone), or their antagonists (ZM 189,154 and flutamide, respectively).	To predict gene expression patterns of two compounds (zearalenone and EE2) that are environmental estrogens.	To develop a computational modeling framework that integrates exposure concentration, gene expression, and proteomic profiles with physiological endpoints.	This STAR Grant is expected to develop a computational model and identify 10-15 molecular and protein biomarkers that are specific and predictive of adverse effects of exposure to estrogenic compounds in reproduction of fathead minnows. This quantitative model will help improve risk assessment of exposure of wildlife and by extrapolation, of mammals to endocrine disrupting compounds.
IA-3 Chemically Induced Changes in Gene Expression Patterns along the HPG axis at Different Organization Levels Using a Small Animal Model (Medaka)	Determine natural backgrounds and variability of gene expressions, optimize methods based on findings, develop methods to identify gene products and begin exposure studies with model chemicals.	Finalize exposure studies and identification of genomic expression patterns based on four model compounds and perform statistical comparisons. Quantify gene products and determine their functionality and biological relevance. Compare responses of test chemicals to model chemicals. Submit final report on project.		A model will be developed to predict the biological relevance of the observed changes in gene expression profiles By identifying the systemic target sites, and the series of biological events from gene expression to the manifestation of an adverse outcome (e.g., reproductive performance), thresholds at the molecular level that are indicative of effects on the fitness of the individual, including survival, growth and reproduction (fertility and fecundity as well as survival of the offspring) will be determined. Understanding the potential of individual chemicals and complex environmental mixtures to interfere with molecular pathways of concern will enhance our understanding of the basic mechanisms of toxicities, and thus, will have the potential to develop better focused, more rapid and cost effective models for quantitative risk assessment. The bioassay can be used to screen for a wider range of endocrine disruptor effects.

I. EPA risk assessors use improved methods and tools to better understand and describe linkages across the source to outcome paradigm *(cont.)*				
A. Development of Data for Advanced Biological Models *(cont.)*				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IA-4 A Systems Approach to Characterizing and Predicting Thyroid Toxicity Using an Amphibian Model	Establish data management system; Formulate initial systems model for thyroid function; Develop in-vitro thyroid and pituitary culture systems.	Characterize molecular changes associated with TH inhibition; Characterize specific molecular responses in thyroid gland culture; Develop initial QSAR for NIS inhibition.	Refine systems model for relating MOA to outcome; Complete first round QSAR hypothesis for testing for NIS.	Development of a sufficient understanding of the HPT so that predictive models can be developed, testing protocols can be abbreviated, and efforts in inter-species extrapolation can be improved. Populate these models with data specifically at the molecular and biochemical levels which will provide a basis to interpret interspecies homology and comparative toxicity.
IA-5 Estrogen Elicited Gene Expression Network Elucidation in the Rat Uterus	Establish estrogenic endocrine elicited dose- and time-dependent changes in rat uterine gene expression using four estrogen receptor ligands; investigate role of ER in mediating changes in gene expression.	Phenotypically anchor changes in gene expression to histopathological outcomes. Develop a computational model that describes the estrogen elicited gene expression network; Submit final project report.		The models developed will identify gene expression changes most highly associated with EED elicited histopathological uterine responses. Examination of multiple environmental endocrine disruptors with varying potencies will also identify key regulatory nodes responsible for eliciting these responses, which could lead to the development of high throughput endocrine disruptor screening assays for chemicals in commerce. The data and resulting models can also be integrated with other algorithms (i.e. PBPK) to create a more comprehensive model of the hypothalamic-pituitary-gonadal axis.
IA-6 Risk Assessment of the Inflamagenic and Mutagenic Effects of Diesel Exhaust Particles: A systems Biology Approach	Generate the first four DEP samples; Conduct chemical characterization of the DEP samples; Conduct a pilot study assessing the inflammogenicity and signal transduction profile of two DEP samples in human and rodent airway epithelial cells.	Generate eight DEP samples; Identify overlaps in human and mouse gene expression patterns associated with DEP inflammogenicity or mutagenicity; Perform a bioassay-directed fractionation on the available DEPs and determine the distribution of mass and mutagenicity among the fractions.	Generate remaining four DEP samples; Define signal transduction pathways involved in inflammogenic and mutagenic responses to DEP exposure based on gene expression patterns; Complete biological models of relevant signaling pathways.	Predictive models that quantitatively describe the relationship between diesel exhaust particle composition and its genotoxic and inflammogenic properties: Well developed and scientifically based risk assessments for diesel exhaust particles including the use of genomics and proteomics in the risk assessment NOTE: for 4th year (09): Complete a database of all biological measurements and analytic results, and the transfer of this data to modelers and risk assessors for the preparation of manuscripts; Develop exposure models based on ambient exposure data, geographic information (locations of roadways and major sources of PAHs and metals), and home/school locations; Compare exposure model estimates with biomarkers of exposure and early effects.

I. EPA risk assessors use improved methods and tools to better understand and describe linkages across the source to outcome paradigm *(cont.)*				
A. Development of Date for Advanced Biological Models *(cont.)*				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
1A-7_Mechanistic Indicators of Childhood Asthma (MICA)	Collect bloods and isolate RNA from rodent lung and blood samples and blood; Develop QA plan for gene expression analysis; Develop and seek approval for Intramural Research Protocol including recommended and standard operating protocols for collecting pilot study gene expression analysis data.	Develop relevant educational modules appropriate for schoolchildren and their parents to enhance participant recruitment and ensure their informed consent; Develop asthma severity score criteria based on asthma diagnosis, medicine and FEV1 in school-based questionnaires and lung function measurements for selection of schoolchildren; Develop standard operating procedures for the collection, processing, and analysis of biological samples from children; Complete monitoring of ambient exposure levels indoor/outdoor.	Develop plan for collection and analysis of biological samples; Collect, prepare process and analyze human biological samples from 200 children; Prepare comprehensive data base NOTE: for 4th year (09): Complete a data base of all biological measurements and analytic results and the transfer of this data to modelers and risk assessors for the preparation of manuscripts; Develop exposure models based on ambient exposure data, geographic information (locations of roadways and major sources of PAHs and metals), and home /school locations; Compare exposure model estimates with biomarkers of exposure and early effects.	Research would be associated with Annual Performance Goal (APG) #6: By 2012, provide risk assessors and mgrs with methods and tools for measuring exposure and predicting effects in children, including adolescents, characterizing cancer and non-cancer hazards and risk to children, and reducing risks to children in schools from harmful environmental agents; Enhancement of quantitative risk assessment, produce better methods and predictive models for quantitative risk assessment and to provide a useful tool for large-scale biomonitoring in humans; Better determination of the shape of the exposure-response curve, especially in the low-exposure region, through the incorporation of gene expression data into experimental systems; The development of more accurate, biologically-based mathematical exposure-response models that predict responses outside the range of experimental values; The identification of regulatory metabolic or physiologic pathways, that may act in concert and lead to adverse health outcomes, through the evaluation of multiple health end points with linkages to gene expression changes.

II. EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation				
B. Information Technologies Development and Application				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IIB-1 Integrated Chemical Information Technologies Applied to Toxicants	05: Incorporate DSSTox standard chemical fields into CEBS; Create DSSTox database and documentation files for NTP Immunotox database; create chemical index file for the EPA IRIS programs; Establish Communities of Practice - Chemoinformatics Workgroup to begin to coordinate efforts across the Agency to inventory, retrieve, and explore chemical information data.; Propose plan to link DSSTox effort with the NLM PubChem project. 06: Create DSSTox standard chemical field index files for public genomics databases and NTP legacy toxicity data; coordinate linkages within the CEBS relational search environment; Create and publish DSSTox database and documentation files for additional published toxicity databases and chemical index files for EPA and NTP programs; Propose plan to structure-index and quality review chemical information in EPA data files currently on the web and to provide for an EPA-wide structure browser; Assist with formation of proposed toxicity chemicals subset for high-throughput testing in ToxCast and the NIH Molecular Libraries Screening Initiative collaboration.	Assist with incorporation chemical structure browser technology into EPA ArrayTrack, and full with DSSTox data files and structure-indexed public genomics data; Establish procedures and protocols for automating the chemical annotation of new data submitted to CEBS or EPA ArrayTrack from the DSSTox Master chemical list; Continue expansion of the DSSTox public toxicity database inventory; Implement plan to uniformly structure-index EPA data files and provide EPA website structure-searchability through curated EPA chemical data files; Create structure-annotated database of test results for toxicity subset chemicals from the NTP/NIH Molecular Libraries Initiative.	Advise and assist with the implementation of chemical definitions (e.g., assigning a chemical to a class) and structure analog searching capability across chemically indexed data files, integrated with toxicogenomics data and bioinformatics capabilities, to serve as the foundation for chemoinformatics capabilities across EPA and NTP data; Advise and assist with the development of procedures and capabilities for deriving chemical signatures for predicting toxicity outcomes from the complete profile of ToxCast data; Begin to tabulate and explore data from NTP/NIH Molecular Libraries screening collaboration in relation to other DSSTox databases; Implement procedures for expanding the structure-annotation of EPA chemical data records and providing methods for flexible structure or analog searching on the EPA website.	Improve capabilities to access, mine, and integrate useful chemical-biological activity information from existing and new data, both within and outside EPA. These efforts have the potential to impact a wide variety of EPA program offices that heavily rely on chemical information resources, such as the High-Production Volume Testing Program, the Premanufacture-Notification Program in OPPTS, ORD's IRIS Program, and the Office of Pesticide Programs.

II. EPA Program Offices use advanced hazard characterization tools to prioritize and screen chemicals for toxicological evaluation *(cont.)*				
C. Prioritization Method Development and Application				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IIC-1 Simulating Metabolism of Xenobiotic Chemicals as a Predictor of Toxicity	For priority chemicals, incorporate available chemical metadata, metabolism data, and metabolic maps into a searchable database for data management and structure/substructure searchable access; Forecast metabolic pathways for selected priority chemicals using existing simulator for liver metabolism; Initiate in-vitro Phase I liver microsome experiments; incorporate new laboratory and literature metabolism data into simulator training sets.	Confirm formation of predicted metabolites for priority chemicals and compare observed maps to forecasted maps; Develop and approach to evaluate and enhance simulator performance through improvement of transformation probability estimates and expansion of transformation reaction domain; evaluation of refined simulator and model predictions in context of Program Office prioritization needs.		Provide effects-based prioritization of chemicals (parent compound and metabolites) relevant to EPA Program Offices: Provide OPPT and OPP the ability to prioritize chemical lists (based upon predicted toxic effects of parent chemical and metabolites) with reliability estimates for use in chemical evaluations and to rank chemicals for in-vitro or in-vivo screening and toxicity testing; provide capability to OPP and OPPT for predicting bioactive metabolites; develop searchable metabolism database for OPP and OPPT use for identification of relevant chemical and/or substructures of interest for risk assessment; provide linkage of effects based toxicity model with metabolic simulation.
IIC-2 Modeling Molecular Targets for Toxicity: a Computational Approach to Understanding Key Steps in the Mechanisms for Toxicity and a Tool for Prioritizing Bioassay Requirements	Results of the use of the target-toxicant paradigm to screen for estrogenicity/androgenicity; Results of the affect of the two binding site model on the cumulative risk of chemicals acting through the enzyme AChE.	Results of the importance of protein flexibility in evaluating the interaction of chemicals with macromolecular receptors.	Evaluation of the use of the target-toxicant method as a tool in a diverse chemical screen; Application of the target-toxicant approach to other targets of toxicity.	Help fulfill the Agency need for predictive models for hazard identification, both the sub areas of QSAR and other computational approaches and High Throughput Screening.
IIC-3 ToxCast: A Tool for Categorization and Prioritization of Chemical Hazard Base on Multi-Dimensional Information Domains	Develop conceptual framework for ToxCast (started in 05); establish initial battery of assays across the information domains, identify list of chemicals to evaluate proof of concept for framework and begin data acquisition.	Report on the utility of statistical clustering techniques on assay results from pilot chemicals to group them according to known toxicity patterns; revise framework as dictated by results.		A biologically and chemically based system to begin to associate chemicals of like properties and activities will provide a number of EPA Program Offices with an extremely useful tool that addresses the mission of improving the efficiency and effectiveness of hazard identification and risk assessment methodologies employed by the EPA.
IIC-4 Development of Microbial Metagenomic Markers for Environmental Monitoring and Risk	Evaluation of metagenomic databases as a source of molecular markers to assess human and animal fecal contamination in surface waters.	Report on the evaluation of PCR-based host-specific assays to confirm the presence of animal fecal sources of pollution in waters impacted by fecal contamination; Report on fecal indicator microorganisms and/or genetic markers from fecal material whose densities in recreational waters best correlate with the rates of illnesses in users of recreational waters.		Development of assays that can be used in source identification, environmental monitoring, and risk assessment and to distinguish between human and nonhuman fecal contamination in Nation's waterways and supplies.

III. EPA risk assessors and regulators use models based on the best science to reduce uncertainties in the dose-response and cross-species extrapolations.				
D. Providing Tools and System Models for Extrapolation Across Dose, Life Stage, and Species				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IIID-1 Statistical Methodology for Estimating Parameters in PBPK/PD Models	Submission of a journal article outlining outstanding statistical issues in the analysis of PBPK models and providing methods for use of formal statistical methods for more reliable and rational estimates of key parameters and for model evaluation.	Joint NIEHS/NCCT workshop on Statistical Issues in PBPK/PD modeling to establish expert based consensus on best practices in statistical analysis and evaluation of PBPK/PD models.	Publication of framework for formal statistical analysis of PBPK/PD models.	Better characterization of methods for estimating parameters and quantifying uncertainties in pharmacokinetic and pharmacodynamic models predictions will remove one major impediment to their more general application. This will allow replacement of default uncertainty factors with transparent mechanism-based statements of scale and uncertainty, in turn decreasing the subjectivity and increasing the transparency of environmental health risk assessments, impacting several program offices including OPP, OPTS, and OW.
IIID-2 Modeling Toxicokinetics for Cross Species Extrapolation of Developmental Effects	Model rat pup lactational exposure for perfluooctanoate (PFOA) using compartmental approaches.	Develop initial physiologically-based pharmacokinetic model for PFOA in rat maternal-fetal-pup unit and identify data gaps in relation to rat and human.	Evaluate modeled dosimetry for rat fetus and pup for a limited number of prototype compounds to inform the uncertainty in use of maternal exposure dose in risk assessments.	Characterization of internal dosimetry will inform the uncertainties attendant to analyses based upon the maternal exposure dose. In the presence of information on the critical window, the models may directly form the basis of the quantitative risk assessment by deriving the relevant internal dose metric for extrapolation to humans. Perfluorinated compounds, including PFOA, are an important proof of concept for this research because they are a class of compounds producing developmental effects of significant regulatory concern to OPPTS and others. These compounds are currently under evaluation by the Agency.
IIID-3 Development of Portable Software Language for PBPK Models	Report on findings of assessing extensiveness required for the XML schema.	NCCT-sponsored meeting on evaluation of the performance of the completed XML schema and status of the visualization program.	Report on proposal to have finalized XML schema included in the next release of SBML and evaluation of visualization program.	Enhance the ability of formulating and investigating quantitative dose-response relationships using mode and mechanism of action knowledge and data: The ability to create, transfer, use, augment, and review PBPK models without the limitation of software compatibility are a long-standing desire in the PBPK community. The ability to seamlessly link PBPK models to biological pathway models is of great interest, since this gives the modeler a fast and efficient means of extending the estimations of tissue dose from the PBPK model to cellular-level responses.

III. EPA risk assessors and regulators use models based on the best science to reduce uncertainties in the dose-response and cross-species extrapolations. *(cont.)*				
D. Providing Tools and System Models for Extrapolation across Dose, Life Stage, and Species *(cont.)*				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IIID-4 Systems Modeling of Prostate Regulation and Response to Androgens	Prostrate function model following castration.	Dose-response relationship with testosterone and antiandrogen exposure.	Biologically based model of prostate androgen dependent gene regulation incorporating genomics data.	Development and demonstration of a biologically based model of endogenous endpoints and their pertubationscaused by exposure to environmental factors. A true pharmacodynamic model that can be used by OPPTS and others in mechanistically based risk assessments.
IIID_5 Systems Biology Model Development and Application	Journal article on use of biologic models to ascertain necessary resolution of exposure measurements; Collaborative groups formed; Formulation of conceptual model and writing the mathematics and code for testing models; Selection and begin model implementation for endogenous biochemical system; Start model development of disease process.	Implementation of model and begin application for cases with exposures to known environmental toxicants – abstracts and presentations; Disease model coded, exercised, and evaluated; Determine through literature search and other means for examples of exogenous exposure that impact the disease process.	Journal articles illustrating some uses of "omics" information in quantitative models; Summary report for Agency use on earliest best practices on "omics" and system biology models; Journal article for disease model; Enhancement of disease model to incorporate exposures to environmental toxicants.	The latest mechanistic information and the interplay between exposure, endogenous factors, pre-existing conditions, and genetic predisposition can be rationally accounted for by using such models. There is a great need for quickly illustrating how the latest molecular information, especially "omic" information and information coming from high through-put studies will be used in risk assessments by NCEA and several program offices, including OPPTS, OW, OAR.
IIID-6 Use of Toxicogenomic Data in Risk Assessment: Case Study for a Chemical in the Androgen-Mediated Male Reproductive Development Pathway	Draft of the case study report (includes scoping exercise complete for case studies; i.e., progress on case study and defining approach for integrating TG data into risk assessment); Discussions with the EPA chemical assessment team, Regions, and Program Offices.	Conduct EPA Colloquium presenting results and lessons learned from the case study; External peer review draft of report; Submit manuscript on project to peer-reviewed journal.		This project was developed in response to a recommendation from the NCEA sponsored Genomics and Risk Assessment Colloquium of 2003. Specifically, it will conduct a case study to provide a practical attempt to incorporate currently available toxicogenomics data that would illuminate issues and the methods development. The results will help EPA prepare for genomics data availability and submission by addressing areas of risk assessment where such data may be particularly useful; analyzing acceptance criteria for inclusion of toxicogenomics data in risk assessment; and incorporating approaches for the use of toxicogenomics in risk assessment.
IID-7 Developing Computational Tools for Application of Toxicogenomics to Environmental Regulations and Risk Assessment	Completion of participation in the Microarray Quality Control (MAQC) with FDA and publication of papers describing best practices- includes description of SPC Genomics Technical Framework. Installation of FDA ArrayTrack database for ORD and Agency use.	Continued development of ArrayTrack database and analytical tools for toxicogenomics in cooperation across Agency, with FDA, and with the NC and NJ Environmental Bioinformatics Centers. Trial incorporations of toxicogenomic data into risk assessments in collaboration with various Program and Regional Office staff.	Publication of examples and principles for integrating toxicogenomic data into risk assessments in peer-reviewed scientific journals and contribution of these principles into Agency science policy.	Development of these toxicogenomic databases and tools, and application of these various toxicogenomic data within Program and Regional Offices will provide EPA staff with valuable, practical training in genomics and associated disciplines. As toxicogenomics grows more important to environmental science and policy, such activities will help EPA develop the computational tools and methods to properly evaluate genomics information.

