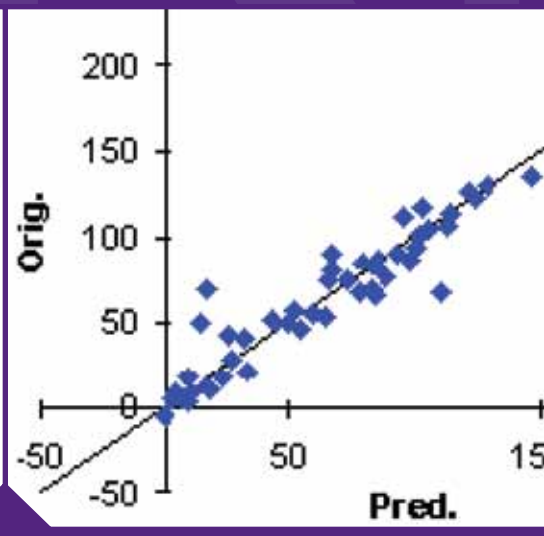
EPA
United States
Environmental Protection
Agency

QSAR/VFAR Workshop Summary

REPORT



Office of Research and Development
National Homeland Security Research Center

August 2007 EPA/600/R-07/095

QSAR/VFAR Workshop Summary Report

By
Smita Siddhanti and Caroline Baier-Anderson
EnDyna, Inc.
McLean, VA 22102

Under subcontract to:
Science Applications International Corporation
11251 Roger Bacon Drive
Reston, VA 20190

EPA Contract No. 68-C-02-067

Project Officers

Chandrika Moudgal
National Homeland Security Research Center
U.S. Environmental Protection Agency
Cincinnati, OH 45268

Douglas Young
National Risk Management Research Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

Office of Research and Development
National Homeland Security Research Center, Threat and Consequence and Assessment Division

Disclaimer

The opinions expressed within the workshop summaries do not necessarily represent the views of the EPA. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Acknowledgements

This report was prepared for the U.S. Environmental Protection Agency, Office of Research and Development, National Homeland Security Research Center and National Risk Management Research Laboratory. The primary authors were Dr. Caroline Baier-Anderson and Dr. Smita Siddhanti of EnDyna, Inc., under subcontract to Science Applications International Corporation (SAIC), Contract No. 68-C-02-067. Lisa Kulujian was SAIC's Work Assignment Manager. Chandrika Moudgal and Douglas Young of EPA's National Homeland Security Research Center and National Risk Management Laboratory, respectively, served as EPA's project officers.

Table of Contents

Acronyms

BMD: Benchmark Dose

BMDL: Benchmark Dose Lower Confidence Limit

BMC: Benchmark Concentration

BMCL: Benchmark Concentration Lower Confidence Limit

CCL: Contaminant Candidate List

CAMRA: Center for Advancing Microbial Risk Assessment

CTC: Computational Toxicology Center

DORIAN: Dose-Response Information Analysis System

EC_{50}: Median Effective Concentration

ECB: European Chemicals Bureau

EPA: Environmental Protection Agency

ER: Endocrine Receptor

EU: European Union

FDA: Food and Drug Administration

FIGUR: Fast Identification of Genomic Unique Regions

GPCRs: G Protein-Coupled Receptors

GSH: Glutathione

LOAEL: Lowest Observed Adverse Effect Level

LD_{50}: Median Lethal Dose

LC_{50}: Median Lethal Concentration

MOA: Mode of Action

NERL: National Exposure Research Laboratory

NCEA: National Center for Environmental Assessment

NHSRC: National Homeland Security Research Center

NOAEL: No Observed Adverse Effect Level

NRMRL: National Risk Management Research Laboratory

OECD: Organization of Economic Cooperation and Development

OPPTS: Office of Pollution Prevention and Toxic Substances

OSWER: Office of Solid Waste and Emergency Response

PBPK: Physiologically Based Pharmacokinetic

PCR: Polymerase Chain Reaction

PD: Pharmacodynamics

PDB: Protein Data Bank

PK: Pharmacokinetics

PNN: Polynomial Neural Network

QSAR: Quantitative Structure-Activity Relationship

RE: Risk Estimation

REACH: Registration, Evaluation, and Authorization of Chemicals

RfC: Reference Concentration

RfD: Reference Dose

SDWA: Safe Drinking Water Act, as amended in 1996

TCAD: Threat and Consequence Assessment Division

TERA: Toxicology Excellence for Risk Assessment

TOXHUM: The Human Toxicology Research Group, Université de Montréal

UMDNJ: University of Medicine & Dentistry of New Jersey

VF: Virulence Factor

VFAR: Virulence Factor-Activity Relationship

VHTS: Virtual High-Throughput Screening

VMG: Virulence and Marker Genes

WOE: Weight of Evidence

Executive Summary

Introduction

The U.S. Environmental Protection Agency's (EPA's) National Homeland Security Research Center (NHSRC) and National Risk Management Research Laboratory (NRMRL) conducted a QSAR/VFAR Workshop, on June 20 – 21, 2006 in Cincinnati, OH. The workshop's main purpose was to explore the application of Quantitative Structure-Activity Relationship (QSAR) and Virulence Factor-Activity Relationship (VFAR) concepts to the risk assessment process in situations where chemical- or biological-specific empirical data are either inadequate or lacking.

The mission of both NHSRC and NRMRL is related to assessment of public health and environmental risk from harmful chemicals. Parts of the Office of Research and Development (ORD), NHSRC and NRMRL manage and support a variety of research and technical assistance efforts. NHSRC focuses on enhancing the ability to detect, contain, mitigate the effects of, and clean up after significant emergency events, terrorist attacks, or natural disasters. NHSRC scientists and engineers seek to identify or develop affordable, effective technologies and methods for addressing the risks posed by chemical, biological, and radiological agents. NRMRL's mission is to develop ways to prevent and reduce pollution of air, land, and water. This mission plays a critical role in EPA's goal of achieving sustainability; several methodologies have been developed within NRMRL to quantify the potential environmental harm of chemical releases.

The overarching goal of this workshop was to evaluate the potential uses of QSAR and VFAR to advance the rapid and efficient evaluation of chemicals and microbes of potential concern. To achieve this goal, the QSAR/VFAR workshop convened toxicologists, microbiologists, chemists, engineers, biostatisticians, pharmacologists, biochemists, and risk assessment scientists to discuss the state of the science, opportunities for advancement, and practical applications. Expert panel members included researchers with expertise ranging from microbial genomics to computational toxicology and risk assessment. The workshop also included EPA scientists with expertise in the development and application of QSAR and VFAR. To facilitate discussion at the workshop, a list of charge questions was made available to the expert panel and the workshop participants.

Specific Workshop Goals

NRMRL

NRMRL will use the outcome of the QSAR/VFAR Workshop to inform its research in developing QSARs in a number of ways: validating the importance of QSAR research, providing guidance for QSAR development, and providing a vision for the future role of QSAR in a regulatory context. The diverse group of participants and panel members that attended the workshop, which included researchers from EPA Program Offices, the European Unions' (EU) Commission involved with the Registration, Evaluation, and Authorisation of Chemicals (REACH), other federal agencies, nongovernmental organizations, industry, and academia, validate the importance of continuing with QSAR research. Based on discussions at the workshop, it is apparent that QSAR has an important role in the future of chemical regulation and industry both here in the U.S. and in the EU. With development of new technological areas, such as bioinformatics (which includes genomics, proteomics, and metabonomics), there has been some question as to whether or not QSAR research has a useful future.

NHSRC

NHSRC will use the outcome of the QSAR/VFAR workshop to streamline its current QSAR research and to initiate its VFAR research. Currently, the Center is developing several QSAR models for predicting acute, subacute, and subchronic benchmarks to address exposure durations that are of key importance during an emergency event, terrorist attack, or natural disaster. The VFAR method is being explored to determine the hazards associated with exposure to highly potent pathogens. Since little is known about the VFAR methodology, this workshop will allow the Center to define VFAR and to assess the state of the science. To aid the discussion process, the Center, in collaboration with NRMRL, developed a set of strategically and technically sound charge questions aimed at key aspects of QSAR and VFAR methods. The discussions on these charge questions will set the stage for future QSAR and VFAR research at NHSRC and NRMRL.

Background

It has long been recognized that chemical substances with sufficiently similar structures and chemical activities exert similar qualitative toxicities with differing magnitudes (Ashby and Tennant 1988; DHHS 1980; Gray and Ostby 1993; Harada et al. 1992; Lewis et al. 1993; Lowell et al. 1989; Rosenkranz and Klopman 1989; Weisburger 1979; Weisburger and Fiala 1979). Thus, analysis of molecular structures and physicochemical properties of chemical substances can provide a rapid means of predicting and quantifying the toxicity of minimally tested chemicals. This fundamental observation is the basis for the qualitative structure-activity relationship (QSAR) method of toxicity analysis. The QSAR method of toxicity analysis assumes that a sufficiently strong structure-activity relationship of chemical substances is indicative of qualitative and quantitative similarity in toxicity (EPA 1994, 1992, 1989). Consequently, the long-term toxicities of minimally tested congeners of a chemical series or type can be estimated from those of better-known congeners on the basis of available information (EPA 1989, 1992; Rosenkranz and Klopman 1989; Weisburger and Fiala 1978). Hence, the QSAR method provides a means by which the toxicity of a candidate chemical substance, for which adequate toxicity

data for risk assessment is not available, can be reasonable inferred from those of a toxicologically better-known structurally and chemically related surrogate chemical substance or congener. This method of analysis is intended to establish a qualitative and quantitative association between the structure-activity and toxicity of a candidate chemical substance and that of the surrogate chemical substance to enable quantitative estimates of the toxicity of the minimally tested candidate chemical substance.

QSAR estimates allow for the prioritization of such chemical substances for more costly and time-consuming toxicological testing or for setting cleanup or media health-based limits in the regulation of minimally tested chemical substances. This method of toxicity assessment is especially useful for many environmental toxicants for which there is a critical lack of adequate toxicity and pharmacokinetics data for risk assessment. The use of the QSAR method for the evaluation and establishment of interim or provisional toxicity references for chemical substances for which there are inadequate toxicity data for use in risk assessment is an EPA-approved practice and the basis of the data found in the EPA's Assessment Tools for the Evaluation of Risk (ASTER) (EPA 1994). This method has also been used by EPA and internationally (under the auspices of the North Atlantic Treaty Organization's Committee on Challenges of Modern Society [NATO/CCMS]) to develop toxicity equivalency factors (TEF) for chlorinated dibenzo-*p*-dioxins and -dibenzofurans (CDDs and CDFs) (EPA-TEF/87, I-TEF/89), wherein the toxicity of 2,3,7,8-tetrachloro-dibenzo-*p*-dioxin (2,3,7,8-TCDD) is used as the central reference with which the toxicity of all the other CDDs and CDFs are qualitatively and quantitatively related (EPA 1989). EPA has proposed a similar QSAR approach for polycyclic aromatic hydrocarbons (PAHs) for which there is inadequate toxicity information as described by EPA (1992).

With thousands of chemicals representing potential environmental contaminants, the need for a framework of effective prioritization for regulatory development and risk characterization is vital. EPA's Contaminant Candidate List (CCL) for drinking water contaminants represents one type of framework, though the selection of chemicals for the CCL has remained problematic due to the large number of chemicals that must be evaluated. Similar challenges are posed by microbes of potential concern. Methods to identify and prioritize these microbes in anticipation of potential health threats from environmental and intentional releases remains a critical unresolved dilemma.

QSARs are based on the relationship between the structure of chemicals and their interaction with biological tissues, leading to adverse effects, whereas VFARs extend this concept to microbial contaminants, suggesting that the pathogenicity of a microbial agent is directly related to the architectural and biochemical components found in that organism. Both QSAR and VFAR have the potential not only to facilitate the prioritization of chemicals and microbes of potential concern, but also to inform the subsequent risk assessment and risk management process.

The practice of risk assessment, which is composed of hazard identification, exposure assessment, dose response or toxicity assessment, and risk characterization generally integrates data from *in vivo*, *in vitro*, and epidemiological studies in the characterization of human health risk assessment. However, as previously mentioned, the chemical universe is large with the majority of these chemicals lacking traditional toxicity measurements. In such instances, risk characterization can integrate data from *in silico* methods in combination with *in vitro*, *in vivo*, and epidemiological studies to develop the weight of evidence (WOE) in the characterization for human health risk assessment. Thus, the theme that was advanced during the workshop was that additional useful evidence can be provided to enhance the overall hazard identification and toxicity assessment based on QSAR and VFAR methodology. Additionally, the analysis of microbial virulence factors can provide critical information for identifying sources of biological exposure and contamination.

Major Themes Discussed

Throughout the workshop, the following themes were discussed by experts in the VFAR and QSAR fields. The QSAR concept, used for chemical toxicity prediction, is more mature than the VFAR concept, which may be used for assessing hazard from exposure to microorganisms. In an introductory presentation on VFARs, Dr. Gerard Stelma of the National Exposure Research Laboratory (NERL) explained that the VFAR concept is related to the architectural and biochemical components of microorganisms that are defined by both genes and proteins. These components are related to pathogenicity and human health risks as presented in a series of National Research Council (NRC) meetings (1999, 2001). As indicated by Dr. Joan Rose, Michigan State University, and Principal Investigator of EPA and the Department of Homeland Security (DHS)-funded Center for Advancing Microbial Risk Assessment (CAMRA), the challenges of integrating potential applications of VFAR into the risk assessment framework is under intense scrutiny. Dr. Syed Hashsham, Associate Professor in the Department of Civil and Environmental Engineering and the Center for Microbial Ecology at Michigan State University, explained how descriptors for the microbial genome and proteome can be applied to evaluate the impact of variability on microbial virulence, broadly defined as the ability of a microbial agent to infect its human host, reproduce, and/or cause disease. Dr. R. Paul Schaudies of Science Applications International Corporation (SAIC) then described a rapid technique for identifying variability in the microbial genome. The technique can be readily used for identification and hazard assessment.

Although the science of QSAR is more mature than that of VFAR, there remain important challenges in the application of QSAR to the risk assessment paradigm. One of these challenges is the determination of how mode of action (MOA) data can be employed more fully to improve prediction of toxicity benchmarks using QSAR. Dr. Mark Cronin, Professor of Predictive Toxicology, Liverpool John Moores University, addressed a challenge that has puzzled researchers for years: the identification of a method for quantifying the

structure-activity relationship of highly reactive electrophiles. Dr. Kannan Krishnan, Director of the Human Toxicology Research Group at the Université de Montréal, described a model that integrates QSAR with physiologically based pharmacokinetic (PBPK) modeling to derive extrapolation capabilities. This model can be adjusted for variations in exposure route, rate, duration, and other factors. Dr. Andrew Maier, Associate Director of Toxicology Excellence for Risk Assessment (TERA), discussed how QSAR could provide critical information in risk characterization based on the WOE approach. In his talk, Dr. Maier emphasized that an understanding of the chemical MOA can enhance the applicability of QSAR in risk assessment. Dr. William Welsh, Department of Pharmacology, University of Medicine & Dentistry of New Jersey (UMDNJ), discussed the assemblage of a wide variety of computational toxicology tools, including QSAR-based methodologies that are applicable to risk assessment. Dr. Andrew Worth, leader of the QSAR Project, European Chemicals Bureau, Institute for Health & Consumer Protection, Joint Research Centre, European Commission, described how the new REACH legislation, which incorporates a preference for alternatives to animal testing, is serving to promote QSAR research and applications.

Charge to the Expert Panel

Following the presentations by the expert panels, a number of questions charged to the panel were discussed. Each charge question under VFAR and QSAR, given below, is followed by highlights of considerations by the panel members.

VFAR

1. Identify selection criteria for virulence factors that should be considered in the VFAR approach. Should certain classes of virulence factors be excluded?

   The initial development of VFAR methodology and technology allows for a very broad array of gene identification. Thus, there is no need to omit any classes of virulence factors (VFs) from consideration and no reason to rule out anything until it can be demonstrated that it is not relevant. The presence of a VF may be necessary but not sufficient for the development of pathogenicity. Other factors, such as those that permit the expression of VFs, the survival and persistence of the microbes, or even a particular array of microbes in the environment, are needed to permit the development of pathogenicity or the occurrence of disease. There is also an urgent need to characterize background levels of common VFs in organisms to better recognize a change in conditions that may pose a human health risk.

2. Compare and contrast the VFAR and QSAR approaches. Considering the similarities to QSAR, should the VFAR approach work with biotoxins? Viruses? Spores? Cysts? What are the strengths of the VFAR concept?

   Host-specific factors (e.g., individual variability in metabolism, sensitive subpopulations, the immune response of the host) alter the dose-response relationship in all traditional toxicity testing protocols. Therefore, there will always be uncertainty associated with such factors, which will be extended to QSAR and VFAR modeling efforts. Due to the variability of individual immune system function, host-specific factors are more important for microbial agents than for chemical agents. However, these limitations should not be a deterrent for using these approaches in the evaluation of the universe of chemical and microbial agents that need to be assessed using nontraditional methods. Because of some commonalities between biotoxins, viruses, spores and cysts, the VFAR approach may be useful in assessing the hazards associated with these different forms of biological agents. For the initial prioritization of chemicals or microbial agents, when toxicological or empirical data are lacking, QSAR and VFAR can be particularly useful.

   The data being collected and models under development could be critical to facilitating a rapid response in the event of an intentional attack. Available empirical data could be linked to predictions regarding virulence and potential adverse outcomes. QSAR and VFAR can provide critical information regarding alerts to human health concerns, and chemical and biological plausibility in terms of potential human health effects, particularly as an input to comprehensive WOE approaches.

3. Discuss how VFARs can be used in the detection of recognized biothreat agents, newly emerging pathogens, and bioengineered pathogens.

   It is unlikely that VFs will be the focus of genetic engineering for the purpose of bioweapon development. However, the analysis of VFs can provide information regarding genetic engineering for both bioweapons and natural evolution. In addition, it should be noted that other characteristics, such as factors that enhance gene expression or environmental persistence, will also play a role in exposure and risk.

4. Describe technology available for examining virulence factors. How can we determine the presence of such virulence factors in water or air?

   There are many tools and technologies available for examining virulence factors, including genomics and gene arrays, polymerase chain reaction (PCR), and proteomics for the analysis of protein products. These technologies are all under development in terms of applicability to VFARs, but there are current limitations in terms of the identification and characterization of VFs that have a hazard associated with them and the background occurrence of VFs. Difficulties in sample collection and processing still exist and must be addressed before these technologies can be applied to surveillance in water or air.

5. Discuss the positive and negative applications of using VFARs in bioengineering. Discuss the construction of highly potent pathogens inserting single gene or combinations of virulence genes into commensal organisms. Do certain classes of virulence genes lend themselves to genetic engineering?

   The genetic changes that occur naturally are an excellent example of the ingenuity of the microbial genome. Most notably, microbes can transfer plasmids, resulting in the rapid exchange of genetic material. Because virulence mechanisms are not completely understood, there is a need to look for unusual combinations of genes, as well as other factors, such as gene arrays and genes that are either up- or down-regulated. In general, a change in potency is accompanied by a string of changes, not just a single change.

6. How can VFARs be used to determine the human toxicity potential of the virulent genes? Is it possible to obtain a quantitative estimate of the virulence along with a qualitative estimate?

   - For the purposes of public health protection, the goal is to be able to use VFARs to aid in the:
   - Identification of the presence of microbes of concern
   - Identification of accessory genes necessary for virulence
   - Identification of environmental conditions necessary for virulence
   - Extrapolation from virulence gene expression to virulence protein expression
   - Prediction of the magnitude of the health hazard it represents
   - Determination of the infectivity or dose-response relationships to gauge the response needed to prevent or mitigate an outbreak

   These characterizations and predictions would provide information critical to an understanding of the magnitude of the public health risk associated with a natural or intentional exposure event.

7. Can a virulence gene be altered so that it is still active but no longer detectable by the gene probes that are typically used?

   The current state of knowledge is focused on the identification of virulence factors and how the virulence factors function in the microbe to express virulence. The capability does not yet exist to link this information to health outcomes, though the potential clearly exists. Due to the degeneracy of the genetic code, alterations in the gene might not result in a change in the synthesized protein. Constant changes in the microbial genome necessitate surveillance for these genetic mutations and evaluation of how virulence is affected.

QSAR

1. In light of emerging technologies (e.g., genomics, proteomics, and bioinformatics), what role will QSAR methods play in the future with regard to EPA's risk assessment/risk management process?

   For the purposes of regulatory prioritization and the development of remedial action strategies, the universe of chemicals must be characterized and reduced to assess the chemical threats to human health. Also, in order for chemical characterization to be most effective, mechanisms of toxicity or MOA must be determined. This is an essential component of expert system based structure-activity relationships where the aspect of the structure of the chemical that results in a particular effect or outcome must be determined. This concept can greatly enhance QSAR model development and interpretation.

2. How can genomic, proteomic, and bioinformatic data be used in QSAR models? Are there examples where the "-omics" technologies in combination with QSAR models have proven to be able to predict, both qualitatively and quantitatively, acute/chronic toxicity across multiple chemical classes?

   In terms of the role of -omics and QSARs in EPA's framework for risk assessment, any useful and valid information will help decrease uncertainty in the context of the overall weight of evidence. Some technologies may be better for screening than for regulatory decision making in that these technologies may not be fully validated or accepted. -Omics technologies and QSARs fit into this category. Currently, -omics technologies serve primarily as hazard identification tools by providing insight into the chemical's potential MOA. Such knowledge can provide informed interpretation of QSARs. The integration of QSARs with -omics technologies may allow these complementary technologies to reinforce each other. Computational toxicologists are working on this integration.

3. Can QSAR methods be used to reduce the uncertainty in extrapolating from acute and short-term benchmarks (such as median lethal dose [LD_{50}]) to subchronic and chronic lowest observed adverse effect levels (LOAELs)? What are the issues that must be dealt with in order to do this?

   There are distinct challenges in using QSARs to inform the extrapolation from acute to chronic effects because the critical endpoints are different. If there is knowledge about the critical effects and MOA, then it may be possible to use QSARs to extrapolate and reduce uncertainty. It is possible that there are cases in which the critical effects and MOA are the same, such that extrapolation using QSARs may be helpful. If there is commonality in MOA, then extrapolation from acute to chronic is more reasonable, but the rationale and the uncertainties must be discussed explicitly. Discriminators also can be segregated by MOA.

Participants acknowledged that many approaches have been suggested for evaluation of chemicals that lack toxicity data. Some indicated it was possible to take the LD_{50} and divide by several uncertainty factors and use this derived dose as a substitute for chronic effects. Others assert that since the MOA for acute effects is generally different from that of chronic effects, it is inappropriate to extrapolate from acute to chronic effects for most chemicals.

4. Since rule-based and expert models are based on congeneric groupings of chemicals (i.e., the training set is a congeneric data set), how can statistical models, which are generally based on non-congeneric training sets, be improved? Can such models incorporate MOA data if available? Can such statistical models provide some insight regarding MOA for a chemical query?

   There are several opportunities to combine QSARs and MOA information to better inform risk assessment, and the panel noted that routine acceptance of QSAR predictions will likely require that they be derived with an underlying mechanistic understanding. As models become more sophisticated, they will further incorporate structural features and property features and therefore allow for fuller evaluation of chemicals through the consideration of MOA data. Several examples of developments in this area were described. The integration of QSARs with PBPK modeling was discussed, where MOA considerations (e.g., identification of appropriate dose metrics based on chemical metabolism prediction) are factored into the PBPK model. Growing use of tools in bioinformatics (e.g., protein structure prediction and libraries) have allowed for the use of shape signatures based on the comparison of surface features to integrate MOA (e.g., receptor binding) into QSAR methodology. MOA data can be applied to large groups of chemicals to identify clusters of closely related chemicals. This is the conceptual basis for decision tree and regression tree approaches. QSAR models can be tailored via selection of descriptors for each cluster to provide more uniform training sets for QSAR development or to aid in interpreting global QSAR predictions.

5. The toxicity of a chemical for any given health endpoint is in general due to an adverse interaction between the chemical and/or its metabolite and the tissue/organ/DNA associated with the endpoint. In developing statistically based QSAR models for chemicals with different modes of action, the descriptor pool contains descriptors that are chemical specific (i.e., they depend on the structure of the chemical alone). Are there any descriptors that can describe the tissue/organ/DNA characteristics and its interaction with a chemical and/or its metabolites?

   The focus of QSARs is on describing the potential interaction between chemicals and biological molecules. There are two basic types of chemical-biological interactions. Receptor-based interactions often are the basis of endocrine disruption effects, and covalent interactions occur with nonspecific macromolecular binding. Mechanistic QSARs for predicting receptor-based interactions are commonly used in drug development and are increasingly being used for toxicity prediction. Nevertheless, many chemicals act via relatively nonspecific covalent interactions, which can be quite complex even within a chemical class, as was highlighted in the context of phenolic electrophiles. To be most useful, QSARs need to account for this complexity more fully. While mechanistic QSARs are preferred, an intermediate step in this direction is to focus efforts on endpoint-specific QSARs, since the specificity of target organs can arise based on local metabolism or the nature of cell/tissue response (toxicodynamics).

6. Current methodology on the statistically based QSAR development for toxicity prediction calls for the inclusion of as many (classes of) descriptors in the descriptor pool as possible to explain the variance in the dependent variables (some measure of toxicity). In developing these QSARs, are there any (class of) descriptors that one should definitely include in the potential descriptor pool (e.g., partition coefficients to account for transfer from blood to tissue)?

   Although certain descriptors (i.e., molecular size or hydrophobicity) are more commonly used, the mechanistic context must be used as a starting point for the selection of descriptors. Since the mechanistic context varies based on chemical class, it is not possible to make blanket statements regarding the selection of descriptors. Examples of descriptors based on chemical mechanisms are those descriptors that describe accumulation or penetration through membranes, reactivity with macromolecules, receptor binding with critical targets, and others.

7. Qualitative SAR models (i.e., models yielding dichotomous or graded responses such as yes/no or low/med/high) do not provide a quantitative measure of a chemical's toxicity while quantitative SAR models (i.e., models yielding numerical potency estimates) do not provide a qualitative measure of the activity of a chemical for any given health endpoint. How does the panel view the feasibility of applying hybrid QSAR models (i.e., capitalizing on the benefits of SAR and QSAR by minimizing the disadvantage, if any, of each approach) for toxicity prediction? If feasible, how does the panel envision EPA applying such models?

   Several approaches for hybrid SAR/QSAR analyses were discussed. Approaches ranged from using MOA descriptors as a screening step for the initial classification of chemicals to help in interpreting global QSARs to direct use of MOA descriptors in developing quantitative endpoint-specific logistic regression models. Semiquantitative QSARs methods include decision trees or modifications of this concept

that use parallel sets of decision trees to improve predictability. Binned chemicals identified through these tools could serve as endpoint-specific QSAR training sets or be used to identify characteristics associated with potency levels for risk assessment using threshold-of-concern approaches. For chemical risk assessment, there is often a need to extrapolate from dose-response data based on exposure durations of less than a lifetime to estimate the effects of lifelong exposure. Traditionally, for EPA risk assessments, a default factor of 10 is applied to adjust adverse effect levels from subchronic (i.e., exposure for roughly 10 percent of the lifetime) to chronic exposure conditions. This default factor of 10 can be useful for the extrapolation of subchronic to chronic toxicity; however, it may be inappropriate for the extrapolation from acute to chronic exposure because the critical endpoints are often different and the MOA is different between acute and chronic exposure. The panel noted that several correlation approaches have been developed to address this situation – but these are not necessarily QSARs. While QSARs may fully address this application directly, they can also provide very important insights that are used in decisions regarding such extrapolations. For example, they are used to predict toxicokinetic parameters (e.g., partition coefficients or metabolism parameters) that impact decisions regarding the potential for increased body burden with longer-duration exposures. Furthermore, QSARs can provide understanding of both acute and chronic toxicity mechanisms — which impact considerations of potential for accumulation of tissue damage with increased exposure duration.