III. EPA risk assessors and regulators use models based on the best science to reduce uncertainties in the dose-response and cross-species extrapolations. *(cont.)*				
E. Advanced Computational Toxicology Approaches to Improve Cumulative Risk Predictions				
Project Number & Title	**Outputs/Outcomes 06**	**Outputs/Outcomes 07**	**Outputs/Outcomes 08**	**Expected Impacts**
IIIE-1 Dose-Time-Response Modeling for Evaluating Cumulative Risk of N-methyl carbamates	Science Advisory Panel review (August 05) of preliminary cumulative risk assessment, including dose-response modeling; release and Science Advisory Panel review of revised cumulative risk assessment; submission of methods and dose-response models for publication in the peer-reviewed literature.	Incorporation of dose-time-response methodology into the Agency's BMDS software.		The dose-response results of this work will be used by OPP in the dose-response assessment portion of the Agency's N-methyl carbamate cumulative risk assessment. Methods developed in this analysis could be used in the future for cumulative risk assessments for agents with ephemeral acute effects.
IIIE-2 Application of visual analytic Tools to Evaluate Complex Relationships between Environmental Factors and Health Outcomes	Demonstration of potential for application of VA to evaluate exposure data and workshop on applying VA to analyze children's cohort data.	Generic conceptual model of complex relationships between environmental factors and human health outcomes.	Demonstration of application of VA to analyze children's cohort data.	Develop multi-factorial analyses methods to conduct national-scale regulatory-based risk assessments (program offices); to conduct community-based risk screening and remediation (regions and states); to support epidemiology studies investigating gene-environment interactions (interagency); and to characterize exposure and risk for public health tracking. The results of this effort may be used to develop concepts and tools for application to the Detroit Children's Study, the North Carolina Cohort, and the National Children's Study.

FTEs From ORD Multi-year Plans Associated with the Computational Toxicology New Start Projects Initially Funded in FY05 (Note: FTEs distributions as indicated in the funded project proposal).														
ID	Title	Lead L/C/O	FTEs	Multi Year Plan										
				PM	AT	DW	WQ	ECO	SP2	EDC	HH	CT	HS	HHRA
IA-1	Linkage of Exposure and Effects Using Genomics, Proteomics and Metabonomics in Small Fish Models	NERL	12					5	5	2				
IA-6	Risk Assessment of the inflammogenic and mutagenic effects of diesel exhaust particles: A systems biology approach	NHEERL	2.6	1.2					0.4		1.0			
IA-7	Mechanistic Indicators of Childhood Asthma (MICA)	NHEERL	5.6	1.9	0.2	1.2			0.2		1.6	0.2	0.3	
IIC-1	Simulating Metabolism of Xenobiotic Chemicals as a Predictor of Toxicity	NERL	3.7						1.2			2.5		
IA-4	A Systems Approach to Characterizing and Predicting Thyroid Toxicity Using an Amphibian Model	NHEERL	6.6						1.8	4.7		0.1		
IIC-4	Development of microbial metagenomic markers for environmental monitoring and risk assessment	NRMRL	4.6				4.6							
IIID-6	Use of Toxicogenomic Data in Risk Assessment: workshop and Case Studies for a Chemical in the Androgen-mediated Male Reproductive Development Toxicity Pathway	NCEA	1.6							0.2				1.1
		Total	36.7	3.1	0.2	1.2	4.6	5	8.6	6.9	2.6	2.8	0.3	1.1

Table 3 -FTE's from ORD's Supporting MYPs

Appendix A

1) The Research Center for Environmental Bioinformatics and Computational Toxicology at the University of Medicine & Dentistry of New Jersey (UMDNJ), Piscataway, which will bring together a team of computational scientists with diverse backgrounds in bioinformatics, chemistry and environmental science, from UMDNJ, Rutgers, and Princeton Universities, and the US Food and Drug Administration's Center for Toxicoinformatics. The team will address multiple elements of the source-to-outcome sequence for toxic pollutants as well as develop tools for toxicant characterization. The computational tools developed through this effort will be extensively evaluated and refined through collaboration between Center scientists as well as with colleagues from the three universities and the EPA. Particular emphasis will be on methods that enhance current risk assessment practices and reduce uncertainties. Researchers will also develop a web-accessible Environmental Bioinformatics Knowledge Base that will provide a user-oriented interface to an extensive set of information and modeling resources.

2) The Carolina Environmental Bioinformatics Research Center at the University of North Carolina, Chapel Hill, will develop new analytic and computational methods, create efficient user-friendly tools to disseminate the methods to the wider community, and apply the computational methods to molecular toxicology and other studies. The Center brings together multiple investigators and disciplines, combining expertise in biostatistics, computational biology, chemistry, and computer science to advance the field of Computational Toxicology. Researchers will focus on providing biostatistician support to the Center by performing analyses and developing new methods in Computational Biology. The Center will also create a framework for merging data from various technologies in a systems-biology approach.

Appendix B

IA-1

Title: Linkage of Exposure and Effects Using Genomics, Proteomics, and Metabonomics in Small Fish Models.

Lead: David Bencic, NERL

Research Issue and Relevance: Over the past decade there has been a focused international effort to identify possible adverse effects of EDCs on humans and wildlife. Effects on reproduction and development mediated through alterations in the HPG axis have been of particular concern (USEPA 1998). There is convincing evidence that fish in the environment are being affected by EDCs both at the individual and population levels (World Health Organization 2002). Unfortunately, because feral fish populations are simultaneously exposed to multiple stressors, it is difficult, if not impossible; to accurately assess the role of EDCs in producing adverse impacts (World Health Organization 2002). As a result, many protocols using fish have been developed and validated, both nationally and internationally, for regulatory programs for EDCs (Ankley and Johnson 2004). In the US, a 1996 congressional mandate directed the US Environmental Protection Agency (EPA) to develop a formal screening and testing program for EDCs. Among the Tier 1 tests recommended (USEPA 1998) to detect endocrine disruption of the HPG axis is a 21-d reproduction assay with adult fathead minnows (Ankley et al. 2001). This approach is also being applied to EDC testing with two other small fish models (i.e., medaka, zebrafish) in other countries via activities coordinated by the Organization for Economic Cooperation and Development (OECD; Ankley and Johnson 2004). While of great utility, an important limitation of these tests is that they require significant investment in time and resources. Furthermore, many of the effects endpoints measured in the assays are not diagnostic of specific endocrine-related MOA. An ideal screening assay for EDCs would quickly identify diagnostic endpoints directly indicative of exposure in adult organisms. Detection of anomalies at the genomic level would enable screening methods of shorter duration to identify effects at the molecular level, soon after exposure, before they are manifested at the population level. This research also directly addresses the Computational Toxicology objective of providing approaches for prioritizing chemicals for testing.

Approach: We propose to use a combination of whole organism endpoints, genomic, proteomic, and metabonomic approaches, and computational modeling to (a) identify new molecular biomarkers of exposure to endocrine disrupting compounds (EDCs) representing several modes/mechanisms of action (MOA) and (b) link those biomarkers to effects that are relevant for both diagnostic and predictive risk assessments using small fish models. These goals will be achieved through a three-phase approach that incorporates expertise across EPA/ORD and capitalizes on partnerships with other federal and academic laboratories. During Phase 1, effects of a candidate list of nine compounds having different MOA within the hypothalamic-pituitary-gonadal (HPG) axis will be characterized using the fathead minnow (*Pimephales promelas*) Phase 2 will take advantage of the well characterized zebrafish (*Danio rerio*) genome, to identify transcriptome and proteome level changes in addition to metabolite

changes, associated with zebrafish exposure to the same suite of EDCs. Phase 2 data will be used to identify relevant molecular changes that could (a) serve as diagnostic markers for various types of EDC exposure and (b) begin to inform a systems-level characterization of the responses to those exposures. In Phase 3, candidate genes/diagnostic markers identified in zebrafish (Phase 2) will be validated in fathead minnows through focused toxicological testing aimed at examination of changes in specific gene expression and protein levels. In this way, changes at the genomic, proteomic, and metabonomic level will be linked to one another, linked across multiple teleost species, and ultimately linked to adverse effects at the individual- and, through modeling, population-level (i.e. linkage back to Phase 1). This three phase effort will identify new, potentially cost-effective, diagnostic exposure markers for EDCs, and developing source-to-outcome linkages critical for effective use of biomarkers for risk assessments. This is one of three objectives of the EPA/ORD Computational Toxicology program (Kavlock et al. 2004).

Progress to Date: This is New Start in the Computational Toxicology Research Program and was initiated in January 2005.

Impact: The research described in this proposal will directly and indirectly impact several EPA Program Offices. The most immediate client for the research will by the Office of Science Council and Policy (OSCP) within OPPTS, who are charged with developing and validating screening and testing methods for chemicals that impact the HPG axis of humans and wildlife. ORD (MED) has been working with OSCP for the past 5 years in developing the fathead minnow as a model species for testing endocrine disrupting chemicals, both in the US and, under the auspices of the Organization for Economic Cooperation and Development (OECD), through the world. Because of this existing relationship with OSCP (and OECD), data from the proposed work have a significant and timely impact on international testing programs for endocrine-disrupting chemicals. The Office of Pesticides (also within OPPTS) also will benefit from this research, in two fashions. First, at least two of the test chemicals to be used for these studies (prochloraz, ketoconazole) are of direct interest to the Program Office with respect to both human health and ecological effects. This work also will benefit the Environmental Fate and Effects Division (EFED) within the Office of Pesticides. EFED is responsible for pesticide registration/risk assessments for fish and wildlife; in doing this, they are moving to probabilistic risk assessments that incorporate consideration of mechanism-specific data in the context of population-level effects, an approach parallel to this proposal.

Partnerships/Collaborations: This research is enabled through an extensive network of EPA and non-EPA partners. The Ecological Exposure Research Division (EERD) in NERL has extensive facilities and trained personnel to conduct state-of-the-art molecular biology research in aquatic animals. The Mid-continent Ecology Division (MED) in NHEERL is recognized throughout the world for cutting edge aquatic toxicology research with small fish species, including the fathead minnow. The Ecosystems Research Division (ERD) in NERL will be responsible for metabonomics on the project; this will be achieved through the use of a new 600 mhz NMR. In addition to EPA partnerships on the project a number of external groups will be involved. The EPA STAR program recently awarded a grant Dr. Nancy Denslow (University of Florida) for a proposal entitled “Systems biology modeling of fathead minnow response to endocrine disruptors”. The research proposed by Denslow et al. and that described in the present proposal are extremely complementary. The University of Florida team includes leaders in the

area of ecotoxicogenomics, and strong systems biology component (contributed to the University of Florida effort by Dr. Karen Watanabe, Oregon Health and Sciences University). Further, their proposal has a significant emphasis on effects of estrogens on the HPG axis, which is not covered in this proposal. Hence, combining the efforts of EPA and the University of Florida on this common research problem will result in synergistic outputs. To facilitate this collaboration, Dr. Denslow has agreed for the University of Florida award to be converted from a grant to a cooperative agreement.

Milestones/Products:

FY06

Development of a conceptual model for the HPG axis in small fish models as a basis for focused hypothesis testing of potential endocrine disrupting chemicals.

FY07

Preliminary results on the effects of model chemicals on the fecundity of fathead minnows and single gene and protein expression

FY08

Gene and protein expression as the basis for extrapolation of the effects of endocrine disrupting chemicals across small fish models (fathead minnow and zebrafish)

QA: QA Plan is being evaluated

IA-2

Title: Systems Biology Modeling of Fathead Minnow Response to Endocrine Disruptors

Lead: Nancy Denslow, University of Florida

Research Issue and Relevance: The U.S. Environmental Protection Agency (EPA) is interested in the application of novel technologies, derived from computational chemistry, molecular biology and systems biology, in toxicological risk assessment. In assessing risk associated with exposure to a chemical or other environmental stressor, a number of scientific uncertainties exist along a "source-to-adverse outcome" continuum, beginning with the presence of the chemical in the environment, the uptake and distribution of the chemical in the organism or environment, the presence of the active chemical at a systemic target site, and the series of biological events that lead to the manifestation of an adverse outcome that can be used for risk assessment. The object of this study to develop a computational model to evaluate molecular and protein biomarkers in relation to reproductive dysfunction in fathead minnows exposed to environmental estrogens. The model will incorporate a number of biochemical endpoints along the entire hypothalamic-pituitary-gonadal axis, direct evaluation of physiological changes and reproductive endpoints and the pharmacodynamics and kinetic distribution of the contaminants. The hypothesis that we will test is that genomic and proteomics biomarkers will be diagnostic of the estrogenic effects of environmental estrogens and that they will provide a global understanding of mechanisms of action that will relate specifically to reproductive endpoints in FHM that are adversely affected by exposure to estrogenic compounds.

Approach: Fathead minnows will be exposed to three concentrations each of ethinylestradiol (EE2), and its antagonist ZM 189,154 and to combinations of the two compounds for 48 hrs or 21 days to measure a battery of biochemical, physiological and reproductive endpoints. Short exposures will be used for gene expression and proteomics studies while both short and long exposures will be used to measure the physiologic, biochemical and reproductive endpoints. These data will be brought together in a predictive computational model for the action of environmental estrogens. The model will then be tested with an exposure of fathead minnows to zearalenone (estrogen mimic and positive testor) and trenbolone (nonaromatizable androgen and negative testor), compounds that are used in the cattle industry.

Expected Results: We expect to develop a computational model and identify 10-15 molecular and protein biomarkers that are specific and predictive of adverse effects of exposure to estrogenic compounds in reproduction of fathead minnows. This quantitative model will help improve risk assessment of exposure of wildlife and, by extrapolation, of mammals to endocrine disrupting compounds.

Specific Aims:

Specific Aim 1: To determine and compare gene and protein expression profiles and physiological and reproductive endpoints for adult FHM exposed to a model estrogen 17 alpha-ethinylestradiol (EE2), androgen (17β-trenbolone), or their antagonists (ZM 189,154 and flutamide, respectively).

Specific Aim 2: To predict gene expression patterns of two compounds (zearalenone and EE2) that are environmental estrogens.

Specific Aim 3: To develop a computational modeling framework that integrates exposure concentration, gene expression, and proteomic profiles with physiological endpoints.

Progress Summary: Endocrine disrupting compounds (EDCs) can target the HPG axis at different levels of complexity. These levels span from the molecular level to the whole organism. Our work is centered on understanding how changes at the molecular and protein levels affect downstream physiological endpoints and how well these biomarkers predict adverse effects on reproduction. There are many published studies that have shown a direct impact of EDCs on the reproductive success of fish and wildlife. These compounds include certain organochlorine pesticides, industrial compounds, pharmaceuticals, plasticizers, surfactants, and metals. These compounds have been shown to alter fertility, fecundity, egg hatchability, survival of young, sex ratio, and other reproductive parameters. Although these endpoints have high ecological value, they do not point to specific mechanisms of action. We are pursuing a "Systems Toxicology" approach in which we apply molecular tools to understand global mechanisms of action that are affected by chemical exposure. We will relate these changes to reproductive endpoints using a novel physiology-based mathematical model.

In Year 1 of the project, significant progress was made in the formulation of a physiologically based model of the HPG axis in male FHM. Following the suggestions of our proposal reviewers, a modular approach to model development is being taken so that useful submodels will be developed in the process of constructing an integrative model representing multiple scales of biological organization (from molecular-level gene expression to physiological-level reproductive effects). The first submodel to be developed is a physiologically based pharmacokinetic (PBPK) model that simulates the disposition of EE2 in male FHM.

Milestones/Products:

FY06

To determine and compare gene and protein expression profiles and physiological and reproductive endpoints for adult FHM exposed to a model estrogen 17 alpha-ethinylestradiol (EE2), androgen (17β-trenbolone), or their antagonists (ZM 189,154 and flutamide, respectively).

FY07

To predict gene expression patterns of two compounds (zearalenone and EE2) that are environmental estrogens.

FY08

To develop a computational modeling framework that integrates exposure concentration, gene expression, and proteomic profiles with physiological endpoints.

QA: QA Plan is being evaluated

IA-3

Title: Chemically Induced Changes in Gene Expression Patterns along the HPG Axis at Different Organizational Levels Using a Small Animal Model (Japanese Medaka)

Lead: John Giesy, Michigan State University

Research Issue and Relevance: The EPA is interested in the application of novel technologies, derived from computational chemistry, molecular biology and systems biology, in toxicological risk assessment. In assessing risk associated with exposure to a chemical or other environmental stressor, a number of scientific uncertainties exist along a “source-to-adverse outcome” continuum, beginning with the presence of the chemical in the environment, the uptake and distribution of the chemical in the organism or environment, the presence of the active chemical at a systemic target site, and the series of biological events that lead to the manifestation of an adverse outcome that can be used for risk assessment. "Endocrine-disrupting" compounds have been defined as exogenous agents that interfere with the "synthesis, secretion, transport, binding, action, or elimination of natural hormones in the body that are responsible for the maintenance of homeostasis, reproduction, development, and/or behavior". Much of the recent concern and energy has been focused on compounds that are hormone direct agonists or antagonists, especially those that interact with the estrogen receptor (ER). Effects consistent with exposure to ER agonists have been observed in fish exposed to natural hormones and some synthetic chemicals such as nonylphenol (NP), nonylphenol polyethoxylates (NPEs), and octylphenol (OP). Because chemicals can cause both direct (receptor-mediated) and indirect effects through changes in signal transduction pathways, methods are needed that permit the screening of multiple effects. Furthermore, methods are needed that can screen for these effects simultaneously in a number of tissues, including during critical windows of development, when tissues may be small and the amount of material available for testing is small and difficult to remove from the organism. We propose to develop a screening method to use molecular techniques such as *in situ* hybridization, *in situ* RT-PCR and immuno-histochemical staining (IHCS) to screen for effects of chemicals on the hypothalamic-pituitary-gonadal (HPG) axis with a special emphasis on steroidogenic pathways and hormonal control mechanisms along the HPG-axis in the Japanese medaka. The proposed method will allow for screening of multiple effects in multiple tissues, even at points in development when the tissues are too small to be accurately dissected for use in more traditional molecular techniques. The proposed methods will be applied to a set of "model" and "test" compounds for a set of target genes. Once the methods have been developed and validated, they can be adapted for use with other genes and/or species of interest and used to efficiently and completely screen for endocrine disruptor effects beyond simple receptor binding.

Approach: To investigate the tissue-specific expression of genes in the most efficient manner a variety of *in situ* molecular techniques will be used to visualize and quantify mRNA and/or gene products. To maximize sensitivity and permit multiplexed gene expression quantification methods will be based on *in situ* hybridization and a variety of *in situ* PCR techniques. The aim of these investigations is to produce the simplest and most 'transportable' techniques so the use of a common technique for all genes investigated will be a priority for the final protocol.

Expected Results: The proposed research program will establish an integrative and quantitative small fish model to identify and evaluate the modes of action, key target sites, and biological relevance of EDCs acting along the HPG-axis. To achieve this we will apply state-of-the-art molecular techniques in a whole animal systems approach (see approach section). The studies will be conducted in two phases. In the first phase normal physiological processes along the HPG-axis will be described in order to establish a natural basis for the evaluation of EDC effects. Information on the natural variability, reproductive cycling, diurnal pattern, needs, etc. will be used to optimize the proposed test system. Thus, we will establish a general model system that can be used in the assessment of single chemical and complex mixture effects on reproductive endocrinology. In the second phase we will use this system to detect and assess changes in tissue-specific gene expression patterns (GEPs) caused by the exposure to endocrine active "model" compounds, and their mixtures. We will identify key genes for each individual exposure scenario and establish a quantitative RT-PCR system as a rapid screening tool for these genes.