Recommendations

Several recommendations of near-term applications of VFAR/QSAR models were discussed by the panel. These include:

- To advance the applicability of VFAR in real-world situations, it is critical to facilitate the characterization of samples collected during natural outbreaks of microbial diseases. This will permit the identification of background levels of VFs and advance understanding of the natural evolution of VFs in addition to providing the framework to test hypotheses pertaining to the dose-response relationships of VFAR.
- Another potential opportunity for the advancement of VFAR research involves the BioWatch Program, which consists of continuous sampling at locations across the country. This would be an opportunity for researchers to obtain material for the characterization of background levels of VFs in urban environments in addition to testing hypotheses.
- The state of the science regarding QSAR modeling is considerably more advanced than that of VFAR modeling; therefore, the key recommendation for near-term applications focused on the integration of MOA and PBPK with QSAR models to enhance biological applicability.
- For both VFAR and QSAR, host-specific factors alter the dose-response relationship (e.g., human variability in metabolism, sensitive subpopulations, immune response of the host); therefore, there will always be uncertainty in the ability to model host factors. Due to the variability in human immune system function, host-specific factors are important considerations when evaluating responses to microbiological agents. However, these limitations should not be a deterrent for using these approaches in the evaluation of the vast universe of chemicals and microbes that require attention. For the initial prioritization of chemicals or microbes, when toxicological data are lacking, QSAR and VFAR can be particularly useful.
- The data being collected and models under development could be critical to facilitating a rapid response in the event of an intentional attack by linking field data to predictions regarding virulence and potential adverse outcomes. QSAR and VFAR can provide critical information regarding alerts to human health concern, and chemical and biological plausibility in terms of potential human health effects—particularly as an input to comprehensive WOE approaches.

1.0 Introduction

1.1 Background

The U.S. Environmental Protection Agency's (EPA's) National Homeland Security Research Center (NHSRC) and National Risk Management Research Laboratory (NRMRL) convened this workshop on June 20–21, 2006, to explore the development and application of Quantitative Structure-Activity Relationship (QSAR) and Virulence Factor-Activity Relationship (VFAR) models in the risk assessment process, specifically as they relate to homeland security needs and contamination associated with natural disasters and accidental or intentional releases. To this end, the workshop convened toxicologists, microbiologists, chemists, engineers, biostatisticians, pharmacologists, biochemists, and risk assessment specialists to address the goals of the workshop. These goals included the identification of data needs for the development of quantitative noncancer and cancer models, that are capable of predicting commonly used toxicity benchmarks, such as the lowest observed adverse effect level (LOAEL), LD_{50}, and benchmark dose (BMD), for various exposure durations. Of particular importance is the prediction of benchmarks and health effects associated with acute and short-term exposure to chemical and biological agents. The workshop explored the development and application of VFAR models to estimate the human health effects of microorganisms and their biological toxins. The workshop also focused on approaches for incorporating mode of action (MOA) data in the development or refinement of such models, including the incorporation of genomic, proteomic, and metabolomic data. In addition, the workshop addressed computational and data mining approaches, such as various regression methods, neural networks, and expert systems for improving QSAR and VFAR development.

The risk assessment process involves four steps as defined by the National Academy of Sciences, National Research Council (NRC, 1983): hazard identification, dose response or toxicity assessment, exposure assessment, and risk characterization. Risk management integrates the results of the risk assessment with other considerations, such as economic or legal concerns, to reach decisions regarding the need for and practicality of implementing various risk reduction activities. NHSRC and NRMRL, both part of EPA's the Office of Research and Development (ORD), are primarily involved in dose response or toxicity assessment, and in developing guidance for risk management. Under these processes, an attempt is made to understand the toxic properties of individual chemicals as well as mixtures of chemicals, and develop appropriate guidance documents. An important goal of research in toxicology is the prediction of the toxic potential of chemicals from acute short-term and long-term chronic exposures.

Globally, the chemical industry and regulatory agencies such as EPA spend millions of dollars on testing and assessing the health risks associated with chemicals. For most chemicals, the risk assessment process is conducted using limited experimental data. In such instances, the ability to rapidly and accurately predict potential health hazards from chemical exposures is needed. One approach to meeting this need is the use of nonempirical parameters, which can be calculated directly from a chemical structure. This can be achieved by the application of computational toxicology or QSAR models, which have proven to be both appropriate and useful for many chemicals. Similar computational toxicology approaches are also being employed to enhance risk assessment processes for exposure to microorganisms and their toxins. This field, involving the methodologies for deriving VFAR, is emerging to estimate the health hazards posed by biological agents via the characterization of proteins, which convey toxicity, infectivity, pathogenicity and/or virulence. The concept of VFAR was developed as a way to relate the structural, architectural, and biochemical components (such as biotoxins) of a microorganism to its potential to cause human disease.

1.2 Purpose and Goals of the Workshop

The workshop was conducted to explore the application of these techniques to the risk assessment (RA) process in situations where chemical-specific empirical data are either inadequate or lacking.

The following list details the goals and objectives of the workshop:

- Identification of data needs for the development of quantitative noncancer and cancer models, including models that are capable of predicting benchmarks such as LOAEL, LD_{50}, median lethal concentration (LC_{50}), BMD, and benchmark concentration (BMC) for various exposure durations.
- Prediction of benchmarks and health effects associated with acute and short-term exposure to chemical and biological agents.
- Exploration of the feasibility of developing and applying hybrid QSAR models.
- Exploration of the development and application of VFAR models to estimate the activity of microbial agents.
- Exploration of the incorporation of genomic, proteomic, and metabolomic data into QSARs in order to incorporate the MOA into QSAR models.
- Assessment of the development of models for predicting the relative toxicity of the parent compound and metabolites for identification of the ultimate chemical effectors.

- Discussion of computational approaches, such as various regression methods, genetic algorithm descriptor selection techniques, data clustering methods, neural networks, and expert systems.

1.3 Charge to the Expert Panel

Following the presentations by the expert panels, the following questions were discussed:

VFAR

1. Identify selection criteria for virulence factors that should be considered in the VFAR approach. Should certain classes of virulence factors be excluded?
2. Compare and contrast the VFAR and QSAR approaches. Considering the similarities to QSAR, should the VFAR approach work with biotoxins? Viruses? Spores? Cysts? What are the strengths of the VFAR concept?
3. Discuss how VFARs can be used in the detection of recognized biothreat agents, newly emerging pathogens, and bioengineered pathogens?
4. Describe technology available for examining virulence factors. How can we determine the presence of such virulence factors in water or air?
5. Discuss the positive and negative applications of using VFARs in bioengineering. Discuss the construction of highly potent pathogens inserting single genes or combinations of virulence genes into commensal organisms. Do certain classes of virulence genes lend themselves to genetic engineering?
6. How can VFARs be used to determine the human toxicity potential of the virulent genes? Is it possible to obtain a quantitative estimate of the virulence along with a qualitative estimate?
7. Can a virulence gene be altered so that it is still active but no longer detectable by the gene probes that are typically used?

QSAR

1. In light of emerging technologies (e.g., genomics, proteomics, and bioinformatics), what role will QSAR methods play in the future with regard to EPA's risk assessment/risk management process?
2. How can genomic, proteomic, and bioinformatics data be used in QSAR models? Are there examples where the "-omics" technologies in combination with QSAR models have proven to be able to predict, both qualitatively and quantitatively, acute/chronic toxicity across multiple chemical classes?
3. Can QSAR methods be used to reduce the uncertainty in extrapolating from acute and short-term benchmarks (such as LD_{50}) to subchronic and chronic LOAELs? What are the issues that must be addressed in order to do this?
4. Since rule-based and expert models are based on congeneric groupings of chemicals (i.e., the training set is a congeneric data set), how can statistical models that are generally based on noncongeneric training sets be improved? Can such models incorporate MOA data if available? Can such statistical models provide some insight regarding MOA for a chemical query?
5. The toxicity of a chemical for any given health endpoint is, in general, due to an adverse interaction between the chemical and/or its metabolite and the tissue/organ/DNA associated with the endpoint. In developing statistically based QSAR models for chemicals with different modes of action, the descriptor pool contains descriptors that are chemical specific (i.e., they depend on the structure of the chemical alone). Are there any descriptors that can describe the tissue/organ/DNA characteristics and its interaction with a chemical and/or its metabolites?
6. Current methodology on the statistically based QSAR development for toxicity prediction calls for the inclusion of as many (classes of) descriptors in the descriptor pool as possible to explain the variance in the dependent variables (some measure of toxicity). In developing these QSARs, are there any (classes of) descriptors that one should definitely include in the potential descriptor pool (e.g., partition coefficients to account for transfer from blood to tissue)?
7. Qualitative SAR models (i.e., models yielding dichotomous or graded responses such as yes/no or low/med/high) do not provide a quantitative measure of a chemical's toxicity while quantitative SAR models (i.e., models yielding numerical potency estimates) do not provide a qualitative measure of the activity of a chemical for any given health endpoint. How does the panel view the feasibility of applying hybrid QSAR models (i.e., capitalizing on the benefits of SAR and QSAR by minimizing the disadvantage, if any, of each approach) for toxicity prediction? If feasible, how does the panel envision EPA applying such models?

1.4 Organization of This Report

The remainder of this report is organized as follows:

- Chapter 2 presents the background of the workshop's sponsoring organizations, NHSRC and NRMRL.
- Chapter 3 provides summaries of the VFAR presentations made by expert panelists, including ensuing discussions from panelists and other workshop participants.
- Chapter 4 provides summaries of discussion based on charge questions related to the VFAR concept posed to the expert panel.
- Chapter 5 provides summaries of the QSAR presentations made by expert panelists, including ensuing discussions from panelists and other workshop participants.
- Chapter 6 provides summaries of discussion based on charge questions related to the QSAR concept posed to the expert panel.

- Chapter 7 includes major considerations to which discussions of the charge questions gave rise.
- Chapter 8 provides references mentioned during presentations on the QSAR and VFAR concepts.
- Appendix A presents a list of workshop speakers.
- Appendix B provides “biosketches” of the speakers and expert panelists.
- Appendix C contains a copy of the workshop agenda, as well as the EPA-distributed flyer for the workshop.
- Appendix D provides a list of all workshop attendees.
- Appendix E includes copies of all presentation materials in Microsoft PowerPoint slides.

2.0 Background and Opening Remarks

This section summarizes the background of the workshop's sponsors, National Homeland Security Research Center (NHSRC) and National Risk Management Research Laboratory (NRMRL). Three main speakers from these two ORD entities set the stage for the workshop discussions, and their presentations are synopsized.

2.1 NHSRC and NRMRL

NHSRC, headquartered in Cincinnati, Ohio, was formed in 2002. It manages, coordinates, and supports a variety of research and technical assistance efforts and develops and delivers reliable, responsive expertise and products based on scientific research and evaluations of technology. NHSRC's expertise and products are widely used to prevent, prepare for, and recover from public health and environmental emergencies arising from terrorist threats and incidents. The center provides a management structure that ensures effective design and oversight of research and facilitates interaction with EPA program offices and regions, other federal agencies, the private sector, and research partners. NHSRC's team of scientists and engineers are dedicated to understanding the terrorist threat, communicating the risks, and mitigating the results of attacks. Guided by the roadmap set forth in EPA's Strategic Plan for Homeland Security, NHSRC ensures rapid production and distribution of security-related products. These products include methodologies and tools to support contaminant detection and characterization, treatment and decontamination, physical security enhancement, risk assessment and communication, as well as numerous papers and technical briefs covering a variety of topics.

The mission of NRMRL is to develop ways to prevent and reduce pollution of air, land, and water. With headquarters in Cincinnati, Ohio, and divisions in North Carolina, Oklahoma, and New Jersey, NRMRLs several hundred scientists and engineers share the mission to solve a wide range of environmental challenges in seven research areas: drinking water protection, air pollution control, contaminated media remediation, watershed management and protection, environmental technology verification, technology transfer, and technology support.

2.2 Opening Presentations

NHSRC and the Workshop Goals

Andy Avel, Assistant Center Director, NHSRC

Mr. Avel stated that in the event of a terrorist attack, both EPA and the Department of Homeland Security will have responsibility for cleanup. However, after first responders leave, EPA will have the primary responsibility for remedial activities. Since the September 11, 2001, terrorist attacks, EPA, under a series of Homeland Security Presidential Directives (HSPDs), has been given specific roles, including decontamination of buildings, public infrastructure, and public areas in the event of biological, chemical, or radiological terror attacks, and protection of the drinking water infrastructure. NHSRC is organized to address chemical, biological, and radiological weapons of mass destruction targeted toward water and the environment. Its primary focuses include:

- Developing detection methods to identify an attack
- Developing risk assessment methodologies to assess, characterize risks, and provide guidance for cleanup and reentry
- Understanding and anticipating chemical and biological warfare agents
- Incorporating the radiological component

Mr. Avel went on to discuss the need to build on what has already been done, particularly in terms of cleanup management. Once contamination occurs, he said, impact has to be minimized via containment. Once contained, impacted media must be assessed for potential human exposure and health risk and must be handled to remove/reduce contamination. Once decontaminated, the removed hazardous materials or residues must be disposed of, using the best available control technology (e.g., landfill, thermal destruction), according to local, state, and federal regulations.

NHSRC has developed a technology verification program to test claims made by industry regarding the technologies for managing chemical, biological, or radiological agents. NHRSC also is expanding its capability for the analysis of these chemicals and agents and increasing lab capacity in general. NHSRC is currently collaborating with other parts of EPA including the Office of Solid Waste and Emergency Response (OSWER), the National Center for Computational Toxicology (NCCT), the National Center for Environmental Assessment (NCEA), the National Exposure Research Laboratory (NERL), and NRMRL.