A model will be developed to predict the biological relevance of the observed changes in GEPs. By identifying the systemic target sites, and the series of biological events from gene expression to the manifestation of an adverse outcome (e.g., reproductive performance), we will identify thresholds at the molecular level that are indicative of effects on the fitness of the individual, including survival, growth and reproduction (fertility and fecundity as well as survival of the offspring). The proposed model test system will provide a new and powerful approach in hazard identification and toxicological risk assessment. Understanding the potential of individual chemicals and complex environmental mixtures to interfere with molecular pathways of concern will enhance our understanding of the basic mechanisms of toxicities, and thus, will have the potential to develop better focused, more rapid and cost effective models for quantitative risk assessment. The bioassay can be used to screen for a wider range of endocrine disruptor effects.

Progress Summary/Accomplishments: A subset of genes has been selected for the development of whole fish *in situ* hybridization methods. Beta-actin was selected as the housekeeping gene, and two isoforms of the aromatase gene, CYP19A and B, were selected as initial target genes due to the fact that their expression is highly tissue specific and responsive to endocrine disruptors. cDNA and riboprobes have been successfully developed for these genes and currently are being tested and optimized using cryosectioned tissues and parafin-embedded sections for whole fish mounts. The *in situ* hybridization procedures are being optimized in the laboratory. The methods have demonstrated a very high level of staining for all probes tested. Although some of the staining seems to be of nonspecific background, the sense and antisense probes do exhibit the expected binding profiles. Certain factors that may decrease the nonspecific binding are in the process of being tested, including high stringency posthybridization washes, high temperature hybridization washes, and high temperature posthybridization washes. At this point, the probes in use are digoxygenin-labeled. They are visualized using standard protocols.

Milestones/Products:

FY06

NCER needs to provide

FY07

QA: QA Plan is being evaluated

IA-4

Title: A Systems Approach to Characterizing and Predicting Thyroid Toxicity Using an Amphibian Model.

Lead: Sigmund J. Degitz, MED/NHEERL

Research Issue and Relevance: The EPA was recently mandated to evaluate the potential effects of chemicals on endocrine function and has identified *Xenopus* as a model organism to use as the basis for a thyroid disruption screening assay. The main objective of this work is to develop a hypothalamic-pituitary-thyroid (HPT) model for this bioassay. This model will provide a rational framework to organize and interpret toxicological data from the molecular to the organismal levels and will serve as a basis for development of predictive tools related to thyroid toxicity. Recent developments in understanding the molecular events involved in TH homeostasis and action suggest that thyroid toxicity might be identifiable using appropriate molecular endpoints in an abbreviated test. This potential improvement, which will lower costs and reduce animal use, is currently more likely to succeed due to expanding genomic information. This project will focus on building linkages between early molecular events associated with exposure and organismal-level effects. Understanding these linkages and their relative importance will be assessed using the HPT model. One of the major goals of this work is to populate the HPT model with data specifically at the molecular and biochemical levels which will provide a basis to interpret interspecies homology and comparative toxicity.

Approach: Development of this assay is progressing, but its widespread use on Agency chemical inventories will be limited due to limited resources. As a consequence, a strategy to objectively rank and prioritize the order of chemical testing needs to be developed. One of the most likely uses for a HPT systems model is to aid in the understanding and discrimination of different toxic modes of action. As such, these models further enable the development of quantitative structure activity relationships (QSARs) by providing a basis for sorting chemicals by mode of action, a necessary step prior to quantifying features of chemical structure associated with a particular type of toxicity. If these relationships can ultimately be established, then predictive models can be developed to rank chemicals for future *in vivo* testing. *In vivo* testing for HPT effects will be improved through this research by providing a basis to link early molecular events to organismal outcomes.

There are four specific aims for this research. The first is to develop an HPT systems model which is capable of integrating data from different levels of biological organization, molecular to organismal, into a coherent system. The second specific aim is to develop an understanding of the compensatory mechanisms at the genomic level involved in TH homeostasis and how they respond to chemical perturbation. The third specific aim is to use *in vitro* models to help define the functions of component systems of the HPT system. Both pituitary and thyroid gland cultures will be used as experimental approaches to define tissue-specific outputs in response to endogenous and xenobiotic chemical inputs. The fourth specific aim is to use the emerging knowledge of the HPT system to develop usable, predictive toxicological tools. Once the input-output relationships are developed for the thyroid glands in culture, this work will be extended to specifically address chemicals known to inhibit T4 synthesis via disruption of iodine uptake by the sodium/iodide symporter (NIS).

Progress to Date: This is a New Start in the Computational Toxicology Research Program that began in January 2005. Past efforts have been focused on the development of a thyroid screening assay for OSCP. This assay, which is based primarily on whole organism responses, is proposed to last 14 - 21 days beginning in late pre-metamorphosis or early pro-metamorphosis using the amphibian, *X. laevis*. We have worked with three T4 synthesis inhibitors (6-propylthiouracil (PTU), perchlorate (PERC), and methimazole (METH)), two receptor agonists (T4 and T3), and a deiodinase inhibitor (iopanoic acid (IOP)) in an effort to test the responsiveness of this model system to chemicals with different modes of action. We have developed dose-response relationships for each of these chemicals. Interestingly, we have found that in the case of the T4 synthesis inhibitors, clear signs of thyroid system disruption, as evidenced by thyroid follicular cell hypertrophy and hyperplasia, occurs at concentrations below those which effect developmental rate. This observation suggests that the HPT executed a successful compensatory response that permitted normal development (Tietge et al, 2004; Degitz et al, 2004). More extreme glandular changes were observed at higher chemical concentrations accompanied by delayed development.

We have initiated studies to examine the changes in molecular and biochemical endpoints specific to components of the HPT system. As a prerequisite to understanding the changes associated with exposure to TH synthesis inhibitors, we have been determining the expression patterns of several genes throughout normal metamorphosis. TSH and type II deiodinase gene expression patterns in the pituitary, for example, have been developed and demonstrate patterns of up-regulation coincident with initiation of metamorphic climax These patterns of gene expression are critically important as they establish the baseline conditions against which changes induced by chemical exposure will be compared and interpreted.

Impact: The primary benefit of this work is to develop a sufficient understanding of the HPT so that predictive models can be developed, testing protocols can be abbreviated, and efforts in inter-species extrapolation can be improved.

Partnerships/Collaborations: This work in being done in conjunction with systems biology modeler at the PNNL.

Milestones/Products:

FY06

Establish data management system
Formulate initial systems model for thyroid function
Develop in vitro thyroid and pituitary culture systems

FY07

Characterize molecular changes associated with TH inhibition
Characterize specific molecular responses in thyroid gland culture
Develop initial QSAR for NIS inhibition

FY08

Refine systems model for relating MOA to outcome
Complete first round QSAR hypothesis for testing for NIS

QA:
QA Plan is being evaluated

IA-5

Title: Estrogen Elicited Gene Expression Network Elucidation in the Rat Uterus

Lead: Tim Zacharewski, Michigan State University

Research Issue and Relevance: The EPA is interested in the application of novel technologies, derived from computational chemistry, molecular biology and systems biology, in toxicological risk assessment. In assessing risk associated with exposure to a chemical or other environmental stressor, a number of scientific uncertainties exist along a "source-to-adverse outcome" continuum, beginning with the presence of the chemical in the environment, the uptake and distribution of the chemical in the organism or environment, the presence of the active chemical at a systemic target site, and the series of biological events that lead to the manifestation of an adverse outcome that can be used for risk assessment. Systems biology involves the iterative development of strategies that integrate disparate physiological and biochemical data into computational models that are capable of predicting the biology of a cell or organism. In order to facilitate hazard identification and risk assessment, a comprehensive quantitative understanding of the molecular, cellular, physiological, and toxicological effects that are elicited following acute and chronic exposure to synthetic and natural chemicals is required within the context of the whole organism. The objective of this proposal is to develop a computational model that will identify and predict critical estrogenic endocrine disruptor elicited changes in gene expression which play a central role in the observed physiological/toxic effects based on systematic and quantitative data obtained from comparative *in silico*, genomic, molecular and histopathological approaches.

Approach: Gene expression and histopathological changes elicited by ethynyl estradiol, genistein, bisphenol A, *o,p'*-DDT will be assessed in ovariectomized immature female Sprague-Dawley rats. Dose-and time-dependent gene expression changes will be determined using customized sequence-verified cDNA rat arrays enriched with estrogen responsive genes. Significant changes in expression will be identified and weighed using Bayes and t-statistic approaches, and verified by quantitative real-time PCR (QRTPCR), western analysis, *in situ* hybridization and/or immunohistochemistry. Chromatin immunoprecipitation (ChIP) assays will further elucidate the estrogen receptor (ER)-mediated mechanisms of action and causal relationships between genes. In addition, histopathological assessments will be conducted to distinguish adaptive and toxic responses using various computational methods including canonical correlation, and Fisher discriminate analyses. Genetic algorithm (GA)/partial least squares (PLS) analysis will then be used to integrate this disparate data into a model that can identify the most relevant genes associated with a histopathological outcome. All data will be captured in dbZach (http://dbzach.fst.msu.edu), a MIAME-compliant toxicogenomic supportive database that facilitates data analysis, integration of disparate data, and sharing with other investigators.

Expected Results: The models developed will identify gene expression changes most highly associated with EED elicited histopathological uterine responses. Examination of multiple EEDs with varying potencies will also identify key regulatory nodes responsible for eliciting these responses, which could lead to the development of high throughput endocrine disruptor screening assays for chemicals in commerce. The data and resulting models can also be

integrated with other algorithms (i.e. PBPK) to create a more comprehensive model of the hypothalamic-pituitary-gonadal axis.

Milestones/Products:

FY06

NCER to provide

FY07

To predict gene expression patterns of two compounds (zearalenone and EE2) that are environmental estrogens.

QA: QA Plan is being evaluated

IA-6

Title: Risk Assessment of the Inflammogenic and mutagenic effects of diesel exhaust particles: A systems biology.

Lead: James Samet, NHEERL

Research Issue and Relevance: Diesel exhaust particulate matter (DEP) is a ubiquitous ambient air contaminant derived from mobile and stationary diesel fuel combustion. Exposure to DEP is associated with carcinogenic and immunotoxic effects in humans and experimental animals. At the cellular level, these health effects are underlain by genotoxic and inflammatory properties of chemical compounds present in DEP. DEP is composed of elemental, inorganic and organic compounds that vary widely in composition with the source of the fuel, engine operating conditions, sampling methods and other parameters. The genotoxic and inflammatory potencies of DEP also vary with its physicochemical properties, and these differences along with multiple health effects impede the development of targeted regulatory strategies for mitigating the impact of DEP exposure on human health. While traditional reductive toxicology approaches are not likely to succeed in quantifying relationships between DEP composition and its numerous health effects, generating a database for modeling the toxicological effects of DEP would provide a framework for quantitative hazard identification. This project proposes a systems approach to developing and applying predictive computational models that quantitatively describe relationships between the composition of DEP and its genotoxic and inflammogenic potencies.

Approach: The objectives of this project will be met by conducting research in three phases. In phase 1 (Specific Aim 1), 16 distinct DEP will be generated using a combination of fuels, engine types, engine loads and collection temperatures. These DEP will then be characterized through extensive chemical and physical analyses. In phase 2 (Specific Aims 2 and 3), the inflammogenic and genotoxic potencies of each of the 16 DEP will be determined quantitatively. Specific bioassays will measure the expression of the pivotal inflammatory mediator IL-8/MIP-2 in cultured human and mouse lung cells in response to DEP exposure. Signaling mechanisms that regulate the expression of IL-8/MIP-2 in response to DEP exposure will also be examined in order to provide mechanistic insight and support for the models. The genotoxicity of the 16 DEP will be assayed using bacterial mutagenicity assays. Human, mouse and bacterial gene expression arrays will be used to provide additional mechanistic insights on patterns of gene expression induced by DEP. Phase 3 (Specific Aim 4) will utilize the generated data to construct a series of statistical and mathematical models that quantitatively relate DEP composition, its inflammogenic and mutagenic effects and the relevant intracellular signaling mechanisms.

Progress to Date: See below

Impact: The information generated by this multidisciplinary research program is intended for use in risk assessment aimed at mitigating the health effects of DEP exposure, including quantitative and computational approaches, cross-species extrapolation and endpoint validation. This proposal is directly responsive to priority research needs identified by NCEA, Office of Air and Radiation (OAR) as well as the ORD PM Research Program. The research will be conducted using a custom-designed diesel emission sampling system, leading edge genomics and

proteomics technologies and the latest tools in bioinformatics and computational modeling software. Beyond providing biological plausibility in support of DEP regulations, the usefulness of the findings from these studies will hinge upon identifying toxicological effects mechanisms in the context of DEP characteristics. Thus, this project's primary aim is to deliver a set of predictive models that quantitatively describe the relationship between DEP composition and its genotoxic and inflammogenic properties. An important impact of the process of developing and applying this set of models via the systems biology approach will be the generation of novel mechanistic data with the specific intent of identifying and characterizing critical paths that lead from DEP exposure to toxicity. It is anticipated that many of these data will be applicable to the study of other sources of particulate air pollution whose effects include mutagenisis or inflammatory responses. Moreover, application of the systems biology approach will represent a case study adaptable to computational studies of exposure and toxicological effects of a broad range of environmental agents. Although the reductive studies will also yield novel mechanistic information about DEP toxicity, the actual deliverable product of this research will be a series of computational models that can be used by client offices in support of assessment and regulatory efforts. These models will have particular practical application in offices responsible for DEP assessment and regulatory programs including the Office of Air and Radiation (OAR) and the National Center for Environmental Assessment (NCEA).

Partnerships/Collaborations: The NRMRL high bay facility is designed for combustion research. Drs Linak and Gilmour have retrofitted a combustion bay complete with dilution tunnel and baghouse collection units. In addition, the diesel exhaust may be piped to inhalation exposure chambers. The facility is equipped with a computer-integrated gas and particle-monitoring bench. The NRMRL research laboratories and the Environmental Carcinogenesis Division also operate core organic chemistry laboratories for routine solvent extraction and GC-MS capability. The laboratories of Drs Samet, Reed and Gilmour are equipped with cell culture areas, nucleic acid extraction and measurement instruments, ELISA, western blot and real time PCR capability. Dr DeMarini's laboratories have capabilities for high throughput screening of bacterial mutagenicity assays using a number of *Salmonella Sp* substrains, and through other members of ECD has access to methodologies for eukaryotic DNA damage and repair testing. Additional core support for genomics and bioinformatics will be available to all investigators through the NHEERL genomics and proteomics Committee (NGPC). As a faculty member of the University of North Carolina at Chapel Hill (UNC-CH), Dr. Reed has access to the resources of the UNC-CH Center for Bioinformatics and the UNC Shared Bioinformatics Resource that provide access to software packages for the analysis of gene expression profiling data, including GeneSpring (Silicon Genetics), PathArt (Jubilant Biosys) and GenoMax (InforMax).

Milestones/Products:

FY06

Generate the first four DEP samples
Conduct chemical characterization of the DEP samples
Conduct a pilot study assessing the inflammogenicity and signal transduction profile of two DEP samples in human and rodent airway epithelial cells.

FY07

Generate eight DEP samples
Identify overlaps in human and mouse gene expression patterns associated with DEP inflammogenicity or mutagenicity
Perform a bioassay-directed fractionation on the available DEPs and determine the distribution of mass and mutagenicity among the fractions.

FY08

Generate remaining four DEP samples
Define signal transduction pathways involved in inflammogenic and mutagenic responses to DEP exposure based on gene expression patterns
Complete biological models of relevant signaling pathways

QA:

QA Plan is being evaluated

IA-7

Title: Mechanistic Indicators of Childhood Asthma (MICA)

Lead: Jane Gallagher NHEERL; Haluk Ozkaynak NERL

Research Issue and Relevance: The US Environmental Protection Agency (EPA) is interested in the interplay of environmental and genetic factors on the development and exacerbation of asthma. The Mechanistic Indicators of Childhood Asthma (MICA) study will use exposure measurements and markers of environmental exposure and health effects. Based on epidemiological studies of air pollution and asthma, there is sufficient evidence to justify investigations that incorporate state-of-the art technologies including genomics and proteomics, to study how and which genes and environmental factors interact in a way that increases the risk of worsening asthmatic responses. EPA scientists will use collected markers of exposure and effects to analyze, characterize, and possibly quantify combined risk factors that relate to asthma severity from multiple agents/stressors. Our study will also provide information on some key molecular events associated with chemical exposures, giving risk assessors more reliable data to assist in defining exposure-response relationships and in making estimates on the range of risks expected in the population compared to data based on biological monitoring and/or screening level hazard data.

Approach: Phase 1 of MICA includes an assessment gene expression data collected from rodent blood and respiratory tissues RNA produced by short-term controlled exposures to concentrated particulate matter. Mobile units housed with rodent exposure chambers (State University) were employed in Grand Rapids, Michigan and in the Detroit Metropolitan Area. Phase 1 will provide 1) information on the reliability of surrogate blood RNA samples to predict target tissues effects and 2) context for the human gene expression data, collected in phase 2 planned for the summer of 2006. Phase 2 is a children exposure assessment/biomarker study focusing on three broad classes of particulate associated chemicals: volatile organic compounds (VOCs), metals, and polycyclic aromatic hydrocarbons (PAHs). MICA will study 150 asthmatic and 50 non-asthmatic children. Blood and urinary measures of these chemicals will be compared to benchmark levels of these chemicals and metabolites from the Center for Disease Control and Prevention's (CDC) National Exposure Report. MICA consists of clinical measurements including (a) a skin prick test for allergen sensitivity; (b) analysis of blood, urine, nail clippings, and DNA; (c) immunological markers, odor testing, lung function and breath analysis; (d) gene expression and protein tests, viewed in the context of environmental assessments and respiratory health history.

Progress to Date: New start Jan 2005. Rodent exposures completed. Blood RNA isolated from two species and two sites in Michigan. Isolation of RNA from matching rodent lung samples ongoing. Collaborations and/or contracts for over 30 exposure and health effects markers have been identified.

Impact: 1). MICA provides linkages in the source-to-outcome paradigm
2). Enhancement of quantitative risk assessment, produce better methods and predictive models for quantitative risk assessment and to provide a useful tool for large-scale biomonitoring in

humans.

3). Better determination of the shape of the exposure-response curve, especially in the low-exposure region, through the incorporation of gene expression data into experimental systems.

4). The development of more accurate, biologically-based mathematical exposure-response models that predict responses outside the range of experimental values.

5)The identification of regulatory metabolic or physiologic pathways, that may act in concert and lead to adverse health outcomes, through the evaluation of multiple health end points with linkages to gene expression changes. •Identification of molecular indicators of exposure and toxicity that can be applied to other epidemiological studies or risk assessment.

More specifically the Office of Air and Radiation (Office of Indoor Air, Office of Air Quality Standards and Planning and the Office of Transportation and Air Quality) and the Office of Environmental Information. Our study will provide information on some key molecular events associated with chemical exposures, giving risk assessors more reliable data to assist in defining exposure-response relationships and in making estimates on the range of risks expected in the population compared to data based on biological monitoring and/or screening level hazard data. Research will support the Government Performance and Results Act (GPRA) goal "Risk Assessment for Susceptible Subpopulations", Long-Term Goal 4 of the Human Health Multi-Year Plan (MYP), Objective 4.4 Science/Research, Sub-Objective 4.4.2 Research. More specifically, the proposed research would be associated with Annual Performance Goal (APG) #6: "By 2012, provide risk assessors and managers with methods and tools for measuring exposure and predicting effects in children, including adolescents, characterizing cancer and non-cancer hazards and risk to children, and reducing risks to children in schools from harmful environmental agents to support EPA risk assessment and risk management"

Partnerships/Collaborations: (e.g., where is the data coming from?). This is a multi-disciplinary proposal requiring the expertise of epidemiologists, modelers, toxicologists, and risk assessors, and the knowledge of computational data analysis and systems biology from NERL and NHEERL investigators, EPA Region 5 and potentially Detroit's City Health Department (Department of Health and Wellness Promotion (a recipient of a CDC grant to conduct asthma education, training and outreach activities). It is expected the DCHD or other recognized community leaders will have a role in effectively educate parents of the enrolled children related to potential asthma triggers and how best to communicate health and research data back to the community participants.

Milestones/Products:

<u>FY05</u>

- Collect bloods and isolate RNA from rodent lung and blood samples and blood
- Develop QA plan for gene expression analysis
- Develop and seek approval for Intramural Research Protocol including recommended and standard operating protocols for collecting pilot study gene expression analysis data.

<u>FY06</u>

- Develop relevant educational modules appropriate for schoolchildren and their parents to

enhance participant recruitment and ensure their informed consent.

- Develop asthma severity score criteria based on asthma diagnosis, medicine and FEV_1 in school-based questionnaires and lung function measurements for selection of schoolchildren.
- Develop standard operating procedures for the collection, processing, and analysis of biological samples from children.
- Complete monitoring of ambient exposure levels indoor/outdoor

FY07

- Develop plan for collection and analysis of biological samples.
- Collect, prepare process and analyze human biological samples from 200 children.
- Prepare comprehensive data base.

FY08

- Complete a data base of all biological measurements and analytic results and the transfer of this data to modelers and risk assessors for the preparation of manuscripts.
- Develop exposure models based on ambient exposure data, geographic information (locations of roadways and major sources of PAHs and metals), and home /school locations.
- Compare exposure model estimates with biomarkers of exposure and early effect.

QA:
QA Plan is being evaluated

IIB-1

Title: Integrated Chemical Information Technologies Applied to Toxicology

Lead: Ann Richard, NCCT (FTE .0.8)

Research Issue and Relevance: A central regulatory mandate of the Environmental Protection Agency, spanning many Program Offices and issues, is to assess the potential health and environmental risks of large numbers of chemicals released into the environment, often in the absence of relevant test data. Models for predicting potential adverse effects of chemicals based primarily on chemical structure play a central role in prioritization and screening strategies yet are highly dependent and conditional upon the data used for developing such models. Hence, limits on data quantity, quality, and availability are considered by many to be the largest hurdles to improving prediction models in diverse areas of toxicology. Generation of new toxicity data for additional chemicals and endpoints, development of new high-throughput, mechanistically relevant bioassays, and increased generation of genomics and proteomics data that can clarify relevant mechanisms will all play important roles in improving future SAR prediction models. The potential for much greater immediate gains, across large domains of chemical and toxicity space, comes from maximizing the ability to mine and model useful information from existing toxicity data, data that represent huge past investment in research and testing expenditures. In addition, the ability to place newer "omics" data, data that potentially span many possible domains of toxicological effects, in the broader context of historical data is the means for optimizing the value of these new data.

The challenges for application of information technologies, including chem-informatics and bioinformatics, are fourfold: 1) to more efficiently migrate legacy toxicity data from diverse sources into standardized, electronic, open, and searchable forms into the public domain; 2) to employ new technologies to mine existing data for coherent patterns that can provide scientific underpinning for extrapolations; 3) to place a new chemical, of unknown hazard, appropriately in the context of existing data and chemical and biological understanding; and 4) to integrate data from different domains of toxicology and newer "omics" experiments to look beyond traditional means for classifying chemicals, inferring modes of action, and predicting potential adverse effects.

Approach: In the area of improving data resources for structure-based mining, the NCCT is supporting further development and expansion of the DSSTox (Distributed Structure-Searchable Toxicity) database network. The DSSTox project is primarily focused on migrating toxicity data from diverse areas of study into structure-annotated, standardized form for use in relational structure-based searching and structure-activity model development. An essential element of this effort involves bridging understanding and forging productive linkages between the toxicology domain experts and the data users and modelers by means of focus on clarifying the chemistry content and summary presentation of the toxicology data. The larger goal of these efforts is to, in effect, overcome inherent and limiting data constraints in focused domains of toxicological study (e.g., cancer, developmental toxicity, neurotoxicity, etc) by expanding the searchable and mine-able data network across both chemical and biological domains.

As an extension of the DSSTox project, NCCT researchers are promoting adoption of standardized chemical structure data fields for public toxicogenomics datasets to enable broader searchability across these data domains, and to enable integration of these datasets with legacy toxicity data and other public data. In particular, collaboration of the DSSTox project with the NIEHS Chemical Effects in Biological Systems (CEBS) project is working towards incorporation of DSSTox data fields and providing structure-searching capability and linkages across CEBS data and public genomics data, as well as DSSTox and National Toxicology Program legacy toxicity databases. Chemical structure and genomic expression patterns provide common metrics for exploring diverse toxicological effects, and can provide the basis for development of predictive patterns or signatures of a toxicological effect. Similarly, biological activity profiles consisting of experimentally determined, or computationally predicted interaction spectra (receptors, proteins, enzymes) could be viewed as expanded "properties" of the chemical and could augment structure-based information for enhancing toxicity classification and prediction algorithms.

Finally, NCCT researchers are taking a lead in efforts to address more fundamental and essential needs to migrate older paper legacy data (such as within EPA Program Offices such as OPP and OPPT) into electronic form suitable for incorporation into standardized, searchable relational databases. New commercial technologies from IBM, SciTegic and others that allow for more automated structure-annotation, and chemical indexing and retrieval procedures are being evaluated to facilitate efficient electronic conversion and structured content-annotation of legacy EPA data. In addition, related issues of quality control of chemical information are being addressed, and Agency-wide chemical structure-browser capabilities are being explored.

Progress to date:

DSSTox Project Status

The EPA DSSTox website (http://www.epa.gov/nheerl/dsstox/), launched in March 2004, provides detailed information on DSSTox standard chemical fields, guidance for creating new DSSTox databases, and links to a wide range of public information resources. A major emphasis of the DSSTox project is on creating field-delimited, content-enhanced data files for diverse toxicity endpoints. Five DSSTox databases are currently published on the website and several others are in progress or currently undergoing review. Toxicity endpoints considered include: rodent carcinogenicity, mutagenicity, estrogen receptor binding affinity, fish acute toxicity, and pharmaceutical maximum adverse effect dose levels. Additional toxicity endpoints slated for DSSTox database publication include: skin sensitization, acute toxicity, nasal and eye irritation, androgen receptor binding, rodent developmental toxicity, DNA intercalation, pesticide ecotoxicity, and immunotoxicity. A large emphasis has been placed on the quality review of chemical information, which has led to the creation of a central DSSTox Master chemical structure reference data file and detailed quality data review procedures.

CEBS DSSTox Project Status

The DSSTox project is collaborating with the NIEHS National Toxicogenomics Center CEBS Knowledge-Base project firstly by the incorporation of DSSTox standard chemical fields into the data dictionary and CEBS data entry system. DSSTox Standard Chemical Fields (SCFs) have recently been revised to better handle the diverse chemical content of public toxicity databases, which include all variety of mixtures and less well-defined substances, and to better coordinate

with other public data standards efforts, such as the ToxML public toxicity data schema project. These DSSTox SCFs will be additionally employed to index the largest public genomics databases, to provide expanded structure-searchability within and outside CEBS data content. A chemical inventory of these databases has begun and will be followed by attempts to encourage external public data standards organizations (e.g., MGED) and database sources (ArrayExpress, NCBI) to adopt more rigorous chemical structure annotations of genomics data. Coordination with other large public database efforts, such as the National Library of Medicine's PubChem, NIH Molecular Libraries Initiative, and NIH National Cancer Institute molecular database projects, will also directly impact on the CEBS DSSTox project collaboration.

EPA Communities of Practice – Chemoinformatics Workgroup Status
As part of the effort to improve our ability to index, search and link chemical information data files across EPA Labs, Centers, and Program Offices, NCCT Researchers have formed a "Communities of Practice" Chemo-informatics Workgroup to begin to forge Agency-wide collaborations and coordination with respect to improving treatment and utility of chemical structure-related information within EPA data files. Additionally, the National Computer Center's Scientific Visualization group is evaluating possible solutions for providing Agency wide structure browsing capability.

Impact: NCCT researchers are involved in efforts that are poised to dramatically improve capabilities to access, mine, and integrate useful chemical-biological activity information from existing and new data, both within and outside EPA. These efforts have the potential to impact a wide variety of EPA program offices that heavily rely on chemical information resources, have large internal stores of data, and have a need for structure-based data exploration, analog searching, and improved toxicity prediction models. These include many programs within OPPTS [e.g., Green Chemistry, Premanufacture-Notification Program (PMN), Office of Pesticide Programs (OPP), High-Production Volume (HPV) Testing Program] as well as EPA's Integrated Risk Information System (IRIS) Program, Office of Water, and Office of Environmental Information. New information technologies that incorporate more flexible and diverse means for assessing of biological and chemical similarity will also improve the identification of toxicologically relevant analogs by enhancing the ability to explore data and quantify associations in diverse chemical and biological domains.

Partnerships/Collaborations*:*
The DSSTox public toxicity database effort is being coordinated and linked with a number of other public efforts involving data-mining companies (e.g., LeadScope), non-profits (International Life Sciences Institute – ILSI) and LHASA, UK (VITIC SAR Toxicity Database Project), and government research laboratories (NIEHS, FDA) that are promoting controlled toxicity vocabularies, adopting data standards, and migrating diverse toxicity data into the public domain. The CEBS-DSSTox collaboration more broadly includes collaborations with FDA's National Center for Toxicological Research, for adoption and use of a chemical structure browser, and the NIEHS National Toxicology Program (NTP), for incorporation and structure-linkage to legacy toxicity data in the NTP data files. Coordination of chemical data information standards that will greatly facilitate cross-platform structure-searching is also ongoing with the National Institutes of Standards and Technology and International Union of Pure and Applied

Chemists (NIST-IUPAC), as well as the National Cancer Institute's Structure-Browser, National Library of Medicine (ChemID) and PubChem. Additionally, dialog has begun with the NIEHS NTP and the NIH Molecular Libraries Initiative to coordinate efforts to use chemoinformatic tools and DSSTox project capabilities to propose sets of chemicals for high-throughput screening testing based on chemical structure-activity considerations.

Through current efforts and the enlarged Communities of Practice workgroup, partnerships and collaborations with scientists across EPA are being forged to better improve chemoinformatic capabilities across the Agency. Collaborations are on-going with MED-Duluth and the ASTER/AQUIRE system, and the IRIS project, and are being continued or initiated with scientists in OPP, OPPT, and OEI to inventory, process, review, and integrate data across various Agency Programs from a unified chemical structure perspective.

Milestones/Products*:*

FY05

- Incorporate DSSTox standard chemical fields into CEBS.
- Create and publish DSSTox database and documentation files for NTP Immunotox database; publish a chemical index file for the EPA IRIS programs.
- Establish Communities of Practice - Chemoinformatics Workgroup to begin to coordinate efforts across the Agency to inventory, retrieve, and explore chemical information data records.
- Propose plan to link DSSTox effort with the NLM PubChem project.

FY06

- Create DSSTox standard chemical field index files for public genomics databases and NTP legacy toxicity, to provide linkages within the CEBS relational search environment.
- Create and publish DSSTox database and documentation files for additional published toxicity databases (skin sensitization, acute tox, nasal irritation, etc) and publish additional chemical index files for EPA programs (e.g., HPV, SRS-TSCA).
- Propose plan for coordinated effort to structure-index and quality review chemical information in EPA data files currently on the web and propose plan to provide for an EPA-wide structure browser on the inter- and intranet EPA website.
- Assist with formation of proposed toxicity chemicals subset for high-throughput testing in the NTP/NIH Molecular Libraries Initiative collaboration.

FY07

- Assist with incorporation of NCTR Array Track chemical structure browser technology into CEBS, and full structure-searching integration with DSSTox data files and all CEBS-linked genomics data.
- Establish procedures and protocols for automating the chemical annotation of new data submitted to CEBS from DSSTox Master chemical list.
- Continue expansion of the DSSTox public toxicity database inventory to include varied databases from the published literature and public sources for a wide range of toxicological endpoints.

- Implement plan to structure-index EPA data files and provide EPA website structure-searchability through curated EPA chemical data files.
- Create structure-annotated database of test results for toxicity subset chemicals from the NTP/NIH Molecular Libraries Initiative.

FY08

- Advise and assist with the implementation of chemical definitions (e.g., assigning a chemical to a class) and structure analog searching capability within CEBS, integrated with toxicogenomics data and bioinformatics capabilities, to serve as the foundation for developing chemoinformatics capabilities within CEBS.
- Advise and assist with the development of procedures and capabilities for deriving chemical signatures for predicting toxicity outcomes from the complete profile of CEBS data, with the integration of chemical structure information and chemical analogy and structure-activity relationship concepts.
- Begin to tabulate and explore data from NTP/NIH Molecular Libraries screening collaboration in relation to other DSSTox databases.
- Implement procedures for expanding the structure-annotation of EPA chemical data records and providing multiple methods for flexible structure or analog searching on the EPA website.

QA:

QA Plan is being evaluated

IIC-1

Title: Simulating Metabolism of Xenobiotic Chemicals as a Predictor of Toxicity.

Lead: – Jack Jones, NERL

Research Issue and Relevance: The mission of the EPA is to protect human health and the environment from adverse effects caused by exposure to pollutants in the water, soil, air, and food. Of the approximately 80,000 chemicals used in the US commerce, relatively few have undergone extensive testing to allow a thorough evaluation of risk requiring an extensive array of data such as physicochemical properties, persistence, bioconcentration factor, exposure, uptake, distribution, metabolism, and toxicity. One of the major uncertainties in evaluating risk is determining exposure of the chemical stressor to the target organism. A compounding factor is that biotic and abiotic transformations of the chemical inside the target organism following an exposure may lead to the formation of reactive metabolites that are toxic. For the vast array of organic pollutants, the availability of transformation rate data and metabolite identification is sparse. Experiments to identify metabolic pathways can be analytically demanding and costly, and are often incomplete with respect to potentially reactive intermediates. Thus, elucidating the metabolism of a chemical and formation of reactive metabolites is a major challenge in determining pollutant exposure and toxicity for risk assessments.

Approach: Using computational approaches to screen and prioritize chemicals for risk assessments and to minimize animal toxicity testing has been a goal of EPA and other regulatory agencies for years. *In silico* simulation of biotransformations and descriptions of metabolic maps have great potential for assessing chemical impacts on human and ecosystem health. Predicted metabolism may also be useful for guiding strategic studies to identify 'bioactive' intermediates for toxicity assessment and for pollution prevention by avoiding commercial use of chemicals forming toxic metabolites. The proposed research will develop a capability for forecasting the metabolism of xenobiotic chemicals of EPA interest, to predict what chemical metabolites are the most likely to be formed, and to interface that information with toxic effect models that predict chemical binding to the estrogen receptor (ER), a well-recognized pathway of toxicity leading to endocrine disruption. This project will (a) illustrate the importance of considering metabolic activation in toxic effects modeling to predict not only parent chemical toxic potential but to identify chemical metabolites of equal or greater potential toxicity than the parent chemical, and (b) demonstrate an approach to provide this capability for large chemical lists of risk assessment concern.

Progress to Date: ERD-Athens has initiated *in vitro* metabolism studies using rat microsomes to track chemical transformation and metabolite formation for a select group of chemicals (e.g. conazole fungicides) of importance to OPP. Analytical methods for metabolite identification, utilizing GC/MS, LC/MS, and NMR spectroscopy, are currently being developed and will be applied where appropriate to metabolism studies in the current proposal. An advanced analytical capability at ERD-Athens, a 600 MHz NMR coupled to a LC/MS/MS for metabolite identification and metabolism profiling, will be available in 2005. The cost of testing chemicals as reproductive toxicants precludes the possibility of evaluating large chemical inventories without a robust strategic approach for setting priorities. A systematic approach has been

developed at MED-Duluth for efficiently expanding the small chemical training sets used for QSAR development to adequately cover the chemical lists of EPA concern. Chemical selection algorithms are being developed and strategic testing is ongoing at MED to develop the ER binding QSAR for chemical prioritization.

Impact: The development of a metabolic simulator for prediction of chemical metabolic maps integrated with a quantitative structure-activity relationship (QSAR) toxic effects model will provide effects-based prioritization of chemicals (parent compound and metabolites) relevant to EPA Program Offices. One of the primary objectives of the CompTox Program and this proposal is the development of computational tools for prioritization of chemicals for toxicity testing with the goal of minimizing dependence on test animals. A primary function of OPP and OPPT is to provide ecological and human health risk assessment for chemicals to anticipate and limit adverse outcomes. The significant challenge that EPA Program Offices (OPP and OPPT) face for performing this prioritization process is due to the large chemical inventories (coupled with minimal data requirements) for chemical registration and screening. Clearly, computational tools that provide the necessary information to achieve these Program Office goals are needed. The following outputs for the proposed research will be of benefit in this regard. (1) Provide OPPT and OPP the ability to prioritize chemical lists (based upon predicted toxic effects of parent chemical and metabolites) with reliability estimates for use in chemical evaluations and to rank chemicals for *in vitro* or *in vivo* screening and toxicity testing. (2) Provide capability to OPP and OPPT for predicting bioactive metabolites. (3) Develop searchable metabolism database for OPP and OPPT use for identification of relevant chemical and/or substructures of interest for risk assessment And (4) Provide linkage of effects based toxicity model with metabolic simulator.

Partnerships/Collaborations: The research will proceed iteratively from modeling to testing to refining models, so frequent communication among ORD scientists and outside partners will be essential. ORD investigators will meet at least monthly by conference call and will communicate much more routinely by email. The lead PIs from the Mid-Continent Ecology Division (MED) in Duluth, MN, and at the Ecosystems Research Division (ERD) in Athens, GA will oversee daily research activities at their respective Divisions. Analytical techniques for metabolite identification at MED and ERD will be initially developed by ERD and shared across groups. Enhanced development of the metabolism simulator will be the primary function of the Cooperative Agreement awardee (scientists at Bourgas University) with input and oversight by the lead PIs from ERD and MED. Dr. Jack Jones will serve as Project Officer for the awarded Cooperative Agreement and will routinely interact with Bourgas scientists via e-mail and conference call. Annual on-site meetings will occur between all investigators. Linkage of metabolic simulator outputs with a QSAR-based toxic effects modeling will be done at MED. The breadth of expertise of this scientific team will enhance the planning and execution of this project. There will be frequent interactions with EPA Program Office staff by phone and email to verify chemical lists, to discuss needs and concerns of all involved, and to communicate progress. Program Offices managers and NERL and NHEERL senior management will be regularly informed of research planning and progress.