Cindy Sonich-Mullin, NHSRC, Director of Threat and Consequence Assessment Division (TCAD)

Ms. Sonich-Mullin discussed the mission of TCAD's research program: to become better prepared to respond to threats and emergency incidents. She stressed the need for rapid response to specific threats and risks from terrorism. Among TCAD's goals are:

- Adapting and developing risk assessment methods for homeland security
- Developing tools for responders to access information
- Developing cleanup advisory levels and methods for achieving cleanup goals or levels

Exposure timeframes ranging from 24 hours, to 30 days, to 2 years are the focus of TCAD's risk assessment efforts. This is in contrast to acute exposure levels (< 24 hrs) and chronic exposure levels, which are traditionally based on reference doses (RfDs) and reference concentrations (RfCs), and are the traditional focus of the Agency. TCAD envisions that QSARs will play a significant role in filling data gaps for this timeframe. There is a need for risk assessment methods for different exposure scenarios, higher concentrations, and unknown agents. With very little data and capability, innovative techniques and approaches are required to develop a credible risk assessment.

The charge questions are the key to the success for facilitating development of TCAD's risk assessment capability. The focus is on short-term exposures, and TCAD is developing and using QSARs and VFARs to:

- Extrapolate from either acute or chronic exposures
- Decrease default uncertainty, traditionally applied in the risk assessment process
- Develop credible or sound cleanup level estimates for emerging chemical and biological agents, in an emergency
- Make these efforts transparent and rapid

NRMRL and the Workshop Goals

Subhas Sikdar, Acting Associate Deputy Director for Health, NRMRL

Dr. Sikdar noted that NRMRL has been involved with the computational toxicology initiative and QSAR methodology from the beginning, which led to the establishment of the National Center for Computational Toxicology (NCCT). NRMRL's goal for QSAR methods research is to predict the environmental outcomes of new chemicals throughout their life cycles, while working with the NCCT to develop analytical, computer-based models that decrease the need for animal testing. Dr. Sikdar reiterated that the goal of this workshop is to enhance QSAR and VFAR activities by bringing together experts in the field to discuss progress and the path forward.

Doug Young, Clean Processes Branch Chief, NRMRL

Dr. Young indicated that NRMRL's computational toxicology program was involved in the original development of QSARs and has representation on the current steering committee for NCCT. The NRMRL engineering lab is working to develop risk management solutions, including alternative solutions such as Life Cycle Assessment and environmental impact tools. Other categories of interest include quantifying impacts on human and ecological health by developing/using toxicological values as indicators. For example, given 2,000 chemicals to rank and prioritize while lacking toxicological values, new tools and techniques, such as QSARs and bioinformatics, are necessary to reduce the uncertainty of estimations. NRMRL is in the early stages of using VFARs and has a particular interest in having the Water Supply and Water Resources Division develop VFAR tools to evaluate recreational and drinking-water quality and potential risks.

3.0 VFAR Presentation Summaries

The following summarizes the presentations made for the VFAR concept, its use in a risk assessment framework, VFAR factors related to genomic variability, and a bioinformatics approach to VFAR. The discussion following each presentation is also summarized.

3.1 Introduction to the VFAR Concept

Gerard Stelma, Senior Science Advisor, NERL

Summary

The VFAR concept, as Dr. Stelma noted, originated during a National Research Council (NRC, 2001) meeting. The resultant report recommended the further development of the concept with a specific challenge to incorporate it into the drinking water program. NRC recommended the concept of VFARs to assist in the development of the Contaminant Candidate List (CCL) under the Safe Drinking Water Act (SDWA), as amended in 1996. SDWA requires that unregulated contaminants in drinking water be identified, prioritized, and reviewed by EPA as candidates for regulation. At the time of SDWA's reauthorization in 1996, there were no methods for CCL development and prioritization. The NRC subsequently developed a framework for the selection of both CCL and pre-CCL chemicals (2001).

Dr. Stelma stated that because priority chemicals must be identified to meet SDWA requirements, there is a basic need to prioritize the universe of unregulated chemical and biological contaminants. For biological contaminants, the goal is to explore the feasibility of using VFARs to evaluate microbes and develop a system that would parallel the QSAR approach for chemicals. The VFAR approach would emphasize emerging pathogens by building on evidence from previous research and developing a list of descriptors tied to pathogenicity. The idea is to focus on elements tied to virulence.

Dr. Stelma said that if the possibility exists to characterize the descriptors (i.e., genes, surface proteins, etc.), then the descriptors could be used to predict pathogens present in water. However, as pathogens are dynamic, gene arrays associated with pathogenic virulence may change over time. Thus, the use of VFARs for pathogen indication/identification may require constant updating to keep up with pathogen evolution. Additionally, there are virulence genes, such as hemolysins, that are necessary but not sufficient for virulence. Therefore, assaying multiple genes using a gene array may be important in determining virulence. The applicability of VFARs may be limited because the current methodology does not incorporate host susceptibility, the role of unexpressed virulence genes, and the effect of virulence factors from dead cells. Despite these limitations, VFARs can be an important tool in the pathogenicity assessment toolbox.

Discussion

Following Dr. Stelma's presentation, it was noted that an array of genes is often needed for the evaluation of potential virulence, which is a significant challenge. A workshop attendee went on to discuss another source of uncertainty of communal pathogenicity, that essentially some microbes may require the presence of other microbes to express their own virulence.

3.2 Using VFAR in a Risk Assessment Framework

Joan Rose, Homer Nowlin Endowed Chair for Water Research, Michigan State University

Summary

Dr. Rose began by discussing risk assessment as a method to qualitatively or quantitatively evaluate the potential for harm from exposure to contaminants or specific hazards. There are four components to risk assessment: hazard identification, dose response, exposure assessment, and risk characterization. Risk assessment principles can be applied to microbes not only to address natural outbreaks, but also to address the needs of homeland security. Hazard identification is the process of identifying the microbe, source of exposure, and the associated virulence. Microbial genetics is key to this process. For exposure assessment, the goal is to quantify exposure concentration, duration, and frequency, though source identification is also important. Genetic elements provide information regarding persistence both in the environment and during disinfection. Monitoring data, indicators, and models can also be used to estimate exposure concentrations.

Dr. Rose stressed the difficulty posed by obtaining dose-response information for microbial agents. Quantification is important, she said, as is the need to extrapolate from less pathogenic to more pathogenic strains, from healthy adults to sensitive populations, and from high dose to low dose. It is necessary to measure data in the same units as they are measured in the environment. Infectivity, the number of microbes needed to trigger infection, is another important characteristic that must be quantified. Infectivity may be related to virulence, though this is not known for certain. The process of risk characterization is the combination of all data to evaluate health risks. It is in dose response that uncertainty comes to the forefront.

Dr. Rose noted the need to characterize the background rate of gene occurrence. For example, to assess biohazards, the following must be understood:

- Why the genome of some microbes are conserved, while others are variable
- Why some genomes are host-specific and others are not

- Why some microbes cause chronic diseases and others do not

Other important factors are what controls occurrence, survival, regrowth, accumulation, attenuation, etc. Dr. Rose concluded her remarks by suggesting that the focus needs to be on future applications, though there will be uncertainty, and the applications may not be appropriate for all microbes.

Discussion

The ensuing discussion focused on the difficulty in characterizing host-organism interactions. There is a need to focus on the mechanisms that drive differential responses in microorganism-related host response. In the past it has been assumed that differential response was driven by host variables, although now the focus has been extended to pathogen factors as well. For example, characteristics such as the presence of housekeeping genes that enhance persistence, the expression of specific receptors, or the production of toxins will become the focus of classification as these characteristics determine virulence. It also may be necessary to categorize characteristics in different ways, such as lump and split techniques, based on health effects.

It was also suggested that the uncertainty that is inherent to the current application of VFAR might be difficult to accept. There is a need for qualitative and quantitative characterization of uncertainty. The understanding of variables that contribute to persistence, what allows a given microbe or family of microbes to survive and thrive in certain environmental conditions, is limited. One meeting participant noted that the Office of Water has an interest in the rapid identification of biological contaminants and questioned the application of the VFAR methodology for rapid screening since there are inherent uncertainties associated with the method (as discussed above). The panelists also indicated that identifying management strategies is essential and ties in with pathogen discovery and subsequent application of the VFAR method for hazard identification purposes. The potential exists to use well-studied pathogens as a starting point for a rapid identification tool. A dual-pronged approach, which focuses on reducing uncertainty while simultaneously developing monitoring/identification concepts, may be important for tool development and refinement.

3.3 VFAR Factors Related to Genomic Variability

Syed Hashsham, Associate Professor, Department of Civil and Environmental Engineering and Center for Microbial Ecology, Michigan State University

Summary

Dr. Hashsham began with a discussion of how the genome, proteins, and toxins of microbes can all be characterized via descriptors and how it is possible to use these descriptors for ranking and uncertainty analysis today. Virulence genes are associated with function (e.g., antibiotic resistance, virulence), and not necessarily microbial identity. Depending on the genome, variability can range from 1.6 percent to 20 percent. There are variable genes and variable effects, but in general, there are correlations between specific genes and adverse health effects that are worth exploiting. The actual link between health effects and gene variability is undefined. There is a need to develop gene-family training sets, which are groups of related virulence factors (VFs). Training sets can be used to demonstrate the applicability of models defining VFAR, drawing from large data sets of virulence and marker genes (VMG) that are under development. However, the link between health effects and gene variability and the quantification of health effects are key to VMG rankings and eventual pathogen prioritization. A possible approach to defining this link may involve tying species and genomic data to known outbreaks using historical outbreak data.

Dr. Hashsham noted that it is possible to look at rankings based on variability within species, length of gene, and number of virulence genes, to determine whether a gene is a potential marker. Some genes are better markers because they are more specific than others. Dr. Hashsham explained that the capacity to map the genome of different strains exists and common and variable regions can currently be identified. Genes that are constantly changing are more likely to be on plasmids. Fewer changes are found on certain parts of the chromosome. This information can be used to understand which genes are associated with virulence and pathogenicity. However, all of these changes in descriptors ultimately must be related to response. Data related to response as a function of differences in descriptors are deficient and require the most attention. For the purposes of monitoring, gene chips have been developed that contain the simultaneous genomic sequences of up to 20 pathogens. It is possible to conduct a high-throughput real-time polymerase chain reaction (PCR) to amplify any number of genes of interest, using multiple probes to ensure that specific virulence markers are identified, in a manner that is economical. This technology is useful for monitoring and identifying pathogens because it can target multiple VFs from each pathogen for enhanced reliability.

Discussion

One meeting participant noted that the small volumes used in chip development can be a problem; when only picoliter sample volumes are used for the chip, not all representative organisms will be in that small sample. More work is needed in terms of sample processing to ensure that the samples are adequately representative.

3.4 A Bioinformatic Approach to VFAR Analysis and Characterization

R. Paul Schaudies, SAIC

Summary

Dr. Schaudies stressed that the significant challenge in microbial risk is the rapid characterization of the microbe. The software program Fast Identification of Genomic Unique Regions (FIGUR) was developed to characterize microbes within hours by identifying unique elements within the entire chromosome. Using DNA microarray technologies, a pattern can be obtained for an organism that can subsequently be compared to other organisms in established databases for

the purposes of identification. If the pattern is not present in the databases, it is possible to determine whether the pattern is similar to known strains. From there it is possible to accomplish empirical generation of a library of these near-neighbor patterns. In contrast to PCR, amplification via microarray technology is random (using random primers, rather than gene-specific primers), which broadens applicability.

The results, generated by computer software, are color coded to identify unique and conserved sequences. Hybridizations on the chip can be included to demonstrate that there is no cross reactivity between genes. Data also can be filtered by hybridization cutoffs to focus attention on genes that represent an appropriate level of similarity.

Dr. Schaudies presented an example of three different species of Yersinia that are associated with the disease plague. Though all three species are 95 percent similar at the genomic level, the VFs differ among the three strains. Thus, it is possible to begin to develop profiles of VFs that define a species. Using VF profiles as a filter increases the chance of finding specific strains. One key feature of this approach is that it does not require the whole genome, but a part of the genome (e.g., 1Kb). The data can be analyzed serially to refine the comparisons, and common factors in each subsequent analysis can be removed to identify what is unique.

Discussion

This technology may be helpful in understanding the association between VF and pathogenicity. The identification of genes that are present within a broad array of genes can be accomplished within hours rather than days. Predictive capacity is not currently programmed into this tool; its development was not funded in the current application.

The discussion focused on the need to test this application in both clinical and environmental settings to help determine research needs. Validation is needed for VFAR in strains with known differences in virulence. A panelist noted that it would be interesting to compare *Bacillus cereus* strains, which exhibit variability in pathogenicity (as demonstrated by Dr. Schaudies), with *anthracis* strains, which exhibit very little variability.

4.0 VFAR Charge Questions

The following summarizes the discussions on charge questions related to VFAR. Discussion under each charge question is summarized as themes discussed related to the question. Each theme is a bullet point under the charge question followed by the summary discussion of the theme.

4.1 Summary of VFAR Charge Questions Discussion

1. ***Identify selection criteria for virulence factors that should be considered in the VFAR approach. Should certain classes of virulence factors be excluded?***

    - **No virulence factor, or selection criteria, omissions should be made at this point in the development of the VFAR approach.**

Participants noted that the VFAR concept is still in the initial stages of its scientific development and it is important to collect as much data as technologically feasible (within the economy of scale) as it may not be possible to go back and retrieve that data later (e.g., following an outbreak or event). Selection criteria should not be reduced, particularly at the outset.

In the initial development of the VFAR methodology, there is no need to omit any known or potential VFs from consideration. Technology allows for a very broad array of gene identification, which is relatively simple and inexpensive. The challenge is in determining which genes represent critical VFs. A single virulence factor may require the expression of multiple genes to be effective. The capacity to make these types of determinations will come only from the collection of a large quantity of data pertaining to the existence and ecology of the VFs. Currently, the library of known virulence sequences is limited; hence, there is no need to limit the collection of VF data.

If VFARs are thought to be analogous to QSARs, where VFARs explore whether specific microbes cause disease, then tools that are developed should include, to the extent possible, all known VFARs associated with disease. Subsequently, a host of microbes can then be prioritized, although the process of identifying and characterizing VFARs is ongoing and remains far from the threshold of utility. To develop predictive capabilities for EPA, Food and Drug Administration (FDA), and other agencies, there is a need to take theoretical, empirical, and Bayesian approaches to the analysis of VFAR-related data, in conjunction with other predictive techniques. At this time, the parameter sensitivities of VFARs are not known.