Milestones/Products:

FY06

For priority chemicals, incorporate available chemical metadata, metabolism data, and metabolic maps into a searchable database for data management and structure/substructure searchable access

Forecast metabolic pathways for selected priority chemicals using existing simulator for liver metabolism

Initiate in vitro Phase I liver microsome experiments; incorporate new laboratory and literature metabolism data into simulator training sets

FY07

Confirm formation of predicted metabolites for priority chemicals and compare observed maps to forecasted maps

Develop and approach to evaluate and enhance simulator performance through improvement of transformation probability estimates and expansion of transformation reaction domain

Begin evaluation of refined simulator and model predictions in context of Program Office prioritization needs

QA:

QA Plan is being evaluated

IIC-2

Title: Modeling Molecular Targets for Toxicity, a Computational Approach to Understanding Key Steps in the Mechanisms for Toxicity and a Tool for Prioritizing Bioassay Requirements

Lead: James Rabinowitz, NCCT (FTE 2.5)

Research Issue and Relevance: The Agency frequently encounters situations where it must make decisions about the potential health and environmental effects of chemicals when all of the relevant data is not available. One rational approach to this problem is to estimate the relevant missing information by extrapolating from existing data. Knowledge of key steps in the potential mechanisms of action provides a template for developing models for this extrapolation. These models may be used to inform experimental studies and provide a tool for prioritizing data requirements. One relevant example is the Agency need to make decisions on the potential of specific anthropogenic chemicals to cause endocrine disruption. Developing molecular models that can be used for rapid screening of the interaction of environmental chemicals with receptors in the endocrine system will provide an important tool for selecting priorities.

Approach: For many mechanisms of toxicity the key differential process is the interaction between the ultimate toxicant and a macromolecular target (receptor-ligand, enzyme-substrate, DNA-genotoxicant, etc.). Modeling this process on a molecular level provides an approach for prioritizing chemical information needs. This interaction initiates a cascade of events leading to the ultimate (adverse) outcome. Computational molecular models of the interaction of a molecule (potential toxicant or its metabolites) with the relevant target, provides insight into the capacity of a chemical to initiate the relevant cascade.

A number of recent scientific and technical advances facilitate this approach. First, many of the relevant targets have been identified experimentally. The molecular structure of some of these targets has been determined and additional information of this type is likely to become available in the future. Second, molecular modeling software for the simulation and analysis of interactions of this type has become more sophisticated in a number of relevant ways and make it possible to more realistically simulate the processes of toxicity. These advances have resulted from both basic computational investigations of the structure and dynamics of macromolecules in biological systems and the requirements of the pharmaceutical industry for the discovery of new therapeutic agents. Third, computational hardware and visualization techniques have steadily improved. Increased processing speed and memory have made it possible to include large segments of macromolecules in classical simulations and even in quantum mechanical calculations. We are applying molecular modeling methods, fueled by these current scientific and technological advances, to investigating chemicals for their capacity to cause toxicity through specific modes of action and using a target-toxicant paradigm.

Initially, this approach is being applied to the study of environmental endocrine disruption. Crystal structures exist in the literature for many receptors in the endocrine system. By removal of the ligand computationally these crystal structures are used to create virtual biomolecular targets for endocrine disruption. The best fit to the target for each of a series of potential ligands can then be determined by computational methods. The properties of this fit may then be used as part of a scheme to predict the potential of an environmental agent to cause endocrine disruption.

The ultimate goal of this research is to develop a library of biomolecular targets for chemical toxicity and methods appropriate for the prediction of the ability of a chemical to interact with these targets. These targets may then by used as part of a chemical prescreen

Partnerships/Collaborations: Scientists from RTD/NHEERL are providing a data base of chemicals that interact with receptors in the endocrine system and also expertise on the endocrine system. Scientists from NTD/NHEERL will provide similar expertise and data and important environmental targets in the nervous system. Commercial molecular modeling software and advice are available from Schrodinger Inc. These products have been developed primarily for the pharmaceutical industry and their use must be adjusted for the needs of this project. Dr. Wei-tao Yang of the Chemistry Department of Duke University is a collaborator for advanced molecular modeling methods.

Progress to Date: There is increasing concern about the potential of environmental chemicals to produce adverse health effects through interaction with the endocrine system. One general mechanism for disruption of the endocrine system involves competition for steroid hormone binding sites by xenobiotic chemicals that may fully or partially mimic natural hormones. Crystal structures of the estrogen and androgen receptors have been used to create macromolecular targets. Multiple crystal structures available for the estrogen receptor have been used to demonstrate the importance of including receptor flexibility in determining the best fit and therefore the chemicals likely have the most effect. Studies on including flexibility are in progress. In other preliminary studies, methods existing methods are being employed. Using a data base containing estrogens, androgens and chemicals that bind to other nuclear, it was found that the results for a single receptor, considered individually, did not correctly order the ligands according to the ability to bind to that receptor. However, when the data for a series of receptors is considered simultaneously and the results analyzed to determine which receptor a chemical interacts with most favorably, all chemicals are classified correctly.

An additional Agency concern, which may be approached with the target-toxicant paradigm and molecular modeling, is the cumulative risk of some specific pesticides on the enzyme acetylcholinesterase (AChE). There is some data that indicates that AChE has two binding sites, the catalytic site and an allosteric site that affects the specificity and efficacy of the enzyme. In experiments involving just a single toxicant it is not possible to identify the site of interaction but we have shown that for mixtures of AChE active chemicals the cumulative toxicity depends on the relative interaction of each chemical with each site. A computational scheme is being developed for the cumulative toxicity of mixtures that takes advantage of the two site model.

Relevance: This research addresses the Agency need for predictive models for hazard identification, both the sub areas of QSAR and other computational approaches and High Throughput Screening.

Milestones:
FY06 – Results of the use of the target-toxicant paradigm to screen for estrogenicity/androgenicity; Results of the affect of the two binding site model on the cumulative risk of chemicals acting through the enzyme AChE.

FY07 – Results of the importance of protein flexibility in evaluating the interaction of chemicals with macromolecular receptors. Report on the importance of protein flexibility in evaluating the interaction of chemicals with macromolecular receptors.
FY08 – Evaluation of the use of the target-toxicant method as a tool in a diverse chemical screen. Application of the target-toxicant approach to other targets of toxicity

QA:
QA Plan is being evaluated

IIC-3

Title: ToxCast, a Tool for Categorization and Prioritization of Chemical Hazard Based on Multi-Dimensional Information Domains.

Lead: Keith Houck, NCCT and David Dix, NCCT (FTE 2.7)

Research Issue and Relevance: Across several EPA Program Offices (e.g., OPPTS, OW, OAR), there is a clear need to develop strategies and methods to screen large numbers of chemicals for potential toxicity, and to use the resulting information to prioritize the use of testing resources towards those entities and endpoints that present the greatest likelihood of risk to human health and the environment. This need could be addressed using the experience of the pharmaceutical industry in the use of advanced modern molecular biology and computational chemistry tools for the development of new drugs, with appropriate adjustment to the needs and desires of environmental toxicology. A conceptual approach named ToxCast has been developed to address the needs of EPA Program Offices in the area of prioritization and screening.

Approach: Modern computational chemistry and molecular biology tools bring enabling technologies forward that can provide information about the physical and biological properties of large numbers of chemicals. The essence of the proposal is to conduct a demonstration project based upon a rich toxicological database (e.g., registered pesticides, or the chemicals tested in the NTP bioassay program), select a fairly large number (50-100 or more chemicals) representative of a number of differing structural classes and phenotypic outcomes (e.g., carcinogens, reproductive toxicants, neurotoxicants), and evaluate them across a broad spectrum of information domains that modern technology has provided (i.e., physical-chemical properties, predicted biological activities based on existing structure-activity models, biochemical properties based on high throughput screening assays, cell based organotypic assays, and genomic analysis of cells or organisms). These domains represent increasing biological relevance, as well as increasing resource requirements. The ultimate goal of the project would be to mine the resulting data for association between and among the various domains and the known toxicological properties of the base set of chemicals in order to provide a structured strategy to identify potential toxicity pathways, and to prioritize chemicals them for subsequent testing based on that information.

The underlying hypothesis is that whether is concerned with the off target effects of drugs, as desired to be understood by the pharmaceutical industry, or toxicity in case of environmental agents of interest to the EPA, the response is driven by interactions with biomolecular targets of one form or another. One needs only to identify those receptors of concern and identify tools for assessing the likelihood of interaction with the chemicals of concern. In moving from the drug development arena (which can be compared to working along one or just a few vectors) to the environmental toxicology arena (which can be likened to working on a matrix instead of a vector), one needs to shift from a specific screening target to a more global agenda, and it becomes necessary to vastly expand the number of potential biomolecular targets, be these obtained from *in silico* assays, biochemical assays, cell based *in vitro* assays, surrogate animal models, or short term studies in traditional species. Hence, a wider

net of endpoints and information sources will be applied, at least initially, as the concept transgresses from a concept to a reality.

Of course, a number of hurdles would need to be addressed before launching such an effort, including: (1) identification of a subset of chemicals for serving as the proof of concept models; (2) developing a chemical inventory management and distribution system; (3) identifying an upper cap on the per chemical cost of obtaining screening level data: (4) selecting assays within the available resources; (5) flexibility, or tiering, of domains based upon pre-existing knowledge; (6) perhaps initially targeting only a few manifestations of toxicity rather the all possible ones to decrease the complexity of the task; (7) evaluating the impact of metabolizing capability, or lack thereof, on the efficiency of the screening assays; (8) developing a bioinformatic approach to mining the resulting data and identifying signatures of concern; and (9) carrying out a prospective assessment of the bioinformatic approach using chemicals currently entering a traditional testing process. These hurdles would be the subject of considerable discussion as the potential feasibility of this concept proposal is discussed further.

Progress to Date: At this point in time, ToxCast is nearly the end of its conceptual design phase. The information domains have been identified, and a number of potential contributing data sources have been investigated (e.g., Iconix, MDS Pharma, CEREP, Amphora, PASS). Recruitment actions are underway to add two staff members to the NCCT who will be responsible for the biological and information processing components of ToxCast. Communications have been established with the NTP/NIEHS which has similar interests and which is beginning to work with the NIH Molecular Libraries Initiative (see also the DSSTox implementation plan). Outreach to the OPPTS, ACC, EDF and other external groups have also begun to help develop a consensus on the specific directions and contents of ToxCast.

Impact: The availability of a biologically and chemically based system to begin to associate chemicals of like properties and activities will provide a number of EPA Program Offices with an extremely useful tool that heretofore has been seriously lacking. The tool may be one of the first broad scale products of the NCCT that addresses the mission of improving the efficiency and effectiveness of hazard identification and risk assessment methodologies employed by the EPA.

Partnerships/Collaborations: The NCCT is working to establish partnerships with a number of external groups that can facilitate development of the information needed in ToxCast. These groups include the OPPTS, the NTP/NIEHS, the ACC, the EDF and a number of commercial vendors that market some of the enabling technologies.

Milestones/Products:

FY06 - Develop conceptual framework for ToxCast

FY07 - Establish initial battery of assays across the information domains, identify list of chemicals to evaluate proof of concept for framework and begin data acquisition

FY08 - Report on the utility of statistical clustering techniques on assay results from pilot chemicals to group them according to known toxicity patterns; revise framework as dictated by results.

QA:
QA Plan is being evaluated

IIC-4

Title: Development of microbial metagenomic markers for environmental monitoring and risk assessment.

Lead: Jorge W. Santo Domingo, NERL

Research Issue and Relevance: The microbiological water quality standards established by EPA depend on culturing fecal indicator bacteria to predict the risks associated with water usage. For decades this has been the favored approach to microbiological monitoring in spite of the fact that culture-based methods tend to underestimate the densities and diversity of microorganisms in environmental samples (Amann et al., 1995). Relevant to public health is the fact that nonculturable pathogens could be the etiological agents of many waterborne illnesses (Sails et al., 2002). In addition, nonculturable organisms may be better indicators of fecal contamination. Moreover, nonculturable bacterial indicators may be useful in the identification of the fecal sources impacting a particular water system. Identifying the sources responsible for the fecal pollution of a natural waterway is important in order to reduce the fecal bacterial levels and in turn reduce the illnesses associated with recreational activities, and food (e.g., fish) and water consumption (Simpson et al., 2002). Consequently, strict dependence on culturing techniques continues to be a notable roadblock in the path towards understanding the correlation between bacterial indicators of fecal contamination and the hazards associated with fecally impacted waters. Nucleic acid-based approaches can circumvent many of the shortcomings of the culture-based methods. For instance, the possibility of rapidly and simultaneously monitoring hundreds of microorganisms and genes relevant to public and environmental health is now becoming a reality in light of the recent advances in microarray technology (Stahl and Tiedje, 2002). Such a capability will categorically improve microbial risk assessment and the framework used in the development of environmental monitoring and risk management tools.

Approach: We propose the development of an innovative metagenomics program to enhance our current capabilities for environmental microbial monitoring, risk exposure, risk management, and risk assessment. The general goal of this study is to develop a novel molecular markers based on fecal microbial genomes to better assess the sources of fecal contamination in natural water systems. Assuming that the evolutionary pressure for the selection of host specific populations must involve ecologically driven processes *we hypothesize that genes playing a role in host-microbial interactions and cell-cell recognition (e.g., cell surface proteins, toxins, and adhesin) are better markers of host specificity. The primary objective of this study is to discover novel and validate the use of fecal metagenomic sequences as MST host specific markers.* By looking at collective genomes we will substantially increase the number of potential genes and microbial species that can be used in the development of molecular microbiological assays (Schloss and Handelsman, 2003). The first phase of this program will focus on constructing fecal metagenomic libraries specific to animals that are known to be relevant pollution sources of watersheds in the United States. The following phases will focus on systematically identifying DNA sequence markers for specific human and bovine fecal sources and testing the specificity of these markers against waters impacted with different pollution levels and pollution sources. We will use genomic subtraction of metagenomic libraries to generate unique sequences from the collective microbial genomes present in animal feces. Due to the complexity of microbial

communities, this approach will facilitate the identification of candidate host specific microbial genes. Real-time PCR assays will be developed to test the host specificity of the latter markers and to quantify the target genes in environmental waters impacted with different levels of fecal contamination. A similar approach will be undertaken with samples from ongoing epidemiological studies. Markers that show host specificity will be included on biochips containing sequences specific to fecal indicator bacteria and pathogens and challenged with DNA extracts of different fecal samples and water samples impacted with different levels of pollution. By combining the epidemiological studies with the host specific assays and pathogenic markers, we hope to gain better predictive capability to determine outcomes from scenarios associated with fecally impacted waters.

Progress to Date: See below

Impact: Distinguishing between human and nonhuman fecal contamination in our waterways has never been more important than in recent years in light of the need by the States to comply with deadlines associated with the Total Maximum Daily Loadings program (Simpson et al., 2002). Accurate assessment of the primary sources of fecal pollution is needed in order to calculate the different risks associated with each of the potential sources impacting water systems and to correctly implement and evaluate Best Management Practices (BMP) used to remediate fecally polluted waters. As currently developed, most MST methods depend on markers that have no ecological meaning. *The main goal of this study is to develop assays that can be used in source identification, environmental monitoring, and risk assessment.* In order to accomplish this goal we will combine a metagenomic-scale approach with high-throughput gene mining (microarrays) and real-time PCR assays. Therefore, thiis research supports EPA's GPRA Goal 2 (Clean and Safe Water). The proposed work will provide a rich source of metagenomic information (i.e., novel molecular markers) for sequenced-based analyses aimed at better monitoring of fecal pollution and improving the assessment of the outcome associated with different pollution sources. Consequently, it is anticipated that the results will provide a molecular-based framework to (1) better predict the relationship between microbial water quality and public health risks, (2) determine the impact of different microbial pollution sources on watershed biology, and (3) effectively control or eliminate pollution in our Nation's watersheds.

Partnerships/Collaborations: As this program is relevant to water fecal contamination and public health issues related to water, we have requested the participation of personnel from EPA's Office of Water and the National Center for Environmental Assessment as consultants. The primary points of contact of the latter organizations will be Dr. Robin Oshiro and Dr. Mary Rothermich. Drs. Oshiro and Rothermich will act as liaisons for their respective organizations. Their role will be to provide advice on how the data can be utilized by their offices. They will be invited to participate in mid year review meetings and conference calls. We have also briefed other personnel from the Office of Water, and they have expressed interested in receiving updates on the progress of this project. As this project has an immediate impact on problems that states and tribes are currently facing, we will keep Drs. Ron Landy and Bobbye Smith (Region 3 and 9 ORD-regional science liaisons, respectively) informed on the current progress of this project. In addition, we will consult with members of EPA's Genomics Task Force Workgroup on issues pertaining Data Management, Data Analysis, and Data Submission and Methods Performance. Drs. Rebecca Calderon and Al Dufour will also be consulted on a regular basis as

they are managing the epidemiological study that is currently been conducted by members of NERL and NHEERL.

Milestones/Products:
FY06 - Evaluation of metagenomic databases as a source of molecular markers to assess human and animal fecal contamination in surface waters

FY07 - Report on the evaluation of PCR-based host-specific assays to confirm the presence of animal fecal sources of pollution in waters impacted by fecal contamination. Report on fecal indicator microorganisms and/or genetic markers from fecal material whose densities in recreational waters best correlate with the rates of illnesses in users of recreational waters.

QA:
QA Plan is being evaluated

IIID-1

Title: Statistical Methodology for Estimating Parameters in PBPK/PD Models

Lead: R. Woodrow Setzer, NCCT (FTE 0.8)

Research Issue and Relevance: PBPK/PD models are large dynamic models that predict tissue concentration and biological effects of a toxicant before PBPK/PD models can be used in risk assessments in the arena of toxicological hypothesis testing, models allow the consequences of alternative mechanistic hypotheses to be predicted in highly non-linear systems. Whether for quantifying the uncertainty of the extrapolation of a dose metric or determining which hypothesis is better supported by the evidence, and by how much, it is critical that the uncertainties inherent in such modeling be properly quantified, and established principles of statistical inference be brought to bear.

However, such quantification is far from simple. Biological systems models generally involve large numbers of parameters whose values are known with varying degrees of uncertainty. Some parameters (for example, tissue-specific fractions of cardiac output and fractions of body mass) can reasonably be thought of as varying among individuals of a population. Other parameter values have been estimated in *in vitro* systems, or may have been calculated from properties of the chemicals involved (for example, partition coefficients in a physiologically-based pharmacokinetic model). It is typically the case that the values of parameters in a biological systems model will be derived from the results of several different experiments, of very heterogeneous designs, and often conducted in different laboratories at different times. Over the last decade or so a small number of statistical investigators have begun to explore the issues involved in applying statistical methods to such systems. While there are examples of the application of statistical methodology to PBPK/PD models in the literature, there is as yet no coherent approach to this activity that takes into account the special issues that arise in PBPK/PD modeling. The NCCT is establishing a program to further develop the statistical methodology to support the rigorous quantification of both the uncertainty of parameter estimates and of model predictions, with the aim of providing a framework for statistical applications in PBPK/PD modeling.

Approach: The outstanding issues are a mix of computational and theoretical statistical problems, and are best investigated through exploring specific models and datasets. Thus, we need to identify PBPK/PD models for chemicals of interest to the Agency, identify the outstanding theoretical statistical issues that need to be solved, and set up a computing platform to handle the large computing loads this exploration will entail. Finally, theoretical statistical exploration needs to precede hand-in-hand with practical PBPK/PD model development, ultimately leading to the publication of a practical framework for applying statistical methodology to PBPK/PD models.

Progress to Date: Investigators in NHEERL, NERL, and CIIT Centers for Health Research have been approached for appropriate models and data sets, and several models have been identified for further work. A graduate student in biostatistics and his advisor have been recruited to work on outstanding theoretical issues and apply solutions to models and data sets of

interest to the Agency. The literature surrounding this topic has been surveyed, and a literature review is in progress.

Impact: PBPK/PD models are of increasing interest in the Agency for use in extrapolation from animals to humans, across routes of exposure, and for integrating acute effects over discrete episodic exposures. Better characterization of methods for estimating parameters and quantifying uncertainties in such models predictions will remove one major impediment to their more general application. This will allow replacement of default uncertainty factors with transparent mechanism-based statements of scale and uncertainty, in turn decreasing the subjectivity and increasing the transparency of environmental health risk assessments.

Partnerships/Collaborations: Collaborations with NERL, NHEERL, and the CIIT Centers for Health Research and the Department of Biostatistics in the UNC School of Public Health are being developed to develop statistical methods for parameter estimation in PBPK/PD models.

Milestones/Products:
FY06 – Submission of a journal article outlining outstanding statistical issues in the analysis of PBPK models.

FY07 - Joint NIEHS/NCCT workshop on Statistical Issues in PBPK/PD modeling.

FY08 - Publication of framework for the statistical analysis of PBPK/PD models.

QA:
QA Plan is being evaluated

IIID-2

Title: Modeling Toxicokinetics for Cross-Species Extrapolation of Developmental Effects

Lead: Hugh Barton, NCCT (FTE 1.4)

Research Issue and Relevance: Animal toxicology studies used to evaluate the potential for effects due to exposures during developmental periods are extrapolated to humans based upon the maternal exposure dose. The approach does not address whether the toxicokinetics are similar across species during the relevant critical developmental window. In this research program we will develop toxicokinetic models for estimating internal dosimetry during early life stages to serve as a basis for improved interspecies extrapolation. Absent information specifying the critical window, estimates of internal dosimetry during the in utero, lactational, and early post-weaning periods will characterize the uncertainty in evaluations based upon maternal exposure dose. With information on the critical window, models may describe toxicokinetic and toxicodynamic processes in the relevant toxicological species and humans permitting improved quantification.

Approach: This research will incorporate a range of pharmacokinetic modeling methods including noncompartmental analyses and classical and physiologically-based pharmacokinetic modeling to describe dosimetry in early life-stages relevant to one-generation toxicity studies (i.e. maternal exposure resulting in fetal exposure, lactation exposure of pups after birth, and exposures of growing pups). Perfluooctanoate (PFOA) and related compounds will be used as initial prototype chemicals due to the importance of effects observed in offspring, which are used as endpoints in risk assessments. Lactational exposure will be modeled first because dosimetry during this period has not yet been evaluated. A physiologically based pharmacokinetic (PBPK) model for PFOA will be developed to characterize dosimetry during these different lifestages and identify data gaps that limit the ability of the model to address both rats and humans. Finally, this early life stage model will be used to evaluate dosimetry during different early life periods for a selected group of other prototype compounds for which developmental toxicities are observed. Prototype compounds would be selected to have a range of pharmacokinetic and physical chemical properties. This analysis will aid in the evaluation of the uncertainties in the default risk assessment approach based upon maternal exposure.

Progress to Date: Literature reviews have been ongoing to obtain physiological parameters for growing rat pups. Compartmental pharmacokinetic modeling of PFOA in adult animals, including pregnant and lactating females, has been completed.

Impact: Biologically based modeling of toxicokinetics and toxicodynamics during developmental periods can improve chemical risk assessments based upon developmental effects. At a minimum, characterization of internal dosimetry will inform the uncertainties attendant to analyses based upon the maternal exposure dose. In the presence of information on the critical window, the models may directly form the basis of the quantitative risk assessment by deriving the relevant internal dose metric for extrapolation to humans. Perfluorinated compounds, including PFOA, are an important proof of concept for this

research because they are a class of compounds producing developmental effects of significant regulatory concern.

Partnerships/Collaborations: There are active collaborations with NHEERL, NERL, and NTP to collect data for perfluorinated compounds and NHEERL to collect data on conazole pesticides. Evaluation of published literature to obtain physiological parameters has been an ongoing collaboration with NCEA. In addition, this project complements ongoing efforts by others to develop models for human children. Sharing of information and physiological parameters has been ongoing through an early life physiological parameters database project at the International Life Sciences Institute with participating by US and Canadian government agencies, academic institutions, companies and non-profit organizations.

Milestones/Products:
FY06 - Model rat pup lactational exposure for perfluooctanoate (PFOA) using compartmental approaches.

FY07 - Develop initial physiologically-based pharmacokinetic model for PFOA in rat maternal-fetal-pup unit and identify data gaps in relation to rat and human.

FY08 - Evaluate modeled dosimetry for rat fetus and pup for a limited number of prototype compounds to inform the uncertainty in use of maternal exposure dose in risk assessments.

QA:
QA Plan is being evaluated

IIID-3

Title: Development of a portable software language for physiologically-based pharmacokinetic (PBPK) models.

Lead: R. Woodrow Setzer, NCCT (FTE 0.1)

Research Issue and Relevance: The PBPK modeling community has had a long-standing problem with modeling software compatibility. The numerous software packages used for PBPK models are, at best, minimally compatible. This creates problems ranging from model obsolescence due to software support discontinuation to difficulty in linking models created in different packages.

The objective of this research effort is to create an XML extension to be used in conjunction with the systems biology markup language (SBML), an XML-based, machine-readable format for representing models of biochemical reaction networks, including metabolic networks, cell-signaling pathways, and regulatory networks. This extension will augment SBML to accommodate PBPK models with reasonable complexity, including full documentation, PBPK modeling vocabulary, and examples for instructional and demonstrational purposes. This extension will be constructed to permit seamless linking of PBPK models to the biological pathway models abundantly available in SBML format. A secondary goal is to develop a pseudo-generalized visualization program that is capable of translating an SBML file for a PBPK model into a visual schematic along with equations representing the schematic in differential equation form, accompanied by parameter values and references for the parameters.

End-users of a PBPK model, such as risk assessors, would benefit greatly from being able to use the software of their choice to run the model and to visualize the model itself and documentation for it, making incorporation of PBPK models in the risk assessment process easier and more efficient.

Approach: We will begin by coding simple to moderately complex PBPK models in SBML and testing the results. This initial step will map out the limitations and weaknesses of SBML for accommodating PBPK models and give insight into the magnitude of the work involved in creating an appropriate SBML extension. This initial step will be an iterative process with complexity of the models increasing until all major necessary attributes of the extension are uncovered. The next step will be to create an initial version of the schema and extensively test it. An NCCT-sponsored meeting will be held with all interested parties where the schema will be presented and a plan for moving forward will be presented and debated. This plan will include formal and informal collaborations and partnerships. Once a final version of the extension has been developed and fully tested, a formal SBML workgroup will be formed and a proposal will be written to the SBML Community at Large for incorporation of the extension into the current version of the base SBML schema. At this point, SBML would fully accommodate PBPK models. Simultaneously, we will be developing the visualization program with the objective of using it not only as an aid in the development and usage of PBPK models, but also as a powerful educational tool.

Progress to Date:

A team has been assembled to develop the schema as well as advise and manage the project. Some necessary features of the extension have been determined and we are currently coding a simple PBPK model in SBML as part of the process of determining all necessary attributes. The interest of the PBPK community has been assessed and a list of interested parties, some possibly for collaboration, has been created.

Impact:

The ability to create, transfer, use, augment, and review PBPK models without the limitation of software compatibility is a long-standing desire in the PBPK community. The ability to seamlessly link PBPK models to biological pathway models is of great interest, since this gives the modeler a fast and efficient means of extending the estimations of tissue dose from the PBPK model to cellular-level responses, thereby facilitating a quantitative investigation of dose-response relationships through mode/mechanism of action. By making the use of PBPK models faster, easier, and more reliable, this product will make the use of these models more common and widespread in risk assessment.

Partnerships/Collaborations:

Formal collaboration has been established with the Lockheed Martin Corporation through the U.S. EPA Office of Environmental Information. The possibility of collaborating with The Aegis Technologies Group (developers of the prominent PBPK modeling software package, ACSLXtreme) is currently being explored. Contact with original SBML developers and project managers have been established with positive feedback. Possible collaboration with CIIT Centers for Health Research is also being explored. Groups that have expressed interest in being immediate clients for the products of this project include EPA/NCEA and NIEHS/NTP.

Milestones/Products:

FY06 – Report on findings of assessing extensiveness required for the XML schema.

FY07 – NCCT-sponsored meeting on evaluation of the performance of the completed XML schema and status of the visualization program.

FY08 – Report on proposal to have finalized XML schema included in the next release of SBML and evaluation of visualization program.

QA:

QA Plan is being evaluated

IIID-4

Title: Systems Modeling of Prostate Regulation and Response to Antiandrogen

Lead: Hugh Barton, NCCT (FTE 0.3)

Research Issue and Relevance: The prostate is an androgen-dependent tissue that is an important site of disease in human males as well as an important indicator of androgen status in animals. The rat prostate is used for studying antiandrogenic drugs as well as for evaluation of endocrine disruption (e.g., Hershberger Assay). Pubertal changes in the prostate have been observed to be as sensitive to environmental antiandrogens as in utero effects. The goal of this research is to model the biology of prostate androgen function on a systems level to determine the factors responsible for the dose-response observable with androgens and antiandrogens in the male rat. This includes investigation of the roles of positive and negative feedback loops in prostatic response following castration and dosing with testosterone and/or antiandrogens.

Approach: A biologically-based, systems-level model will be developed describing the regulation of the prostate by androgens. The model will extend an existing model for the male rat central axis, which describes feedback between luteinizing hormone and testosterone production in the testes, to include the prostate and conversion of testosterone to dihydrotestosterone (DHT). The prostate model will describe binding of androgens to the androgen receptor, 5α-reductase catalyzed production of DHT, and gene regulation affecting cell proliferation, apoptosis, and prostatic fluid production. The model will combine pharmacokinetic models for endogenous hormones (i.e., testosterone, DHT, LH) and exogenous antiandrogens (e.g., finasteride, flutamide or casodex, vinclozolin), and a pharmacodynamic model for androgen-dependent prostate functions. Linkages of this model with genomics data obtained with castration, testosterone treatments, or antiandrogen treatment will be explored to assist in developing perspectives on how such data would fit into quantitative risk assessments.

Progress to Date: A model has been constructed that recapitulates the time-dependent changes in prostate following castration of male rats resulting in decreased serum testosterone, prostatic androgen receptor, and prostate weight. The model currently poorly describes the dose-response for less drastic changes in testosterone; this is under investigation. A classical compartmental model for finasteride, a therapeutic 5α-reductase inhibitor, has been linked to the biologically-based model for the rat endogenous hormone function. Opportunities to link the model with genomics data are being explored.

Impact:

This project is designed to evaluate how perturbations of endogenous biological systems produce observable dose-response behaviors using a widely used endpoint, rat prostatic response to androgen status. In addition, insights will be obtained concerning the potential utility of genomic data for dose-response analysis in relation to more traditional toxicological endpoints.

Partnerships/Collaborations: (e.g., where is the data coming from?). The initial modeling utilizes data published in the peer-reviewed literature. Collaboration with Dr. Mitch Rosen (NHEERL/RTD) is under development to obtain prostate genomics data. Collaboration with Dr. Robert Chapin (Pfizer) also is being developed to obtain genomics data following testosterone implantation.

Milestones/Products:

FY06 - Report on the prostate function model following castration.

FY07 - Report on dose-response with testosterone and/or antiandrogens.

FY08 - Evaluate linkage of the biologically-based model of prostate androgen-dependent gene regulation with genomics data.

QA:

QA Plan is being evaluated

IIID-5

Title: Systems Biology Model Development and Application

Lead: Rory Conolly, NCCT (FTE 2.0)

Research Issue and Relevance: System biology models holistically describe, in a quantitative fashion, the relationships between different levels of a biologic system. Relationships between individual components of a system are delineated. System biology models describe how the components of the system interact to give rise to the physiologic function of the system. For the realm of toxicology these models will be developed to not only describe such interactions but to also describe how exposure to toxicants can perturb these interactions and the normal physiology of the system. A hallmark of these models is that they are designed to allow study of the multiple components of the system simultaneously. The resolution of such models depends upon the problem being studied. They can describe interaction between molecules, between molecules and tissue, organs, and whole systems. They can even extend to interaction between different species within ecosystems. Most populations, including humans, are simultaneously exposed to numerous potential toxicants under a myriad of conditions. Further those populations have a variety of other processes occurring during and with those exposures. Underlying disease, nutritional factors, and genetic predisposition are just a few examples of underlying factors that can greatly influence an organism's or population's response to environmental toxicants. System biology models offer the opportunity to describe and understand some of these mechanisms so that risk assessments can eventually be based on the most relevant biological information and not just on default assumptions for which the uncertainty is not easily identified nor quantified. The models are also excellent tools to help analyze data and test different hypotheses. These models will use and depend upon complex data such as is generated from genomics, proteomics, and metabolomics. Iteration between experimental measurements and computational modeling is necessary to understand the function of complicated biologic systems.

Approach: This project will progress along four broad levels each informing and helping develop one another. First, and already on-going, are a number of tasks using existing physiologic pharmacokinetic and pharmacodynamic models to develop and then test different hypotheses describing the adverse affect that may result from environmental exposures. This work is being, at this time, applied to humans. Models describing enzymatic changes, such as cholinesterase inhibition, are being used to show the relative impact of different exposure scenarios. Further, these models are also being developed and used to help design the most useful and cost-effective exposure measurement studies. Collaborative work is being performed to test the suitability of using in-vitro and computationally derived parameters in models such as these. This is another important aspect as given the complex nature of future models and exposure scenarios methods to rapidly estimate key physiologic, thermodynamic, and biochemical parameters will be necessary, especially for those conditions that preclude practical laboratory measurements. In addition, a number of tasks are underway or being planned that will build more complicated pharmacodynamic models for this purpose. (See project description on pharmacodynamic modeling of the prostate as an example.)

Second, a task is being formulated that will begin to build a conceptual model that would used various types of data, such as pharmacokinetic data, mode of action data, "omics" data, etc.

for a specific case. A mathematical model will then be built and implemented. The model will be used to show the importance of relevant data. That is, the question will be answered "how much can the risk assessment be improved and uncertainty reduced as more data demanded and become available?" Some work in this task will begin to develop the mathematical constructs necessary to incorporate "omic" data and information into quantitative models.

The third task will start to describe a fairly complicated endogenous physiologic or biochemical process in detail. Organisms have many biochemical processes that help maintain the homeostasis of the system. Understanding and describing such processes may be crucial to eventually describing and predicting the adverse effects resulting from perturbation of those processes resulting from exposure to exogenous substances and factors. Although this task is still in formative stages examples include the development of kinetic model of the microsomal oxygenation system in hepatoctyes, or describing the glutathione system in physiologic detail. After model formulation, implementation, and testing the model will be expanded for use with pharmacokinetic models to describe what occurs in the endogenous system after exposure to environmental toxicants. Again, such a model can be used for hypothesis testing as well as for predictive risk assessments.

The fourth task will seek to develop a model for a disease process affecting many people. By modeling a disease process with the subsequent pharmacologic exposures a model will be in place which can then be used to investigate the potential of exacerbation by environmental exposures. Again, in the very formative stages one such disease state being considered is type 2 diabetes.

These latter two tasks will develop examples of how endogenous processes, disease states, and exposure to endogenous environmental factors may interact to cause adverse affects or exacerbate pre-existing conditions. It is obvious how all four of these tasks are related on how they build upon each other. Further some over arching themes are important to all of them. For each, eventual application to probabilistic systems will be an anchor. As such, advanced statistical techniques will be used (as developed here, in other computational toxicology projects and elsewhere) to account for variability, sensitivity, and uncertainty. Also, as mentioned, these models are excellent tools for hypothesis testing. That makes them also good tools to keep a balance between describing the detailed complexities of a biologic system and the parsimony that is often practically necessary real world risk assessment process.

Progress to Date: In FY 2005 key staff has come into the center including an ST and two biomedical engineers (one as a post-doc). Further staffing plans are being made. Collaborations (see section below) are being established. Work on the first task and work in related Center modeling projects are well underway. Potential collaborators and staff for the third and fourth tasks have also been identified.

Impact: As described in previous sections, this work will help take risk assessment to the next higher level. The latest mechanistic information and the interplay between exposure, endogenous factors, pre-existing conditions, and genetic predisposition can be rationally accounted for by using such models. There is a great need for quickly illustrating how the latest molecular information, especially "omic" information and information coming from high through-put studies will be used in risk assessments. NCEA and several program offices, including OPPTS, OW, OAR will all be faced with these types of biologic data and currently have few examples of tools for using such information in a beneficial way. This work will be

accomplished over several years will gradually add more and more tools to their arsenal for risk assessment.

Partnerships/Collaborations: National Health and Ecological Effects Laboratory, the National Exposure Research Laboratory, the National Institutes of Environmental Health Sciences, Department of Energy, University of Michigan, Moscow State University (Russia).

Milestones/Products:

FY06 - Journal Article on use of biologic models to ascertain necessary resolution of exposure measurements. Collaborative groups formed for task 2, 3, and 4 Formulation of conceptual model and writing the mathematics and code for that model for task 2, above. Selection and begin model implementation for endogenous biochemical system for task 3. Start model development of disease process for task 4.

FY07 - Implementation of model for tasks 2 and 3 and begin application for cases with exposures to known environmental toxicants – abstracts and presentations. Disease model coded, exercised, and evaluated for task 4. Determine through literature search and other means for examples of exogenous exposure that impact the disease process selected for task 4.

FY08 – Journal articles illustrating some uses of "omics" information in quantitative models. Summary report for Agency use on earliest best practices on "omics" and system biology models. Journal article for disease model. Enhancement of disease model to incorporate exposures to environmental toxicants.

This project will continue beyond this three year period and more products relevant to Agency clients will result.

QA:
QA Plan is being evaluated

Related Tasks:
Systems Biology Modeling

Computational Modeling of Interleukin 1 (IL-1) Mediated Intracellular Signaling
(Collaboration with Jim McDougal, Wright State University)

IL-1 signaling is an important component of the molecular mechanism by which the acute inflammatory reaction in skin develops in response to a dose of skin irritant. The short-term goal of this project is to understand how best to develop computational models of signaling pathways given limitations of the databases describing network topology, protein concentrations, reaction rate constants, etc. The longer term goal is to develop a computational model that can provide useful dose- and time-response predictions of the acute inflammatory response in skin.

Computational Modeling of Extracellular Signal-Regulated Kinase (ERK) Signaling in Cerebellar Granular Cells
(Collaboration with Bill Mundy, NTD, and Qiang Zhang, CIIT Centers for Health Research)

Extracellular signal-regulated kinase is important for regulating neuronal survival and growth. ERK its activation is susceptible to perturbation by environmental chemicals. The objective of this study is to 1) characterize the dynamics of ERK activation in response to BDNF and NMDA; 2) use computational modeling to promote understanding and dissecting of the signaling network underlying ERK activation. This project will improve our understanding of molecular mechanisms which stressors such as BDNF and NMDA perturb neuronal cells, thereby enriching the database needed for identification of predictive biomarkers.