It was pointed out that the fate of accumulated data might depend upon the questions that are asked. Are the questions related to basic monitoring or source identification? Development of a database that will aid in the identification of sources of microbes would facilitate understanding of how genes combine from different populations (e.g., Zoonotic transmission) and how this can lead to virulence. Ultimately, such database development can aid in the prediction of outcomes and the development of management options.

- **Data gaps on background occurrence of VFARs are a current challenge.**

If a background sample from the natural environment is analyzed for the presence of genes representing candidate VFs, VFs will be found to be present. Therefore, background conditions need to be better characterized and understood. Similarly, when the genetic signatures of an organism of interest are characterized, these sequences will also be found in background samples. To interpret the significance of these genetic signatures, it will be necessary to sample and analyze the background environment to see how those known genes correlate. This will help to develop the database using a more focused approach.

Another consideration raised is that any of the most deadly bacterial toxins can be engineered into multiple species; this actually occurs in nature in cases where certain toxins transcend species. The categorization of microbes may have to be rethought. As one attendee asked, is the concern over species identification, or should the species be characterized based on potential health effects? For example, if symptoms characteristic of plague were encountered, would one first look for toxins associated with the plague?

- **In VFAR development, proteomics can be used to inform genomics.**

Proteins are less conserved for screening, but it may be useful to start with proteins and work back toward the genes. VFARs are based on the understanding that function follows structure. There is an obvious role for proteomics in VFAR analysis. This approach has been used with viruses, where proteins were characterized first and then characterization moved back toward the genome. Though protein is less conserved for screening technologies, collection of more data to identify and characterize VFs will result in a better understanding of the protein structures to enable their direct use.

- **A dual-pronged approach, combining short-term practical applications using the current knowledge base with ongoing research, development, and refinement of the methodologies, may work best to advance the science of VFARs.**

Another consideration raised was the need for short-, mid-, and long-term goals and approaches. In order to achieve long-term goals, all approaches to develop VFARs should be considered. However, in the short term, with limited knowledge and limited research funding, it may be helpful to select a group of genes that are known to be associated

with pathogenicity and focus on them during development of limited monitoring and assessment tools. For example, it may be necessary to focus on a subset of data collected to achieve specific short-term goals such as development of monitoring or screening tools. While work proceeds toward short-term goals, mid- and long-term research can be planned and means to reduce uncertainty, expand capability and capacity, and increase applicability can be laid out.

In the short term, one approach is to use what is known to develop and demonstrate concepts. In the mid-term, data can be collected to add to what is known and to determine what works for predictability. Such approaches can be applied or tested on the growing body of data.

2. ***Compare and contrast the VFAR and QSAR approaches. Considering the similarities to QSAR, should the VFAR approach work with biotoxins? Viruses? Spores? Cysts? What are the strengths of the VFAR concept?***

There are several common factors for the use of VFARs and QSARs in risk assessment. For both chemical and biological threats to human health, the chemical and microbial universes need to be characterized and reduced.

For both approaches to be most effective, mechanisms of toxicity or modes of action must be determined. This is an essential component of expert system based structure-activity relationships, where the aspect of the structure of the chemical that results in a particular effect or outcome must be determined. This concept can greatly enhance QSAR and VFAR model development and interpretation. In the case of microbial virulence, the structure may refer to a physical structure resulting from protein expression and subsequent processing, carbohydrate metabolism, or genetic coding.

Unlike chemicals, microbes are dynamic. Chemicals may exhibit different properties in different environments and can be metabolized in the body, producing a range of metabolites that may or may not be toxic. Microbes, as living organisms, can exhibit rapid evolution. The flexibility of microbes and viruses, which refers to their ability to transfer genes on plasmids or into the bacterial chromosome, as well as their rapid evolution over short periods of time, present unique challenges to the development of a VFAR methodology.

It is challenging to use VFARs and QSARs in dose-response determinations as they require large quantities of data derived from multiple testing approaches. In the near term, it may be easier to predict hazards by identifying the potential for adverse health outcomes and looking to VFARs and QSARs as more robust tools for screening.

There is a deficiency of tools for rapid or instantaneous identification of biological organisms for use in emergency situations. The available tools, which use culture methods and genetic techniques to identify microbes that are present, can be used to determine whether illnesses are caused by intentional events (e.g., Salmonella in the salad bar). However, these techniques take time. There may be other characteristics in addition to VFARs, such as factors that enhance gene expression or environmental persistence, that indicate the presence of weaponized forms of biological species or an intentional exposure event. The application of molecular techniques is likely to be the most sensitive, specific, and rapid approach.

- **There are several analogous short-, mid-, and long-term goals.**

Panelists discussed that in regard to long-term goals, the current state of VFARs is analogous to that of QSARs many years ago. VFAR tools for identification are under development; however, it is possible that scientists are spending too much time on tool components and not enough on tool composition. In regard to VFARs in particular, it is necessary to move beyond the academic arena and test hypotheses to reveal data gaps.

In the short term, the goal may be to construct a framework for VFAR analysis rather than to focus on details. Mid- and long-term goals could focus on details and, using an iterative approach, make updates and modifications to the framework of analysis as more data are collected.

It is important to articulate the questions that need to be answered. For example, the question of whether the intended purpose of a VFAR tool development is monitoring, classification, screening, or risk assessment should be determined up front. Answering these questions may require different levels of detail, and different techniques may be more or less appropriate. It was suggested that proteomics might be particularly useful for screening, followed by a search for different genetic signatures that give analogous structures. Bioinformatics tools can be used to solve the question of the relevancy of genes and to predict structure-activity relationships.

There is an advantage in trying to develop these frameworks now to identify areas that require research. These concepts should work for viruses and other organisms. In fact, it may be advantageous to work with viruses because they are simpler organisms.

EPA emergency management staff are interested in the practical applications of these tools and, in particular, in opportunities for quick detection in the field, especially for engineered organisms. There is also a need for the scientific community to evaluate persistence to determine appropriate decontamination methods and develop cleanup levels.

As first conceived by the NRC, these questions define the framework for decision making, particularly with respect to weaponized agents, which are very different from naturally occurring agents. Naturally occurring microbes may or may not aerosolize, while weaponized agents, such as the agent used in the U.S. Senate anthrax event, are readily aerosolized.

EPA's Office of Water also is particularly interested in the rapid detection of microbes of concern. It is a challenging problem that needs a solution. From a public health standpoint, as one panelist illustrated, the search for a probe capable of identifying a contaminant and signaling an alarm prior to consumption of the water would potentially save lives by eliminating the time lapse required by current detection technology.

3. ***Discuss how VFARs can be used in the detection of recognized biothreat agents, newly emerging pathogens, and bioengineered pathogens.***

- **Analysis of VFs may indicate bioengineering, but may not be the sole indicator or focus of engineering.**

The analysis of VFs can provide information regarding genetic engineering for both bioweaponization and for naturally occurring genetic evolution. However, VFs may not be the focus of genetic engineering for the purpose of bioweapon development; there may be other characteristics that are altered to increase exposure and risk. For example, a gene or genes may be altered in a way that allows a microbe to persist in an environment, which will result in higher human exposures and lead to increased risk of disease. Persistence factors traditionally are not considered VFs.

VFARs can be used to identify biothreat agents, newly emerging pathogens, and bioengineered pathogens when applied to a surveillance system. However, it may be difficult to determine whether an agent was bioengineered based on VFARs alone. The approach for determining whether an agent has been bioengineered is classified, although, as panelists discussed, the approach goes beyond virulence factors to look at survival, the degree to which the agent can be cultured or stored, and other nongenetic factors. Genes are only one part of the equation.

The tools in use for bioengineering may have nothing to do with virulence; for instance, the bioengineering process may entail gene manipulation for eliciting a protein.

Similarly, bioengineering is not necessary for an intentional attack. For example in 1984, Salmonella was found in a salad bar in Oregon. It was a commercially available American Type Culture Collection (ATCC) strain, and initial efforts focused on determining whether the contamination was intentional. The use of VFARs or another genetic approach would not necessarily help in this type of investigation.

- **To use VFARs, the question needs to be defined — is it for detection or risk characterization?**

Discussion focused on the applicability of VFARs to detection and risk characterization as well as the overlap between the two. The emphasis on which elements are most important may be slightly different. Specifically, for applications pertaining to the detection of pathogens in the environment, the key factors of interest may be the array of VFs that are present. Other factors that may not be directly related to virulence are necessary for identification of pathogenic strains or species, based on their association with those strains or species. For quantitative risk assessment, one participant noted the need for a more rigorous definition of the relationship between the VFs and health effects. For qualitative applications in risk characterization, it may be possible to glean significant information based on the presence of VFs in the sample.

VFARs can be useful for the BioWatch program, which uses a series of pathogen detectors co-located with EPA air quality monitors. Currently, the BioWatch program is based on the collection of airborne particles on filters, which are removed and tested using PCR for the presence of select pathogens (Shea and Lister, 2003). The development of libraries of VF markers, coupled with more rapid and economical technologies, could facilitate the rapid identification of airborne pathogens of concern. Part of this surveillance could potentially be used to look for VFs as markers of bioengineering. The application of gene array technologies, rather than PCR, could yield test results within hours or, possibly, minutes. This type of technology transfer could fit into the surveillance and characterization of background conditions and for BioWatch applications.

Perhaps the most important role of VFARs, as has been identified by the NRC, is to characterize microbes that cannot be cultured and/or are novel to assess the potential for pathogenicity. New technologies discussed in this workshop can aid in the early detection of the presence of pathogenic microbes.

- **Bioengineering vs. Nature**

The explanation of engineered pathogens is complicated by the natural rapid changes that occur in the microbial genome. Rapid changes in this genome can come about via the transfer of plasmids. However, in terms of genetic engineering, the challenge is to get the specific proteins expressed, which involves the coordination of multiple genes, and is therefore a very complicated process. VFARs may have greatest applicability in developing an approach to screening genes associated with potential health effects.

- **What is the definition of the "VFAR Approach"?**

There was additional discussion regarding the meaning of the "VFAR approach" and whether there was a consensus as to its exact meaning. The VFAR approach, in its broadest sense, implies the creation of a database of VFs (descriptors), related health effects (response), and data analysis tools that relate and rank the pathogens (mathematical models for VFAR). Many different tools are being used to construct this database, though the VFAR approach should not, at this time, be limited to one technology, such as PCR or gene arrays. This will allow maximum flexibility for developing applications of VFARs, in terms of structure-activity relationships, that are parallel to those used in QSARs. While clearly the major goal is to use the structural relationship to identify and characterize pathogenicity and subsequent health effects (e.g., develop dose-response scenarios), VFARs can also be used for detection and hazard assessment through the identification of microbes that pose a potential health risk based on VF presence.

- **Can "most important" VFs be defined?**

VFs can provide critical information about a pathogen. However, given the limited knowledge available to the scientific community today and in light of data gaps, it is still a challenge to state what the "most important" VFs might be.

As with QSARs, application to dose-response assessment is still a major challenge. Based on what the scientific community knows about VFs, one participant encouraged testing hypotheses of VFAR application to various elements of the risk assessment paradigm. The participant presented a current example of the challenges inherent in monitoring

the risks associated with the evolving strains of bird flu. The participant asked how VFARs can be used as a tool for understanding the way risks change with the evolution of the virus. For example, despite all of the research on the Spanish flu of 1918, few clear-cut answers for explaining its virulence exist.

Another participant stressed the scientific community's need to know the biological characteristics that make organisms virulent. This likely involves VFs in conjunction with other factors, such as accessory genes and housekeeping genes. Once the characteristics of virulence are known, technologies can be developed for an early warning system that could be applied to both natural and terrorist events.

4. ***Describe technology available for examining virulence factors. How can the presence of such virulence factors in water or air be determined?***

- **Focus on technologies for identification and detection**

Tools and technologies available for examining virulence factors include genomics and gene arrays, high-throughput real-time PCR, and proteomics for the analysis of protein products. These technologies are constantly under development. All may be applicable to VFARs, but currently there are limitations in terms of sample collection and processing. Limitations include low concentrations in the environment, sample processing losses, and minimum detection limits associated with the molecular technologies. Such issues must be addressed before these technologies can be applied to surveillance in water or air.

- **Media-specific sampling issues**

There are media-specific problems with extraction of microbial material for the purpose of analysis and detection. Samples need to be processed, prepared, and concentrated. Water may be the simplest media with which to work. The greatest challenge is extracting the sample from the media for analysis. The amount of the sample must be sufficient for biological and statistical analysis. One panelist stressed that the scientific community ideally needs to be able to identify the biological agent in any given volume. To improve analyses, the sample may be concentrated or subjected to processing, depending upon the media from which it is taken. For drinking water, concentration typically is required. For surface water, some processing is needed in addition to concentration. Other media samples could require additional processing due to interference by other media constituents.

The closer the scientific community gets to the source of contamination, the easier it is to use molecular methods. Quantitative or reverse transcriptase PCR gives robust quantification capacity. It is useful for analyzing the presence of microbes in sewage and ground water, particularly those that are nonculturable. However, it requires prior knowledge of which VFs might be present so that appropriate oligonucleotide primers will be selected. It is important to solve sampling issues or look for targeted, specific genes or organisms. Sensitivity is improved if targeting specific organisms, but concentration or enrichment is often needed. As with chemical contaminants, it is difficult to determine transport, fate, and exposure concentrations. Furthermore, background levels of genes are not well characterized and will likely continue to be a problem until better data are available. Background occurrence of VFs can provide an important perspective on what constitutes a change in occurrence that signals the presence of a potential hazard.

Once it is understood what makes organisms unique to specific environments, it will be possible to target specific VFs within the organisms to determine whether exposure is occurring. Additional research and discovery is required to target for occurrence and exposure.

- **Use of -omics technologies in VFAR development**

The current focus in VFAR development is on genomics because of its sensitivity. Sometimes genomics is overly sensitive as the gene may be present, but not expressed. However, if the gene is not available, neither will be the message or protein.

It may be possible to use proteomics. Proteomics represents a complementary approach that can be initiated with the protein product, followed by an examination of the structural features of importance. Researchers can then work backwards, mining the genome for similar genes. With current technologies, proteomics may be more costly and time-consuming than other –omics technologies. In addition, there may be limited database availability for comparing and identifying proteins.