Modeling of Mammalian Biomolecular Responses: Computational Core in support of a Superfund Basic Research Project (SBRP) at Michigan State University
(Collaboration with Norb Kaminski, Michigan State University, and Mel Andersen, CIIT Centers for Health Research)

Starting in the summer of 2006, Dr. Conolly will work closely with a Computational Core housed at the CIIT Centers for Health Research in support of the SBRP at Michigan State University. The initial effort will be to develop computational descriptions of signaling pathways involved in B cell maturation and which are perturbed by exposure to low levels of TCDD.

Mathematical Model of Steroidogenesis: Molecular Response to Endocrine Disruptor Exposures
(Collaboration with Michael Breen, NCCT, Gerald Ankley and Dan Villeneuve, Mid-Continent Ecology Laboratory)

Ankley et al. are using a systems biology approach to characterize the HPG axis in small fish. The project will increase our understanding of the basic biology that is perturbed by endocrine disrupters. A computational model of steroid hormone biosynthesis, starting from cholesterol in support of this effort is being developed. This project is expected to provide rich opportunities for iterative interaction between computational modeling and data collection.

Visualizations of Computational Systems Biology Model Predictions
(Collaboration with Michael Breen, NCCT, and the National Computer Center)

The goal of this project is to develop software tools that facilitate effective interactions between computational modelers and laboratory experimenters. As an initial project, we are developing simulated western blots to visualize model-predicted concentration-time histories. These visualizations will be used to correlate model predictions with literature data typically shown as western blot images.

Risk Assessment Modeling

Dose-response modeling of formaldehyde carcinogenicity
(Collaboration with Fred Miller, Cary, NC)

An updated version of an existing dose-response model for the carcinogenicity of formaldehyde is being developed. This effort focuses on refinements directed at specific issues that were raised after publication of the earlier model.

IIID-6

Title: Use of Toxicogenomics Data in Risk Assessment: Case Study for a Chemical in the Androgen-Mediated Male Reproductive Development Toxicity Pathway

Lead: Susan Euling, NCEA

Research Issue and Relevance: The goal of this project is to address the question, "Can existing toxicogenomics (TG) data improve Environmental Protection Agency (EPA) chemical health or risk assessments?" Although genomics data promises to impact multiple areas of science, medicine, law, and policy, there are only a few areas where genomics data currently has application (e.g., biomarkers of disease). As the technology continues to advance, EPA will need to prepare for genomics data availability and submission by: 1) identifying areas of risk assessment where such data may be particularly useful; 2) developing acceptance criteria for inclusion of toxicogenomics data in risk assessment; and 3) developing approaches for the use of toxicogenomics in risk assessment. These needs will likely require an iterative and collaborative research process between risk assessors and scientists (inside and outside the Agency). At the NCEA sponsored Genomics and Risk Assessment Colloquium in 2003, one of the recommendations was to conduct case studies that could provide a practical attempt to incorporate currently available toxicogenomics data that would illuminate issues and the methods development. This project is responds to this recommendation.

Approach: To address the question of whether TG data can improve health risk assessments, a case study will be performed in which TG data for one chemical will be incorporated qualitatively within the hazard characterization step of a recent or ongoing EPA chemical health or risk assessment. Integrating TG data into an assessment case study will identify areas that may be impacted by TG data and contribute to the development of criteria and approaches for incorporating TG data in assessments. The chemical for the case study will be selected from among those that affect the androgen-mediated male reproductive development toxicity pathway. This pathway was selected because it is well-characterized, there are published TG studies, and a number of genes in the pathway have been identified. Criteria for chemical selection for the case study will include: 1) a relative abundance of available TG data; 2) a recent or ongoing EPA assessment; and 3) an interest by EPA Program and/or Regional Offices. Using the most recent or ongoing (depending on the selected chemical) EPA assessment as a starting point, the team members will conduct an evaluation of the data presented in the assessment, focusing on the hazard characterization and dose response sections. The following questions will be considered: Is the mode of action fully understood for all of the endpoints of concern? Could gene expression information aid in a further understanding of the MOA? Does the TG data provide insights into other aspects of the assessment (e.g., dose-response)? The TG data analysis will then be integrated into the assessment. A report of the case study will be developed and conclusions will be drawn about whether the TG data strengthened or corroborated the risk assessment, qualitatively, and the utility of the approach for incorporating TG data for future risk or health assessments.

Progress to Date: Dibutyl phthalate (DBP) was selected among five candidate chemicals as the chemical for the case study. This decision was based on DBP having 1) the largest and most consistent toxicogenomics database including data indicating gene expression changes in genes known to be in the androgen-mediated male reproductive toxicity pathway, one dose-response gene expression study with low to high dose, and "phenotypic anchoring" (i.e., linkage between *in vivo* alterations and gene expression changes) for some of the gene expression data; 2) an ongoing IRIS assessment that allows our case study to address some broadly applicable questions about the use of toxicogenomics in risk assessment; and 3) an external review draft of the DBP IRIS assessment that our team could utilize. Two subgroups (Data and Approaches Subgroups) were formed to assess the toxicity data and the toxicogenomics data. The Approaches Subgroup has assessed the draft DBP IRIS assessment and associated toxicity studies. The toxicity database of studies with male reproductive and developmental effects has been assembled. The Data Subgroup has assembled the toxicogenomics database from eight publications (PCR and microarray studies) and summarized each of the studies. Possible questions to focus the case study were developed. For each question, its relevancy to the ongoing IRIS assessment and whether the available toxicogenomics data could address the question were considered. A collaboration with the STAR funded Bioinformatics Center at the Robert Wood Johnson Medical School University of Medicine & Dentistry of New Jersey (UMDNJ) Informatics Institute has been initiated. The collaborative project involves performing a genetic pathway analysis of the eight toxicogenomics studies. The results of this analysis may suggest additional pathways affected by DBP. In light of any additional genetic pathways affected by DBP treatment, the in vivo toxicity data will be further assessed. The team has begun writing drafts of chapters for the case study report. In addition to toxicogenomics and toxicity database assessments, our approach to integrating the toxicogenomics data into the DBP assessment is described. This approach may be useful to future health and risk assessments that utilize toxicogenomics data.

Milestones/Products:

FY06: Draft of the case study report (includes scoping exercise complete for case studies; i.e., progress on case study and defining approach for integrating TG data into risk assessment); Discussions with the EPA chemical assessment team, Regions, and Program Offices.

FY07: Conduct EPA Colloquium presenting results and lessons learned from the case study; External peer review draft of report; Submit manuscript on project to peer-reviewed journal.

QA:
QA Plan is being evaluated

IIID-7

Title: Developing Computational Tools for Application of Toxicogenomics to Environmental Regulations and Risk Assessment

Lead: David Dix, NCCT (FTE 1.0)

Research Issue and Relevance: Toxicogenomics is the study of changes in gene expression, protein, and metabolite profiles within cells and tissues, complementary to more traditional toxicological methods. Genomics tools provide detailed molecular data about the underlying biochemical mechanisms of toxicity, and could represent sensitive and precise approaches for detecting effects of exposures, or methods for comparing these effects between species or individuals. Thus genomics, proteomics and metabonomics can provide useful weight-of-evidence data along the source-to-outcome continuum, when appropriate bioinformatic and computational methods are applied towards integrating molecular, chemical and toxicological information. The *Interim Policy on Genomics* (http://www.epa.gov/osa/spc/genomics.htm) recognizes that if genomics is to become useful in regulatory decision-making, risk assessment, and environmental monitoring, the Agency will require the computational methods to handle such data. Measuring changes in gene expression using DNA microarrays has proven useful for identifying biological processes and informing hazard identification and mode of action in toxicological research. Similar microarray data have already arisen in Agency environmental decision-making, and regulatory applications of genomics are likely to increase. EPA's Science Policy Council (SPC) paper on the *Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA* (http://www.epa.gov/osa/genomics.htm) highlights the potential of toxicogenomics in chemical prioritization and risk assessment. To realize this potential, EPA must have the ability for proper analysis and storage, as well as the computational tools to incorporate these types of data into regulatory decisions. Development of these databases and tools, and application of these various toxicogenomic data within Program and Regional Offices will provide EPA staff with valuable, practical training in genomics and associated disciplines. As toxicogenomics grows more important to environmental science and policy, the NCCT will help EPA develop the computational tools and methods to properly evaluate genomics information

Approach: To address the need for development of tools for managing and analyzing toxicogenomics data, the National Center for Computational Toxicology (NCCT) is working across the Office of Research and Development (ORD), the Program and Regional Offices of EPA, and with other Federal and extramural partners. The NCCT is coordinating it's toxicogenomics efforts with the rest of the Agency through the SPC's Genomics Technical Framework and Training Workgroup. This Workgroup has drafted an *Interim Guidance for Microarray-Based Assays: Regulatory and Risk Assessment Applications at EPA*, that recommends continued collaboration with other federal agencies and stakeholders in developing management and analysis tools for genomics data, and the execution of a series of case studies of genomics applications to chemical prioritization or risk assessment. The NCCT intends to follow these recommendations through a series of projects and partnerships within the Agency, and with the FDA and the STAR-funded Environmental Bioninformatic Centers (EBC) in NC and NJ.

First, the NCCT will develop a federated database(s) and analytical tools for the management and analysis of toxicogenomic data for both research and regulatory applications. This project was initiated in FY2006, and is building on the success of FDA's ArrayTrack database. It is NCCT's goal that this effort provides a complete data management solution that addresses requirements unique to scientifically-based risk assessments, confidential and proprietary data security, public access, and other aspects of regulatory application. Consistency, scientific and operational robustness, common access, and availability in a scalable environment are all part of these data management requirements. The expected result is an Agency-wide data management solution integrating genomics, toxicological, and other key data required for both research and regulatory applications. This EPA database containing gene expression profiles and toxicological data for a wide variety of chemicals will facilitate creation of the statistical and computational methods for predictive toxicology. As part of this effort to develop microarray and toxicogenomic analysis tools, the NCCT is continuing to participate in the Microarray Quality Control (MAQC) project. The MAQC is a comprehensive study of microarray quality control and cross-platform comparison, executed by a consortium of many commercial, government (FDA, EPA, NIST, NIH), and academic participants. The MAQC objectives include measuring intra-platform performance; inter-platform comparability; relative accuracy; and concordance of expression measurements to other technologies (e.g. TaqMan PCR).

The second part of NCCT efforts in toxicogenomics are a series of specific, model applications of toxicogenomics data to environmental chemical prioritizations or risk assessments. This includes some of the toxicogenomics elements of the ToxCast program for chemical prioritization, that are generating genomics and metabolomics data from cell cultures which can be loaded into the EPA toxicogenomics database. Also, toxicological data from EPA Program Offices for pesticides and other environmental chemicals will be captured into the toxicogenomic database for use for both chemical prioritization efforts (i.e., ToxCast) as well as risk assessment applications of toxicogenomics. Additional risk assessment applications of toxicogenomics data include ORD-wide work on the conazole fungicides cancer and non-cancer mode(s) of action; pyrethroid pesticides neurotoxicity; perfluorylalkyl acids (PFAA) developmental and hepatotoxicity; the antiandrogenicity of phthalates; the immunotoxicity of diesel particles; and the role of urban air particles in asthma. The NCCT will coordinate across ORD and the Program and Regional Offices, as well as with the EBC in NC and NJ, on how to manage and analyze these datasets in ways that maximize their utilization in risk assessments.

Milestones/Products:

FY06: Completion of participation in the Microarray Qaulity Control (MAQC) with FDA and publication of papers describing best practices- includes description of SPC Genomics Technical Framework. Installation of FDA ArrayTrack database for ORD and Agency use.

FY07: Continued development of ArrayTrack database and analytical tools for toxicogenomics in cooperation across Agency, with FDA, and with the NC and NJ Environmental Bioinformatics Centers. Trial incorporations of toxicogenomic data into risk assessments in collaboration with various Program and Regional Office staff.

FY08: Publication of examples and principles for integrating toxicogenomic data into risk assessments in peer-reviewed scientific journals and contribution of these principles into Agency science policy.

QA:
QA Plan is being evaluate

IIIE-1

Title: Dose-Time-Response Modeling for Evaluating Cumulative Risk of *N*-Methyl Carbamate Pesticides

Lead: R. Woodrow Setzer, NCCT (FTE 0.4)

Research Issue and Relevance: EPA's Office of Prevention, Pesticides, and Toxic Substances (OPPTS) is required by the Food Quality Protection Act to completely reevaluate pesticide registrations by the end of August, 2006. This evaluation must include the evaluation of cumulative and aggregate risk of compounds that act through a common mechanism. One such group of pesticides is the *N*-methyl carbamates, whose common mechanism is the reversible inhibition of acetylcholine esterase (AChE) in nervous tissue. The inhibition of AChE is generally rapidly reversible, with half-lives in rodents at low doses of 1 – 3 hours or so. Accounting for episodic exposures of the carbamates, for example through the diet, then requires a more complex risk assessment model for combining the effects of multiple exposures (for example, an exposure of chemical A at breakfast, followed by an exposure of chemical B at lunch), than it would for exposures of compounds with longer-lived common effects. It also requires the estimation of parameters for models that describe both the dose-response and the recovery of AChE activity from multiple studies of the same dose response relationship.

Approach: The approach to this problem is to adapt the dose-response model already used for modeling AChE activity after organophosphate exposure to include a model for recovery; to use a hierarchical model to account for the variability among studies, so that all the datasets for a chemical can be used to estimate dose-response parameters; to develop a model that accounts for AChE inhibition subsequent to multiple non-simultaneous exposures, using the parameter estimates derived from dose-response modeling.

Progress to Date: A dose-time-response model has been developed that can be used both for modeling individual chemical dose-time-response data and for predicting AChE inhibition; preliminary parameter estimates have been generated.

Impact: The dose-response results of this work will be used in the dose-response assessment portion of the Agency's *N*-methyl carbamate cumulative risk assessment. Predictions of AChE inhibition under realistic exposure scenarios may be used in the hazard characterization portion of that risk assessment. Methods developed in this analysis could be used in the future for cumulative risk assessments for agents with ephemeral acute effects.

Partnerships/Collaborations: This work is being done in collaboration with scientists in the Health Effects Division of OPPTS and from the Neurotoxicology Division of the National Health and Environmental Effects Research Laboratory.

Milestones/Products:

FY06 – Science Advisory Panel review of preliminary cumulative risk assessment, including dose-response modeling.

FY07 – Release and Science Advisory Panel review of revised cumulative risk assessment.

FY08 – Submission of methods and dose-response models for publication in the peer-reviewed literature. Incorporation of dose-time-response methodology into the Agency's BMDS software.

QA:
QA Plan is being evaluated

IIIE-2

Title: Application of Visual Analytic Tools to Evaluate Complex Relationships between Environmental Factors and Health Outcomes

Lead: Elaine Cohen Hubal, NCCT (FTE 1.0)

Research Issue and Relevance: Characterizing cumulative risk and understanding the complex relationships between environmental exposures and human health outcomes requires collection and analysis of a wide range of data. Information on the characteristics of multiple stressors (chemical, physical, biological and psychosocial), the characteristics of the human receptor (genetics, health status, life stage, behaviors, social factors, etc.) at multiple levels of organization (individual, community, population), and the temporal and spatial patterns of exposures and outcomes must be combined to assess cumulative risk.

Current approaches for designing studies and evaluating collected data tend to focus on a limited number of environmental factors and/or measures of outcome. As such, it is difficult to understand the potential impacts of the full range of factors on environmental health. More holistic approaches for interrogating this multidimensional data space are required to identify potentially important relationships for further study and analysis. Application of emerging computational tools will allow us to optimize utility of collected data, improve understanding of complex exposure-outcome systems, and improve risk assessment.

Approach: In this project, we plan to use engineering principles to develop conceptual and mathematical models of the human-receptor, source-to-outcome system, and visual analytic tools to address the significant challenges associate with characterizing cumulative risks. First, we will apply a systems approach to develop a human-receptor based conceptual framework. We plan to adapt the strategy for creating conceptual models for complex ecological risk assessments presented by Suter (1999).

We will also identify and evaluate visual analytic tools required to address analysis needs for characterizing multi-factorial relationships between environmental factors and human health outcomes. Visual analytics is a new branch of visualization that merges scientific and information visualization and includes technologies from other fields, including information extraction, knowledge management, and statistical analysis. Visual analytics tools can be developed and applied to represent complex multidimensional data. Large, dynamic and complex data sets containing text, measurements, and images, can be effectively combined to reveal significant relationships and trends and to enhance discovery. Visual analytics can be used for outcome analysis and visualization, to find patterns and subtle relationships in data, and to infer rules that allow predictive analysis to prevent and mitigate environmental disease.

Finally, we will test and demonstrate VA tools using existing available data. We propose doing this demonstration using data from children's cohort studies to explore the potential of visual analytics to facilitate evaluation of the effects of environmental exposures on child health and development. Another possible demonstration could involve using data collected for a community-based cumulative risk assessment.

Progress to Date:

To date some preliminary (and very conceptual) research has been conducted to consider how biomonitoring data can be used to characterize cumulative risk and how psychosocial factors can be incorporated into cumulative risk assessments. In addition, a novel data mining tool is being tested and visual analytic software packages are being evaluated using extant children's exposure data to identify relationships between exposure factors and body burden for a selected set of chemical agents.

Impact:

Across EPA, program and regional offices are being called on to assess cumulative risk resulting from real-world exposures. The Agency is also being required to identify vulnerable populations, characterize life-stage risks, and evaluate gene-environment interactions. Multi-factorial analyses of one form or another are required: to conduct national-scale regulatory-based risk assessments (program offices); to conduct community-based risk screening and remediation (regions and states); to support epidemiology studies investigating gene-environment interactions (interagency); and to characterize exposure and risk for public health tracking (all of the above). The approaches and tools developed through this research will help the Agency meet the increasingly complex needs for cumulative risk assessment. In addition, results of this effort may be used to develop concepts and tools for application to the Detroit Children's Study, the North Carolina Cohort, and the National Children's Study.

Partnerships/Collaborations:

In this project, the NCCT will take advantage of visual analytic capabilities that are being developed in the Scientific Visualization Center at the EPA National Environmental Scientific Computing Center. In our initial evaluation of Visual Analytic software packages, we will build on statistical analysis of children's exposure data conducted in NERL. We also hope to collaborate with investigators conducting children's cohort studies to explore the potential of visual analytics to facilitate evaluation of the effects of environmental exposures on child health and development. As a first step in developing these collaborations, we presented this research concept during a breakout session on Computational Toxicology at the 2005 Collaborations for Children's Environmental Health Research Workshop (part of the EPA/NIEHS Children's Centers Scientist-to-Scientist Workshop, July 11, RTP). Finally, we are exploring a potential collaboration with PNNL's National Visualization and Analytics Center.

Milestones/Products:

FY06 - Demonstration of potential for application of VA to evaluate exposure data. Workshop on applying VA to analyze children's cohort data.

FY07 - Generic conceptual model of complex relationships between environmental factors and human health outcomes.

FY08 - Demonstration of application of VA to analyze children's cohort data.

QA:

QA Plan is being evaluated and will be inserted in the next version.

MEMORANDUM OF AGREEMENT
BETWEEN THE
NATIONAL CENTER FOR COMPUTATIONAL TOXICOLOGY, NATIONAL HEALTH AND ENVIRONMENAL EFFECTS RESEARCH LABORATORY, AND NATIONAL EXPOSURE RESEARCH LABORATORY

Background

On October 7, 2004, the Office of Research and Development (ORD) announced the establishment of the EPA National Center for Computational Toxicology (NCCT) headquartered in Research Triangle Park, NC. The Center will provide scientific expertise and leadership related to the application of mathematical and computational tools and models to high priority Agency needs. These needs include improving the Agency's data reporting requirements for environmental fate and transport and for toxicity testing, setting priorities for the acquisition of those data based upon predictive models, and for understanding toxicity and risks posed by environmental agents. The tools and models are derived from modern technological advances in the general areas of computational chemistry, mathematical and systems biology, and similar systems.