- **VF ranking and application in prioritization and risk assessment**

Although VFARs may be used in the prioritization of microbes and in microbial risk assessment, the databases to support such applications are still being developed. At the current time, it is not clear how VFs will be used to rank microbes and for application to risk assessment or prioritization.

5. ***Discuss the positive and negative applications of using VFARs in bioengineering. Discuss the construction of highly potent pathogens inserting single gene or combinations of virulence genes into commensal organisms. Do certain classes of virulence genes lend themselves to genetic engineering?***

- **Bioengineering vs. Nature revisited**

The changes that occur in the natural environment are an excellent example of how genetic factors change; however, genetic engineering is delicate. There are many examples in which genetic engineering resulted in unanticipated results. Most notably, microbes can transfer plasmids resulting in the rapid exchange of genetic material. The growing presence of antibiotic-resistant bacterial strains is an example. Also, some members of *Burkholderia* (earlier grouped under *Pseudomonads* that are generally known to be benign) are now of major concern to cystic fibrosis patients.

It has been demonstrated that *Pseudomonads* can be altered in the laboratory for various engineering applications. The simplest approach may be to co-culture organisms to facilitate the transfer of plasmids. The presence of genetic material

does not guarantee that it will be expressed. Additional steps, whether bioengineered or inherent to the microorganism, are required to translate the genetic code into proteins, and further modification may be required to ensure that the protein is functional. One participant recommended caution, citing an example in which genes were inserted into mouse pox with the intention of making a better vaccine; however, the resulting product was lethal.

- **Factors that change/increase potency**

Increases in potency are not always understood. In general, a change in potency is accompanied by a string of changes, not just a single change. It might not just be VFs that change to increase potency. There is a need to look for unusual combinations of genes, as well as other factors.

- **Host-specific effects**

Although extrapolation from animal studies introduces uncertainty, animal studies are, and will continue to be, an important avenue of research to identify potential human health risks. For some pathogens, outbreaks in other hosts precede infectivity in humans. Therefore, there needs to be an understanding of specific activity changes both in animals and in humans. These changes could occur on either the genotypic or phenotypic level. More importantly, the process of infecting a host can induce changes in the microbe. For example, in laboratory studies, passage through mice is frequently used to increase potency. In laboratory studies, passage through the animal is sometimes needed to identify new genes.

6. ***How can VFARs be used to determine the human toxicity potential of the virulent genes? Is it possible to obtain a quantitative estimate of the virulence along with a qualitative estimate?***

    - **The predictive capability comes from characterization and linkage to known health effects.**

Although it may currently be possible to begin to rank gene sequences, the capacity to link gene sequences to health effects is still being developed. For the purposes of public health protection, where it is necessary to gauge the response needed to prevent or mitigate an outbreak or reduce endemic disease, the goal is to be able to use VFARs to aid in the identification of the presence of microbes of concern, the prediction of the magnitude of the health hazard represented, and the determination of the infectivity or dose-response relationships. The scientific community needs to be able to answer questions such as, "How many people are likely to be affected?"

The current state of knowledge is focused on the identification of virulence factors, and how these virulence factors function in the microbe to explain virulence. As one panelist noted, the scientific community does not yet have the capacity to link this information to health outcomes, however the potential clearly exists.

An example of how these connections can be made is *Escherichia coli*. The 0157:H7 strain carries Shigella toxin and is much more virulent than other *E. coli* strains. By analyzing the genetics of this strain and comparing it to other *E. Coli* strains that lack Shigella toxin, the basis for strain potency can be developed. The characterization of these associations will drive the development of hypotheses regarding VF-activity linkages.

- **Virulence of organisms, not just genes**

It is not solely the virulence of genes that is important, but also the virulence of organisms (i.e., the genes need to be understood within the context of the organism). This is how VFs are tied to dose response. There is clearly a relationship between VF and dose response; however, dose response is more highly variable for biological agents than for chemicals.

Factors that contribute to the definition of the dose response of microbes include:

- Factors that control infectivity
- The evasion of the host immune system
- The ability to colonize within the host
- The initiation of the disease process

For example, poliovirus, in comparison to other disease viruses, requires high concentrations of the virus to initiate infectivity. In the case of poliovirus, the disease is not perpetuated at the site of infection, as may be the case with other viruses. The scientific community needs to understand the relationship between what happens at the site of infection and where the microbe exerts its health effects.

There are genetic factors that control all of these processes. The goal is to illustrate the relationship between VFs and dose response, recognize the complexity in this relationship among different organisms, and use the relationship as a proxy for the virulence of the organisms.

As with chemical exposures, variability of individual factors such as sex, age, and the presence or absence of chronic conditions, can be an important factor in host response. Furthermore, the genetic diversity of the immune system among individuals, which involves somatic mutations in the development of the specific immune response, increases the variability. Therefore, individual variability in terms of host response to biological agents is much broader and more challenging to characterize than it is for chemical agents.

The process of weaponization can be targeted at altering factors controlling dose response, including infectivity, evasion strategies, colonization, and pathogenicity. Successful bioengineering is not just a matter of altering genes alone. Gene expression and protein synthesis within the context of the organism are critical challenges.

- **Importance of exposure pathway and the relationship between exposure pathway and dose response**

As with chemicals, the exposure pathway is an important determinant of potential health effects. Anthrax exposure pathways, for example, include dermal absorption, ingestion, and inhalation, the last of which is the most potent. However, there is insufficient information on dose-response relationships via direct dermal and ingestion routes to determine the health impacts of anthrax via these routes.

Due to the lack of such dose-response data, anthrax inhalation dose-response relationship data are extrapolated to produce estimates for dermal and ingestion exposures. In addition to naturally pathogenic microorganisms, weaponized forms of microbes have the potential to alter exposure pathways and the associated dose-related response.

- **Pathogenicity is based on a complex set of factors, some related to the microbe, some related to the host.**

One panelist raised the need to define VFs more broadly because of related factors that confer or enhance virulence, pathogenicity, and persistence.

Further discussion from participants included several examples that provide insight into dose-response relationships, though it was acknowledged that in each case critical information was missing. For the anthrax contamination that occurred at the Washington, D.C., post office in 2001, the mortality rate was 1/20,000 (based on the exposed population), and only a small number of individuals became ill. Many of those exposed were treated prior to showing signs of infection, so it is not possible to measure infectivity. In Boca Raton, Florida, in 2001, many locations within a building tested positive for anthrax. One person died; however, no one else became ill even though the anthrax spores were presumed to have been present for many days. One woman in Connecticut died from exposure to anthrax-contaminated mail. In Philadelphia (1976), Legionnaires' disease was spread through the ventilation system and many people died.

In summary, the relationship between the organism and the host is extremely complex. As with chemical contaminants, there may be a threshold below which infectivity does not occur, while for others the threshold may be so low that it is negligible. In general, there is a lack of dose-response data; hence it is difficult to predict dose response, particularly at low-level exposures.

7. ***Can a virulence gene be altered so that it is still active but no longer detectable by the gene probes that are typically used?***

   - **VFs can be altered, but expression is not always predictable.**

It is possible for VFs to be altered so that they are still active but no longer detectable; however, oligonucleotide primers can be made for PCR and microarrays when alterations cannot be made without changing function.

Because of the degeneracy of the genetic code, alterations in the gene may be possible while preserving activity. With constant changes in the microbial genome, it is necessary to maintain surveillance for these changes and determine how they will affect virulence.

It is also possible that subtle changes over time will eventually affect the protein. The point at which activity actually changes depends on the organism, the protein, the specific function of the protein, and its biological interaction.

- **Specific tools have advantages and disadvantages in the identification of VFs.**

With carefully designed gene arrays using large numbers of probes, it may be easy to detect changes in the genome. Because microarray technology allows for multiple markers and probes, the probability of detecting genetic changes is increased as compared to PCR, which normally detects one gene target at a time.

To advance the use of VFs in the evaluation of health risks, discussion indicated the need for multiple VF descriptors. Gene occurrence and expression should be the initial descriptor. In addition, participants pointed out a need to characterize exposure routes (i.e., ingestion, inhalation, and dermal contact), survival and persistence, and attenuation in the environment. In addition, algorithms that relate genes to function need to be developed. With an initial focus on the use of VFARs to conduct quick screening, available data can be used to test the applicability of known VFs. However, researchers will need to develop computer models to determine the sensitivity of specific descriptors and the correlation of the descriptors to endpoints of concern.

It is possible that the scientific community has sufficient data to begin to develop a proof of concept that would take available data and demonstrate its applicability to detection, hazard identification, dose-response assessment, and risk characterization. For detection, a collection of VFs could be applied to predict the presence of pathogenic organisms in unknown samples. The analysis could also include predictions regarding potential sources. For risk assessment, the VFs could be used to qualitatively predict pathogenicity or health effects from the unknown samples. Although the results may not be fully accurate, these types of exercises could identify data gaps and help prioritize research to advance the field.

4.2 VFAR Closing Remarks

Factors that relate the virulence of microbes to adverse health effects should be determined. There are factors that control the ability of the microorganism to persist in a given environment, infect a host, evade the host's immune system, colonize within the host, and then initiate the disease process. These factors should be characterized. Factors may include receptor proteins, binding proteins, invasion capability, or toxin production. They may include components that aid in survival under different circumstances (e.g., in the presence of ultraviolet light or commonly used disinfectants such as chlorine). Many tools are available to characterize these factors, though more exist in the area of genomics than in proteomics.

The challenge lies in evaluating genes and proteins within the context of the organisms and their ecology. There are also important considerations regarding the manipulation of genes for the purposes of bioterrorism. In developing this understanding, the scientific community will be better able to identify and prioritize microbes to ensure the protection of human health.

VFARs also have the potential to provide important information for risk assessment. The identification of factors, genes, or proteins that confer an advantage to the microbe, which impacts pathogenicity, can assist with hazard identification and priority ranking, as well as the characterization of dose-response relationships under different exposure scenarios. Although panel members presented impressive current advances, there is a need to collect more data and develop analytical algorithms as the concept moves forward.

To advance the understanding of VFARs within the context of the microbes' ecology, researchers need to make an attempt to collect data during outbreak conditions. Doing so will help identify factors that were important in the outbreak, who will be affected by illness and why, the dose-response relationship, etc.

Given what is known now, there are opportunities to begin to test the concept of VFAR application. Although initial efforts will be challenging, they will help to identify critical data gaps for a more comprehensive study. Future efforts should begin with a broad definition of VFARs as factors that confer an advantage to organisms for their survival and success, identify background levels of known VFs, and track changes in the microbial community.

The focus should continue to be on the development of a set of tools, based on molecular techniques that can be used in the short- or medium-term to facilitate scanning for VFs. The level of stringency can be varied to collect a large amount of information in a short period of time, resulting in algorithm generation and analysis of the data to understand pathogenicity.

While it is expected that VFARs can help to prioritize microbes for the CCL, the existing datasets are not sufficiently robust for this application at the present time. However, since the concept is sound, development and testing of hypotheses to advance the science should be initiated.

5.0 QSAR Presentation Summaries

The following summarizes the presentations made for the QSAR concept. These include the integration of physiologically based pharmacokinetic (PBPK) modeling with QSAR models to reduce uncertainties in the chemical risk assessment process, the use of MOA and WOE in predictive toxicity, the application of reactivity as a descriptor in the development of more accurate QSAR models, a discussion of innovative and varied approaches to QSAR model development, and the role of a regulatory agency in advancing the development and implementation of QSARs. The discussion following each presentation is also summarized at the end of each presentation summary.

5.1 From Reactivity to Regulation: Integrating Alternative Techniques to Predict Toxicity

Mark Cronin, Professor of Predictive Toxicology, Liverpool John Moores University

Summary

Dr. Cronin began by describing the challenge reactive electrophilic compounds have posed to toxicologists in terms of identifying descriptors that accurately define their parameters and quantify their characteristics. Electrophilic chemicals are highly reactive and extremely toxic. Conventional QSAR methods consistently under-predict toxicity for this group of chemicals. Dr. Cronin stated that by using an enzyme assay, it is now possible to quantify electrophilicity to predict reactivity in biological systems. The assay is based on the chemical reaction with glutathione (GSH). There is a strong correlation between cytotoxicity and GSH reactivity. Quantification is based on the measurement of the reactivity index. Reactivity works well as a descriptor to rank a group of related chemicals based on this mode of action, however, it is still a challenge to translate this into a usable tool. As there is a spectrum of electrophiles, the first step is to define the domain, correlate it with toxicity, and model it. Dr. Cronin said that this process is expected to be particularly valuable under REACH and has direct application to regulatory issues. However, it will still require the use of multiple tools to characterize risk, and as with all chemicals, Dr. Cronin conceded, it is still a challenge to quantify uncertainty. The initial focus in the development of this process is to develop a model for fish toxicity and skin sensitization.

Discussion

It was noted that reactivity works well as a descriptor in ranking a group of related chemicals based on this MOA, but challenges remain to translate this into a usable tool. Other chemicals with different modes of action will require different descriptors.

Reactive chemicals may be metabolites, although, as workshop participants discussed, this is not currently the focus of the research. It should be possible to include a model of metabolism prior to GSH reaction.

5.2 Integrated QSAR – PBPK Modeling for Risk Assessment

Kannan Krishnan, Director of the Human Toxicology Research Group (TOXHUM), Université de Montréal.

Summary

Dr. Krishnan stated that based on the risk assessment paradigm, animal toxicity testing is evaluated to determine no observed adverse effect levels (NOAELs) for the derivation of risk-based criteria. QSARs can be used to predict the differential responses based on variation of chemical substitutents, but they are context specific and dependent on exposure route, rate, duration, etc. When conditions are varied, different QSARs need to be derived or extrapolations need to be made. The goal is improving derivation or extrapolation capabilities, and as Dr. Krishnan emphasized, integrating QSARs with PBPK modeling can do just that. PBPK models facilitate extrapolations of one of the two key components for the exposure-response relationship: pharmacokinetics (PK), representing external dose to internal dose; and pharmacodynamics (PD), representing tissue dose to effect. As components of dose response, PK and PD both can be related to QSARs to enhance extrapolation capability. Since there are more data available for PK, the focus of Dr. Krishnan's research is on the development of QSARs for PK profiles that change as a function of species, exposure route, dose, and duration. In the QSAR, given a set of related chemicals, the model begins with an administered dose and calculates changes in blood concentrations with chemical substituent changes. The model uses an easy-to-use spreadsheet to test how kinetics change with the related class of chemicals (using VOCs as a test case). The user enters chemical structure and duration of exposure into the spreadsheet to estimate tissue exposures, which will aid in the estimation of toxicity. The program will also allow modifications of exposure concentrations, routes, and exposure scenarios to evaluate how these changes impact tissue dose.