The reorganization package for the NCCT staff identified nineteen full time positions staff, including 11 who were initially detailed from other ORD National Labs/Centers/Offices. After implementation of the reorganization, these individuals will be permanently reassigned to the NCCT until the proposed sunset date in October 2009. Although two administrative positions exist in the NCCT, many of administrative and research support functions will be provided by the Program Operation Staffs of the National Health and Environmental Effects Research Laboratory (NNHERL) and the National Exposure Research Laboratory (NERL). This MOA outlines the administrative support services that will be the primary responsibility of the NCCT and those that NHEERL and NERL will provide to the NCCT. NCCT management is responsible for the decisions and actions implemented by the NCCT. Since good communication will be an essential element in the success of these interactions, the parties agree to bi-lateral discussions at least bi-monthly with the senior managements of NERL and NHEERL to discuss implementation issues and resolve any ambiguities or issues that arise as the relationships between the organizations develop and mature. This agreement is effective as of January 31, 2005 and will remain in effect until amended by mutual agreement of the participants.

Purpose of the Memorandum of Agreement and Services Provided

The delegation of authority for ORD's NCCT exists under the ORD Policies and Procedures Manual, Chapter 7.4 dated March 11, 2005 and this agreement refers to the administrative support that will be supplied by NHEERL and NERL. This support includes but is not limited to: Funds Control, Freedom of Information Act requests, Training Agreements, Senior

Environmental Employment Program, Technical Qualifications Board, Buildings and Facilities, Capital Equipment. Extramural Management. Technology Transfer, Contractual Support. Health and Safety, and Quality Assurance.

Administrative Support Services that will be primary responsibilities of the NCCT:

1) Budget

- Monitor funds utilization, prepare budget reports and plans, and prepare justifications for budget initiatives
- Create WCF Service Agreements and Funding Levels
- Prepare FTE/PC&B Projections and Workforce Management/Strategies
- Initiate reprogrammings to transfer funds to other organizations or across budget structure dimensions
- Create and maintain IRMS Execution Implementation Plans
- Develop NCCT budget narratives for OMB and President's budget development (with ORMA)

2) Records Management

- Provide new and/or updated NARA/Agency/ORD guidance
- Represent the NCCT's interests with the Project Officer on Center file inventories, records awareness week, Center training needs, negotiation/funding of records management contractor
- Assist the NCCT staff with interpretation of records guidance. schedules. etc.

3) Timekeeping

4) Travel authorizations, itinerary planning, and preparation and submission of Travel Vouchers: purchase card holder and authorizing official

5) Ethics

- The NCCT Director will serve as the Deputy Ethics Official (DEO).
- The NCCT Program/ Management Analyst will serve as the DEO Assistant.

6) Human Resources

Manage all Human Resources activities, including performance. orientation, awards. training. career development. leadership development, post doc program. form preparation, and HR liaison with OARM Human Resources Management Division.

7) Federal Managers' Financial Integrity Act (FMFIA)

- Develop and submit NCCT's Mid-year and the Annual Assurance Letter
- Represent NCCT on ORD workgroups related to Management Integrity
- Keep Center up to date on new and/or changed guidance, policies and/or procedures

8) Management Reviews

- Conduct Management Reviews of the Center
- Prepare review reports
- Develop new protocols, review guides, etc. if needed due to the nature of work being conducted by the Center

9) Extramural Instruments

- Prepare procurement documents including those for contracts, interagency agreements and cooperative agreements.
- Work with Office and Science Policy on Technology Transfer needs.

Administrative Support Services provided to the Center by NHEERL

1) Funds Control Officer

- NHEERL will provide funds certification and data entry into appropriate ORD and Agency systems. NHEERL commits to provide this function for the duration of FY 2005 and will evaluate the impact of the workload associated with this function for future FY's. If the workload is determined to be more than the FCO can manage comfortably, NHEERL will notify NCCT management with sufficient notice to make other arrangements for this support.

2) Electronic Purchase Card System

- NHEERL will allow NCCT to utilize its lotus notes-based electronic purchase card system. If NCCT chooses to use the system, it will provide funding to support its use as well as maintenance of this system.

3) Freedom of Information Act requests (FOIA)

- NHEERL will coordinate and prepare responses to FOIA requests in accordance with FOIA regulations. The Center will gather the responsive information, prepare it for submission, and submit it to NHEERL.

4) NRC and other Training Agreements

- NHEERL will allow the Center access to the training agreements with local

universities and its agreement with the National Research Council.

- NCCT will be included in calls for needs associate with those agreements and will work with NHEERL PO's and coordinators.
- NCCT will be responsible for funding any individuals being trained by its staff through those agreements. An agreement will be reached on the level of funding the Center will place on particular agreements.

5) Senior Environmental Employment Program (SEEP)

NHEERL's SEEP Coordinator will:

- Coordinate recruitment packages (Position Description Form, Requisition Form, Cost Analysis Worksheet, Commitment Notice and IFMS screen, and memo) for processing and forward to selected grantee for recruitment.
- Request Commitment Notices for renewals of SEEP enrollees.
- Provide Center with new/revised guidance, policies and/or procedures regarding the program.
- Forward the quarterly report to the appropriate SEEP Monitor(s) for reconciliation. The Coordinator will provide information on the reconciliation procedures.

6) Technical Qualifications Board (TQB)

- NHEERL will include NCCT in calls for candidates requiring review through the TQB process and will provide support for scheduling and holding the TQB review meetings.
- NCCT will provide funding for the professional services contracts for the ad hoc reviewers used for the review of NCCT's candidates.

7) Building & Facilities Support

- The NHEERL Facilities Coordinator will provide support for B&F and R&I processes as well as the ORD Capital Equipment process.
- This includes updating ORD and Agency systems and databases related to space and facilities.
- This includes liaison with OARM on space issues and assistance in facilities design and renovation.
- NCCT funding will be allocated to RC-6 (NCCT) and NCCT will provide the appropriate funding for its projects.

8) Graphics, Web Development and Photography Support

- Support for graphics, photography, and web development and maintenance will be provided in coordination with NHEERL.
- NCCT will have access to NHEERL support contracts, or NHEERL will assist in referring NCCT to appropriate available contracts.
- NCCT will provide funding for the contractual support that they require.

Administrative Support Services provided to the Center by NERL:

1) Extramural Management

- The appropriate NERL EMSers (Extramural Management Specialists) will provide advice and guidance on appropriate extramural vehicles: keep NCCT current on requirements pertaining to extramural policies and procedures; and include NCCT in calls and due dates related to extramural vehicles and procurement.
- The appropriate NERL EMSers will review all NCCT packages that require EMS review and/or concurrence.

2) Contractual Support Mechanisms

- NERL will allow the Center access to various technical support contracts that it maintains provided the requested effort is within the Statement of Work (SOW) and that the Center covers all costs associated with that use.

3) Health and Safety

- NERL's Health and Safety Team will support the Center.

4) Scientific Quality Assurance (QA)

- NERL's QA Director will ensure that all Center research projects comply with Agency QA requirements
-

5) Information Management (IM)
NERL will provide support in the area of IM to NCCT.

Administrative Support Services provided to the Center by Consolidated Support

The following Administrative Support Services are currently provided to all Laboratories and Centers at RTP by consolidated services. If there is a change to how the support in these areas is provided, support for the NCCT will be addressed at that time.

1) Public Affairs and Scientific Communications Support

- The ORD Communications Team provides support for media. events, and local congressional coordination. The Team also provides the interface with ORD and EPA Communications staff.
- NCCT will provide funding on an allocated basis to support the communications effort.

2) Information Technology

The ORD Consolidated Center's information technology (IT) staff provide a wide range of IT, and information security (IS) to the labs and centers at RTP. General IT and IS support for NCCT will be provided by the Consolidated Center. NCCT will provide funding for IT support for its full-time employees. IT support for NCCT part-time employees will be provided by the "home" laboratory.

3) RTP Shared Cost Budget

The RTP Shared Cost Budget covers services that all of RTP uses. For the Center, these services include the EPA Wellness Center, special emphasis program support, Management Council initiatives, and supplies for the TCO.

The RTP Management Council will agree on a formula-based allocation. The Center will receive a breakout for their portion of the total shared cost budget.

CONCURRENCES:

Jerry Blancato, Deputy Director
National Center for Computational Toxicology

6/21/2005
Date

Jewel Morris, Deputy Director for Management
National Exposure Research Laboratory

6/28/05
Date

John Jones, Deputy Director for Management
National Health and Environmental Effects Research Laboratory

6/29/05
Date

EPA Communities of Practice:
Chemoinformatics

The Chemoinformatics workgroup is proposed to be formed in Summer 2005 to facilitate, coordinate and integrate efforts to address the challenges of chemical structure annotation (or indexing), retrieval, and mining of chemically-related data and documents, including newer toxicogenomics and metabonomics data, across EPA Program Offices, Labs and Centers. Much of EPA's public and internal chemical data records and databases currently are indexed only by chemical name (imprecise and non-unique) and CAS registry numbers (proprietary) and are not searchable by the more universal and informative metric of chemical structure. Additionally, there is unnecessary duplication of efforts and lack of coordination in the area of chemical structure-annotation and quality review of chemical information (e.g., structures, CAS) across diverse EPA databases. Finally, no Agency-wide chemical structure searching capability currently exists to enable EPA scientists, regulators, and outside parties to efficiently and precisely locate Agency chemical information based on chemical structure.

The Chemoinformatics workgroup would invite members of various Offices, Labs and centers currently involved in either chemical database construction or use, to coordinate ongoing efforts in the area of chemical-structure annotation and structure-searchability of EPA records and data, and to help chart a path forward for expanded Agency-wide adoption of capabilities in these areas. One of the first tasks of this workgroup would be to consider and refine the elements of the draft charter below, and create a finalized charter for the workgroup.

A draft charter of this Chemoinformatics workgroup is suggested to include the following functions:

- Facilitate communication and co-ordination among the Agency personnel who are directly using, developing, or managing chemical information data records, including those associated with newer toxicogenomics and metabonomics technologies, particularly in ORD, OPP and OPPT whose duties involve chemical searching and structure-activity assessments.
- Propose adoption of application-independent standards for chemical structure representation (including information pertaining to the characterization of mixtures, racemic chemicals, polymers, and stereochemistry) that will facilitate the broadest possible utility and compatibility of such information, both within and outside of EPA.
- Guide the creation of a consolidated EPA database of chemical identification information (structure, CAS, name), obtained from public sources (chemfinder.com, PubChem, NLM ChemID) but having undergone additional levels of quality consistency review, to serve as a common resource for the structural annotation of EPA records or databases. *[Note that physical/chemical properties can be measured or derived from the accurate chemical structures by a variety of algorithms and means; hence, such property annotation of chemical*

information resources will be considered as a separate and distinct task of chemical annotation not included in the initial charter of this workgroup.]

- Create procedures and guidelines for chemical annotation of Agency data and records, as an adjunct to EPA Information Quality Guidelines, to improve consistency, quality, and efficiency related to storage, retrieval and mining of chemical information within the Agency.
- Exchange information and experiences regarding available public and commercial approaches and initiatives in the areas of chemical structure representation, chemical searching, and managing and modeling of chemical data.
- Evaluate and recommend which of the available approaches and initiatives might be adopted or coordinated to best serve EPA's long-term needs in this area, and serve as an information resource to other Agency personnel on these matters.
- Evaluate and provide guidance concerning the needs and requirements of an Agency-wide chemical structure-searching capability, possibly from publicly available or open-source parties, that would offer low-cost, flexible solutions to enable EPA scientists or others to locate records or data pertaining to chemical information on EPA's intranet or internet websites.
- Facilitate the development of training materials and programs for scientists, chemical managers and risk assessors that assist in the implementation and wider use of these chemical information resources and technologies.
- Report regularly (every six months) to the ORD science council on issues, progress, and outputs

Office/Division/Branch managers across various research offices, including ORD, OPP and OPPT, will be invited to appoint representatives to this workgroup who would benefit most from, and contribute most to these discussions, as well as who would serve as a conduit of information to and from their respective Branch/Division/Office.

EPA Community of Practice:
Biological Modeling Working Group (BMWG)

The EPA Biological Modeling Working Group is proposed to be formed in Summer 2005 to advance the principals for development and application of dosimetry and other biologically based models within the Agency. Dosimetry modeling includes multiple forms of toxicokinetic modeling (e.g., physiologically based toxicokinetic (PBTK) modeling, compartmental modeling), respiratory tract dosimetry modeling (e.g., computational fluid dynamics), and related modeling (e.g., dermal absorption modeling). The working group will also focus on biologically based response modeling with special emphasis on using the newest "omics" information in biologically based models . The goal of this work group is to foster adoption of modeling science by Agency clients in regulatory decision making. A cross-ORD group that also has representation from outside of ORD can carry great influence throughout the Agency, helping assure that ORD efforts are viewed as relevant and important to the Agency mission.

The work group would include representatives of the ORD Labs and Centers with expertise in this area and have a representative from the Risk Forum. The BMWG would draw on the expertise of other scientists within ORD and the Agency as needed for specific issues. The BMWG would ask the Risk Forum to provide broader Agency-wide comment and review on specific matters as it determined necessary. Committee members would serve for a period of two years, following which the lab/center would renew their membership or appoint a new member.

The charter of the BMWG includes the following functions:

- Facilitate communication and co-ordination among the ORD biological modelers to foster their role as the major technical resource for the Agency in this area.
- Develop consensus positions around issues for how analyses are conducted using models to provide consistency (e.g., processes for application of uncertainty factors in non-cancer evaluation, methods for using models to describe population distributions, durations over which AUC is calculated)
- Develop consensus positions around issues of model evaluation and documentation to provide assistance to model developers inside and outside the Agency.
- Develop recommendations for research needs to be addressed by ORD labs and centers (e.g., methods for statistically evaluating alternative model structures and characterizing model output uncertainties)
- Facilitate the development of training materials and programs for chemical managers and risk assessors that assist in the implementation of these technologies.
- As the science advances develop recommendations and guidance on the use of omics information to enhance dosimetry and pharmacodynamic models
- Facilitate communication to the larger modeling and scientific on issues for modeling and applications to risk assessment (e.g., SOT symposia, specialty meetings)

.

The work group will:

- Hold meetings: quarterly to identify issues/projects to be taken on and discuss progress
- Organize internal meeting/conference calls around specific issues/projects with goal of identifying state of the science, research needs, consensus positions.
- Prepare publications for peer-reviewed literature on specific issues/topics.
- Identify research needs that could be addressed internally or through extramural contracts or grants.
- Organize outreach to broader modeling and scientific community to demonstrate EPA's interest and obtain input and participation
- Report regularly (every six months) to the ORD science council on issues, progress, and outputs
- The group's representative from the Risk Assessment Forum will request time on the Forum's Agenda to brief regular Forum members of issues, progress, and outputs.

November 28, 2005

EPA Categorization and Prioritization Community of Practice

The EPA Categorization and Prioritization Community of Practice (CPCP) formed in December 2005 to advance research into the utility of computational chemistry, high-throughput screening (HTS) and various toxicogenomic technologies for Agency use. Modern computational chemistry and molecular biology technologies can provide information about the physical and biological properties of large numbers of chemicals. The goal of the CPCP is to advise on the development of a research program within the National Center for Computational Toxicology (NCCT) generating data from these technologies and interpreting it in order to categorize chemicals and predict toxicity. If proven accurate, these toxicity predictions could then be used for prioritization of limited testing resources towards chemicals and endpoints that present the greatest risk to human health and the environment.

The primary function of the CPCP is to advise on the NCCT's planning, conduct and interpretation of a chemical categorization and prioritization research project entitled "ToxCast". The ToxCast project will be designed to have the ability to predict, or forecast toxicity. Initially, this would entail a demonstration project based upon a set of chemicals with a rich toxicological database (e.g., registered pesticides, or the chemicals tested in the NTP bioassay program). This set of 200 or more chemicals would represent a number of differing structural classes and phenotypic outcomes (e.g., tumorigens, developmental and reproductive toxicants, neurotoxicants, immunotoxicants). The ToxCast project would evaluate chemical properties and effects across a broad spectrum of information domains: physical-chemical properties, predicted biological activities based on existing structure-activity models, biochemical properties based on HTS assays, cell based phenotypic assays, and genomic analysis of cells or organisms. The ultimate goal of the ToxCast project would be to mine the resulting data for associations between and among the various domains and the known toxicological properties of the base set of chemicals, in order to provide a structured strategy to categorize chemicals, identify potential toxicities and pathways, and to prioritize chemicals for subsequent testing based on that information.

The Community would include individuals from the Office of Research and Development (ORD) Labs and Centers, as well as scientists outside of EPA with expertise in HTS, toxicogenomics, predictive toxicology or bioinformatics. The CPCP would be chaired by a member of the NCCT, but draw on the expertise of scientists across ORD and the Agency as needed for specific issues. The Community would ask ORD's Computational Toxicology Implementation Steering Committee (CTISC) to provide broader Agency-wide comment and review on specific matters as necessary. Community members would serve for the period required for planning and execution of the ToxCast demonstration project. It is expected that the CPCP would work closely with the Chemoinformatics Community of Practice (chaired by Ann Richard) on issues relevant to both groups.

The charter of the CPCP includes the following functions:

1. Organize periodic meetings or conference calls (at least quarterly) around specific issues or projects, with goal of identifying state of the science, research needs, and consensus positions.
2. Organize outreach to broader HTS and toxicogenomics scientific community to demonstrate EPA's interest and obtain input and participation.
3. Report regularly to the Director of the NCCT and the ORD CTISC on issues, progress, and outputs.

Specific goals of the CPCP are to advise on the following:

1. Development of key partnerships and collaborations with external groups that can facilitate development of the information needed in ToxCast. These groups include the Office of Air and Radiation (OAR), Office of Prevention Pesticides and Toxic Substances (OPPTS), Office of Water (OW); the NTP/NIEHS and NIH/MLI; the ACC, CropLife, EDF, and other external groups to help develop a consensus on the specific directions and contents of ToxCast.
2. Identification of a set of chemicals for the ToxCast demonstration project.
3. Selection of data domains and specific assays based upon pre-existing knowledge and within the available resources.
4. Selection of key target toxicities for initial focus of the ToxCast proof of concept.
5. The impact of metabolizing capability, or lack thereof, on the efficiency of the screening assays.
6. Development of a bioinformatic approach to mining the resulting data and identifying signatures of concern.
7. Reporting the utility of assay results and analysis techniques to categorize pilot chemicals according to known toxicity patterns; revise methods and approaches as dictated by results.
8. Expanding the ToxCast project beyond proof of concept, and carrying out a prospective assessment of the approach using chemicals currently entering a traditional testing process.

Impact: The availability of a biologically and chemically based system (ToxCast) to categorize chemicals of like properties and activities will provide a number of EPA Program Offices with an extremely useful tool that heretofore has been seriously lacking. ToxCast may be one of the first broad-scale products of the NCCT that addresses the mission of improving the efficiency and effectiveness of hazard identification and risk assessment methodologies employed by the EPA.