Discussion

Following Dr. Krishnan's presentation, discussion focused on the development of the PD component that incorporates MOA, which is in the early development stage. With the inclusion of MOA, prediction of effects should be possible. Gene microarray data, or other data relating to gene expression, cannot yet be incorporated.

For the physiological component the input variables are volumes, only. Partition coefficients in various tissues have been derived via *in vitro* testing. Tissue cultures, such as data from a liver slice can be used for QSAR development, though the scale of tissue levels must be increased for dose considerations.

5.3 Weight of Evidence and Mode of Action in Predictive Toxicology

Andrew Maier, Associate Director, TERA

Summary

Dr. Maier began by stating that the importance of QSARs is growing, in part, due to the incorporation of the concept into risk assessment methods using a WOE approach. WOE emphasizes decision making based on the totality of toxicological evidence. The WOE concept is being driven by improved biological understanding, such as knowledge of MOA, the increasing sophistication and validation of alternative study designs, and several quantitative tools, including SARs and QSARs. In the QSAR field, the consensus modeling concept embodies WOE principles. While WOE approaches use QSARs as an input for decision making, application of the results of the WOE can also be used as feedback in an iterative way to enhance the SAR and QSAR models. Another possibility for enhancing the QSAR concept is to link SARs and QSARs via the integration of MOA data. For many chemicals, the detailed mechanism of toxicity is not known, though the MOA data are available for a number of chemical classes. In lieu of waiting for a full mechanistic understanding, which will rarely be available, research should capitalize on the degree of biological understanding available to refine QSAR approaches. Several approaches for accomplishing this objective are available. On the simplest level, MOA data (including -omics data) provide a tool for interpreting the outputs of global QSAR methods. In addition, MOA data can be used to separate chemical groups using qualitative or quantitative decision-analysis approaches as an initial step in developing endpoint (or MOA-specific) QSAR models. MOA parameters can be used as chemical descriptors in building logistic regression models.

Discussion

Participants concluded that there is a need to consolidate what is known about chemical MOA to allow researchers to rank and prioritize their ability to integrate biology with QSAR. Biomarkers, particularly early effect biomarkers, can be useful in understanding the MOA for enhancing QSAR development. Both genomics and proteomics can be used as tools for MOA identification to aid in QSAR development or interpretation. These -omics technologies are also complementary with SARs and QSARs for reaching WOE conclusions for risk assessment.

5.4 Novel Approaches to QSAR and VFAR Modeling

William Welsh, Norman H. Edelman Professor in Bioinformatics and Computer-Aided Molecular Design, Department of Pharmacology, University of Medicine & Dentistry of New Jersey (UMDNJ)

Summary

Dr. Welsh stated that no one QSAR fits all, and that one way of dealing with this is by integrating consensus modeling, experimental data, bioinformatics, and -omics into WOE decision making. The New Jersey Environmental Bioinformatics and Computational Toxicology Center is developing computational toxicology tools such as Dose-Response Information Analysis System (DORIAN). There also are numerous chemical toxicology tools under development, including the following QSAR-based approaches:

- Decision forest, which makes predictions and evaluates prediction confidence
- Shape signatures, used for large-scale screening based on similarity in three-dimensional shape and bio-relevant surface properties
- Polynomial Neural Network (PNN), developing linear and nonlinear QSAR models
- Virtual High-Throughput Screening (VHTS) to assess the binding affinity of small-molecule compounds inside the positive binding pocket of protein receptors

The goal is to develop new methods that work in concert with established ones, while developing a hierarchy of strategies. The hierarchy will begin with fast, easy-to-use tools, such as structural filters and alerts, and then proceed to more computationally demanding tools such as classification models, followed by segregation using chemical activity. If the compounds are active, they will be selected for additional study. For example, within the decision forest, each tree includes a series of descriptors that segregate chemicals into active or inactive compounds. As the descriptors are independent, this results in consensus predictions. Each branch of the decision tree represents an "if-then" formatted query, thereby allowing for rapid evaluation. The shape-signature model begins with the molecule or receptor pocket. Shape and biorelevant features are converted into compact shape signatures for comparison. A data bank containing the shape signatures of greater than 5 million small-molecule compounds is then used to compare and contrast these features. A separate data bank for screening contains the shape signatures of more than 5,000 ligands extracted from the high-resolution X-ray crystal structures of proteins found in the publicly available protein data bank (PDB). The data bank is a repository for protein crystalline structures, and there is a library for screening. This process allows for the explanation of mechanistic clues of a molecule with an unknown MOA through comparison with chemicals in this PDB-extracted data bank of protein ligands. In theory, this process can also be applied to chemicals or proteins from bacteria of interest, such as *Escherichia coli*.

Discussion

The challenge is to move from structure to function. The next generation of shape signatures will tackle this problem, for example, distinguishing a receptor antagonist from an agonist.

The traditional application of three-dimensional QSARs requires subjective molecular alignment for the comparison of structures. With the new shape signatures program, the comparison is rotationally invariant, which removes subjectivity.

Other properties besides shape that can be used in the comparison include surface charge, polarity, hydrogen bonding capability, or any property that is mappable on the external structure of a molecule. Shape coupled with polarity has proven to work well.

In this schematic, there is a conformation generator and clustering tool that can generate multiple conformers and compare them using shape signatures. It may be more efficient to compare the shape signatures of clusters of conformers for a single molecule with multiple degrees of freedom. Shape signature works for a wide variety of molecular entities, including organic, inorganic, and organometallic molecules; neutral or charged species; proteins; and even nanoparticles.

5.5 Role of the European Chemicals Bureau in Promoting the Regulatory Implementation of Estimation Methods

Andrew Worth, European Chemicals Bureau, Institute for Health & Consumer Protection, Joint Research Centre, European Commission

Summary

The implementation of REACH legislation will depend on the efficient evaluation of chemicals of concern, using QSARs and methods for grouping chemicals. Authorities require that companies demonstrate the safe use of their chemicals. The WOE approach is needed, and animal testing is used only as a last resort. The focus is on developing the WOE approach by means of integrated testing strategies. If the model is scientifically validated and applicable to substances of interest, QSARs can be used for the purposes of classification and labeling and/or risk assessment, provided there is adequate and reliable documentation. The category approach can be used to group chemicals according to chemical similarity (e.g., structural properties, three-dimensional structure) to avoid the need to test every member of the group for every endpoint. Certain conditions apply; if categories are too large, it may not be applicable for every chemical, but the concept of subcategories is foreseen. The European Chemicals Bureau (ECB) is currently developing a guidance document on the use of grouping methods, including insights from the current practices of EU regulators, and introducing new approaches, such as computational toxicology and other new methods. All of the ECB's guidance development (which includes many other guidance documents for REACH) is conducted to be transparent to regulated industry, thereby permitting access to and use of the most advanced state of the science in preparing submissions for new chemicals. ECB is also building an online inventory of publicly available models, intended to be useful to EU industry and the future Chemicals Agency. The current emphasis is on model validation, documentation, consensus building, and capacity building.

Discussion

The adaptation of standard information requirements and the replacement of traditional test data using QSARs, reactivity data, -omics, etc. is a priority under REACH as a means of reducing animal testing. Integrated testing strategies based on a WOE approach will be used to combine the use of multiple approaches. Gaining consensus among industry organizations and 25 EU members on methods and approaches for risk assessment is extremely challenging.

To this end, QSARs must be scientifically validated and applicable to substance(s) of interest for the purposes of classification and labeling and/or risk assessment. In addition, adequate and reliable documentation must exist.

6.0 QSAR Charge Questions

The following summarizes the discussions on charge questions given to the expert panel. Discussion under each charge question is summarized as themes related to the question. Each theme is a bullet point under the charge question, followed by the summary discussion of the theme.

6.1 Summary of QSAR Charge Questions Discussions

1. ***In light of emerging technologies (e.g., genomics, proteomics, and bioinformatics), what role will QSAR methods play in the future with regard to EPA's risk assessment/risk management process?***

    - **It is important to have multiple tools for the evaluation of chemical toxicity.**

Participants expressed that any useful and valid information obtained through the application of emerging technologies will help to decrease uncertainty in the context of the overall weight of evidence. Genomics can aid in the identification of the MOA. For chemical reactivity, it is useful to have a genomic and proteomic fingerprint of the chemical since the genomic fingerprint may offer insight into a chemical's MOA. Some technologies may be better for screening than for regulatory decision making because they may be more readily validated, accepted, etc. Currently, the integration of QSARs with -omics technologies will result in an iterative approach, whereby these complementary technologies reinforce each other. Computational toxicologists are working on this integration to serve primarily as a hazard identification tool by providing insight into the potential chemical's MOA. Such knowledge can provide informed interpretation of QSARs.

There are several opportunities to combine QSARs and MOA information to better inform risk assessment, and members of the panel noted that routine acceptance of QSAR predictions will likely require that they be derived with an underlying mechanistic understanding. As models become more sophisticated, they will incorporate nontraditional structural features and property features and, therefore, allow for evaluating chemicals completely through the consideration of MOA data. Several examples of developments in this area were described. The integration of QSARs with PBPK modeling was discussed, wherein MOA considerations (e.g., identification of appropriate dose metrics based on chemical metabolism prediction) are factored into the PBPK model. The growing use of tools in bioinformatics (e.g., protein structure prediction and libraries) has allowed for the use of shape signatures based on the comparison of surface features to integrate MOA (e.g., receptor binding) into QSAR methodology. MOA data can be applied to larger groups of chemicals to identify clusters of more closely related chemicals. This is the conceptual basis for decision tree and regression tree approaches. QSAR models can be tailored via selection of descriptors for each cluster to provide more uniform training sets for QSAR development or aid in interpreting global QSAR predictions.

It was noted that -omics data have been applied in pharmacology and toxicology for the purposes of drug discovery, prognostic and diagnostic methods, biological pharmacological activity, and the toxicity-based landmark studies of John Weinstein (e.g., Bussey et al. 2006, Nishizuka et al. 2003, Blower et al. 2002). The foundation paper on this subject (Blower et al. 2002) reviewed the linkage between chemical and -omics technologies. However, there are inherent uncertainties in -omics technologies as well, in terms of interlaboratory variability and chip-to-chip errors. Judicious interpretation remains important in the use of -omics data as a supplement to or as an input into QSAR development. The field of single nucleotide polymorphisms (SNPs) is an exciting area of development that could provide information for QSARs. QSARs can also be used to understand -omics and focus on critical variables. This would, in turn, promote development of QSARs for critical molecules.

Currently, -omics data are not used directly as the primary basis for EPA risk assessment decisions, though they can lend support to the overall descriptions of toxicity mechanisms and are part of the Agency's risk assessment documents. In the EU, there is a placeholder in REACH legislation for the use of alternative methods, such as -omics technologies, either alone or in combination with other methods. -Omics approaches have yet to be standardized so that they are reproducible, and the need exists currently to categorize, document, and define these approaches.

- **One size does not fit all.**

Although the identification of a single technology for all chemical evaluation would greatly streamline the risk assessment process, no single technology can provide the necessary information for all chemicals. REACH requires consensus building and acceptance among industry and regulators. Toxicologists are often in a position in which they must explain that although QSARs may be easy to use, expertise and judgment are needed in the interpretation of the results. Increasingly, the concepts of consensus modeling and WOE are being incorporated into risk assessment guidance in recognition that no single technology is likely to provide all the answers.

- **Although QSARs can play an important role in risk assessment, there is a need to consider the WOE to evaluate chemicals.**

Panelists noted that the reliability of nontraditional risk assessment methods needs to be quantified. Even with 95 percent accuracy, the consequence of using incorrect predictions needs to be carefully assessed considering the

large number of chemicals that must be evaluated. When applied to human health outcomes, tolerance for uncertainty is very small and accuracy must increase to higher levels, such as >99 percent. All technologies have limitations, and no single method will provide all the data needed. With -omics, as with QSARs, there are layers of uncertainty — measurement error, unexpected patterns — and it may be very difficult to interpret the results. Altered gene expression does not necessarily mean that there is an effect. Most up and down regulated changes in genes are attributable to housekeeping genes. A fusion of technologies is needed and is occurring. Another panelist stressed the need to work together, maintain skepticism for all technologies, and verify the results, by considering the WOE, rather than focusing on one technology. There is a need to be transparent when communicating how the conclusions of a hazard/risk assessment depend on underlying results and the methods used to generate those results.

2. ***How can genomics, proteomics, and bioinformatics data be used in QSAR models? Are there examples where the -omics technologies in combination with QSAR models have proven to be able to predict, both qualitatively and quantitatively, acute/chronic toxicity across multiple chemical classes?***

    - **QSARs and -omics technologies are complementary and can be used to reinforce or refine estimates of toxicity.**

It was reiterated that -omics data have been applied along with QSARs in pharmacology and toxicology for the purposes of drug discovery, prognostic and diagnostic methods, biological pharmacological activity, and toxicity assessment for many years. The integration of QSARs with -omics technologies will result in an iterative approach, whereby these complementary technologies reinforce each other. Computational toxicologists are working on this integration.

One example of such iterative use of these technologies is that -omics data can help explain MOA and mechanisms of toxicity, which can then serve as inputs for defining QSAR parameters, building more closely aligned training sets or explaining variability in model predictions. Furthermore, data from -omics can be used as descriptors in QSARs. In theory, it should be both possible and useful to use data from -omics research as descriptors in QSARs. MOA descriptors may be informed by genomics and proteomics. Caution is needed in the use of genomics because genes that are transcribed may not necessarily be translated into functional proteins (e.g., post-translational modification). Proteomics data may provide more directly relevant information, but the experimental methods are more cumbersome. Metabolomics may fit more readily with the use of QSARs, but this growing area has not yet been fully explored in the context of QSAR application.

- **QSARs and -omics technologies can be particularly useful for hazard assessment.**

QSAR models can more readily predict a potential toxic outcome, which is equivalent to predicting hazard. If researchers are trying to develop correlations between exposure and hazard using -omics as an endpoint or outcome, this is a feasible approach. In other words, -omics can be used as biomarkers of exposure to identify hazard, which will feed into other elements of the risk assessment paradigm. A recent publication (Ekins et al. 2005), reviewed the use of Absorption, Distribution, Metabolism, and Excretion (ADME) and drug metabolism software to build in toxicogenomics, proteomics, metabonomics, and pharmacogenomics, using a systems biology approach. This is an example of integration that may work for chemical toxicology hazard assessment. To date, potency estimates (i.e., dose-response estimates) based on -omics have not yet been defined; hence, they have not been widely used in QSARs. -Omics data, therefore, remain largely a tool for MOA or hazard identification.

3. ***Since rule-based and expert models are based on congeneric groupings of chemicals (i.e., the training set is a congeneric data set), how can statistical models, which are generally based on noncongeneric training set, be improved? Can such models incorporate MOA data if available? Can such statistical models provide some insight regarding MOA for a chemical query?***

    - **Examples where MOA can be integrated into QSARs**

Panelists noted that, as discussed in an earlier presentation, QSARs can be integrated with PBPK modeling, where MOA is factored into the PBPK model. In addition, the use of shape signatures allows for the comparison of surface features and integrates MOA (e.g., receptor binding) into the methodology.

QSAR models are sophisticated, incorporating structural and property features. However, they should be sufficiently flexible to add MOA considerations directly into the model for chemical evaluation. Alternatively, given a large group of chemicals, one approach is to develop and apply MOA-based tools to subdivide chemicals into clusters. Global QSAR models can be developed for a variety of chemicals, or QSAR models can be tailored via selection of chemical clusters belonging to a certain chemical classification. Expertise is needed to make these decisions. For developing class- or cluster-based models, strict descriptor definitions are required for that class or cluster. A mechanism or MOA-based approach can be used to define these descriptors, but this can be challenging. Nevertheless, this approach has been successfully applied. For example, Knaak et al. (2004) integrated physicochemical and biological data for the development of predictive QSARs and PBPK models for organophosphate pesticides. In ecotoxicology, there have been examples where mechanisms of toxicity were generated from QSAR data. In addition, work has been published on the cytotoxicity of phenols, assigning modes of action and mechanisms of toxicity on the basis of QSARs (Schultz et al. 1997, Cronin et al. 2002).

- **QSARs for ecotoxicology are more widely accepted than in human health.**

It is easier to validate QSAR descriptors by experimentation in ecotoxicology than in human toxicology. There are existing databases (e.g., from studies in the EPA laboratories in Duluth) that facilitate QSAR development for aquatic toxicity

endpoints. Comparable databases for prediction of human health effects are sparse and not readily available. In addition, there is a lack of mechanistic data for the often more complex human health endpoints than for ecotoxicology endpoints. This is due, in part, to the existence of more mechanisms of action in human health endpoints. In the EU, an attempt was made to use a simpler classification of chemicals based on 17 different modes of action; however, the classification proved to be quite complex.

- **Development of appropriate and meaningful chemical grouping techniques requires knowledge of the model's purpose.**

There are thousands of descriptors available for each chemical. These descriptors, such as molecular weight and number of carbon atoms, can be physically meaningful or they can be physically uninterpretable constructs based on graph theory. To enhance the biological meaning for analysis, it is important to select methods that identify descriptors that are biologically meaningful and defensible. Developing QSARs based solely on statistical identification carries the potential risk of developing circumstantial correlations that may be highly predictive but biologically meaningless.

- **With the advent of toxicogenomics, the transfer of this technology to computational toxicology should help us understand the potential effects of chemicals on sensitive populations.**

Efforts are under way in pharmacology and toxicology to understand the interaction between variations in the human genome and variability in response to understanding how individual variability impacts chemical toxicity and risk assessment. Mechanistic QSARs can help define variations in chemical structure or properties that impact interactions with polymorphic receptors or xenobiotic metabolizing enzymes.

4. ***The toxicity of a chemical for any given health endpoint is, in general, due to an adverse interaction between the chemical and/or its metabolite and the tissue/organ/DNA associated with the endpoint. In developing statistically based QSAR models for chemicals with different modes of action, the descriptor pool contains descriptors that are chemical specific (i.e., they depend on the structure of the chemical alone). Are there any descriptors that can describe the tissue/organ/DNA characteristics and its interaction with a chemical and/or its metabolites?***

- **QSARs focus on describing the potential interaction between chemicals and biological molecules.**

There are two basic types of chemical-biological interactions. Receptor-based interactions often are the basis of endocrine disruption effects, and covalent interactions occur with nonspecific macromolecular binding. The latter are relatively nonspecific, but it is useful to focus on covalent interactions and characterize their diversity. This illustrates why endpoint-specific QSARs are useful. The specificity of target organs, where metabolism generally occurs, or the nature of cell/tissue type, provokes a reaction. Tissues introduce repair capacity, buffer capacity, etc., which modulates effects.

- **Metabolism is one of the keys to predictive success.**

Ideally, descriptors should relate to the toxic moiety (parent or metabolite). In cases where the toxic moiety is a metabolite, consideration of tissue characteristics related to metabolism (e.g., the presence of relevant metabolizing enzymes) can enhance model predictivity. For most endpoints, descriptors are not available to include relevant tissue characteristics. Statistical QSARs may implicitly include metabolism; however, metabolism will be correlated to various structural features. Software has been developed (Madden and Cronin, in press) to aid in the prediction of metabolites, although there are limitations inherent in the software. However, when applying such metabolism prediction models, biological understanding is still required to identify the metabolites associated with toxicity.

Complementary Ligand Based Receptor Interaction is a type of descriptor that considers the ligand and the receptor docking or binding. Researchers can use this descriptor and then screen potential ligands against known ligands. There are also models that account for the electron properties that map to the surface of DNA to model the binding of transcription factors to DNA.

- **Advancement of models that incorporate MOA and health effects data**

Although the pharmaceutical industry has been using mechanistic QSARs for years, these often have limited applicability outside the specific receptor or molecular endpoint being studied. Furthermore, much of the advanced work is proprietary. In terms of global QSARs, commercial software is available, but in many cases the underlying algorithms or databases are not transparent. Currently, there are research initiatives in chemical informatics (e.g., at Rensselaer University) to improve public domain data and modeling as well as software techniques. There are also nonpharmaceutical industry models available. Therefore, alternative approaches are needed to advance QSAR model applications that incorporate MOA and health effects data.

There are published examples of QSAR development in the literature pertaining to organic chemicals and human health. (Beliveau et al. 2005, Béliveau and Krishnan 2003, Waller et al. 1996). There are also examples in ecotoxicology; however, the endpoints are not likely to be highly relevant to human health (e.g., lethality).

Tissue microarrays are used in medical diagnostics to determine anticancer therapies and in the testing of drug cocktails, and these data are transferable to toxicological applications. Access to tissues from repositories would be required to generate experimental data from which QSAR models could be developed.

Physical and chemical descriptors can be used to predict interactions with a biological target as a pharmacodynamic (PD) approach. The scientific community needs to identify additional PD descriptors, although they may already be correlated, resulting in unnecessary redundancy. For example, given a QSAR model for breast cancer, researchers can add PD factors, including endocrine receptor (ER) binding and

prolactin release. The goal, essentially, is to model a series of steps that define a complex event.

- **Ratio of descriptors to compounds**

As a rule of thumb, the number of descriptors should be limited to 1 descriptor for 5 compounds. Thus, given 40 compounds, there should be no more than 8 descriptors. In the selection of descriptors, less is better. The QSAR equation describes a mathematical relationship that maps the target based on the descriptors. Descriptors may be correlated to the endpoint being predicted, but this does not indicate a causal relationship. In other words, a statistically derived QSAR may not be related to the pertinent MOA but may still accurately describe the relationship. To derive meaning from these types of descriptors may result in over-interpreting the model. In addition, since the QSAR models are mathematical equations — regardless of the chemical structure — the equations will predict some response. This is inconsistent with biological knowledge, where many chemicals will have no meaningful effect on certain endpoints. To overcome these problems in model parameter definition, approaches for selecting a few descriptors that may be most relevant from the MOA standpoint have been suggested.

5. ***Current methodology on the statistically based QSAR development for toxicity prediction calls for the inclusion of as many (classes of) descriptors in the descriptor pool as possible to explain the variance in the dependent variables (some measure of toxicity). In developing these QSARs, are there any (classes of) descriptors that one should definitely include in the potential descriptor pool (e.g., partition coefficients to account for transfer from blood to tissue)?***

    - **When selecting descriptors, start with the mechanistic context.**

Although certain descriptors are commonly used, the use of the mechanistic context as a starting point for the selection of descriptors is advisable. Since the mechanistic context varies based on chemical class, it is not possible to make blanket statements regarding the selection of descriptors. Examples of descriptors based on chemical mechanisms are factors that describe accumulation in a certain tissue (hydrophobicity), reactivity, receptor binding, etc. As described in an earlier presentation, modeling reactivity is difficult, and it is easy to miss subtleties. For example, given a reactive group on an aliphatic compound, if a stearic group is added near the reactive site, there will be stearic hindrance that is not captured using conventional descriptors. However, novel types of descriptors (e.g., atom-based fragments and certain fingerprint-based descriptors) may capture this information. To develop QSAR models, branching of groups that incorporate mechanistic-based descriptors may be needed to ensure that the molecule geometry has been adequately interpreted.

- **Graph theory as an alternative to define QSARs**

The use of graph theory to define QSAR descriptors is an important alternative when information about a chemical is lacking. These descriptors are easy to compute and are not subject to variability. One of the attractions in the use of descriptors derived from graph theory is that they are not conformation dependent, so researchers do not need to know anything about the conformation. However, these descriptors may not have any obvious mechanistic interpretation.

- **Integration of ligand-receptor interactions in QSAR models**

A recent evaluation of ligand-receptor interactions found significant differences in the properties of ligands. A recent article in the *Journal of Medicinal Chemistry* (Vol. 49, 3451–3453 [2006]) by a group of researchers from Ely-Lilly evaluated ligands that bind to different classes of proteins, such as kinases, nuclear receptors, and G protein-coupled receptors (GPCRs). They found differences in the properties of ligands. It would be interesting if the query were posed, "What is the target tissue?" and then have the software determine what descriptors have been successful for similar applications. There are instances where compounds were run against a panel of receptor proteins at single concentrations, creating a biospectrum of binding affinity. This can be used to characterize ligands.

Another option is to develop LOAEL models that are specific for specific endpoints, leading to the development of a suite of QSARs based on mechanistic considerations.

- **What is the status of the QSARs field (exploration vs. comparability and refinement)?**

Panelists noted that the answer to this question depends on where the researcher is in the life cycle of a given QSAR model. In regard to the development of QSARs, the initial stage can be characterized as exploratory — the gathering of data to develop correlations between chemical structure and outcome. As the field matures, models that are developed for different sets of related chemicals can be compared and refined. Most of the available models are in the comparability/refinement stage. As the available models continue to be advanced, the expectations for QSARs are very high. As more ideas are developed, models will be able to incorporate more complex biological considerations.

6. ***Qualitative SAR models (i.e., models yielding dichotomous or graded responses such as yes/no or low/med/high) do not provide a quantitative measure of a chemical's toxicity while quantitative SAR models (i.e., models yielding numerical potency estimates) do not provide a qualitative measure of the activity of a chemical for any given health endpoint. How does the panel view the feasibility of applying hybrid QSAR models (i.e., capitalizing on the benefits of SAR and QSAR by minimizing the disadvantage, if any, of each approach) for toxicity prediction? If feasible, how does the panel envision EPA applying such models?***

    - **Qualitative analyses can be useful for the purpose of comparing chemicals.**

Qualitative SAR comparisons may be useful for the hazard identification of chemicals with very little toxicity data.

Other qualitative analyses of SARs require the subjective classification of toxicity (e.g., low, medium, high), and there may be no biological significance. The context must be considered. It is better to have biologically meaningful classifications based on measurable biological events.

Hybrid analyses can also be applied; one type of hybrid analysis could begin with the initial classification based on MOA, followed by the application of the QSAR model. Semiquantitative QSARs can take the form of regression trees, using decision logic to inform the interpretation, such as binning or identifying the threshold of concern. There are also models based on quantitative relationships.

As experimental techniques improve, very low levels of toxicity (e.g., ER activity) can be measured. In some instances, QSARs are expected to quantify activity at such a low level that it is beyond the sensitivity of the model.

7. Can QSAR methods be used to reduce the uncertainty in extrapolating from acute and short-term benchmarks (such as LD_{50}) to subchronic and chronic LOAELs? What are the issues that must be addressed in order to do this?

- **There are distinct challenges in using QSARs to inform the extrapolation from acute to chronic effects because the critical endpoints are different.**

The extrapolation of subchronic to chronic exposure is frequently based on Haber's law, which states that concentration times duration is a constant, and this gives a ratio of exposure duration of about 10 (e.g., in rodents 800 days/90 days, which provides a rational definition for the extrapolation value of 10). This can be useful for the extrapolation of subchronic to chronic toxicity; however, it is inappropriate for the extrapolation from acute to chronic exposure because the critical endpoints are often different. Also, the MOA is different between acute and chronic exposure.

- **If there is knowledge about the critical effects, and MOA, then it may be possible to use QSARs for extrapolation and reducing uncertainty.**

It is possible that there are cases where the critical effects and MOA are the same, such that extrapolation using QSARs may be helpful. If there is commonality in MOA, then extrapolation from acute to chronic is more reasonable, but the rationale and the uncertainties must be discussed explicitly. Discriminators also can be segregated by MOA.

In particular, this may work for noncumulative reversible effects. If the target tissue and MOA are the same, this forms a basis to build an extrapolation algorithm. This information can also help decrease uncertainty. For instance, the default duration uncertainty factor in the EU is 100 (whereas it is 10 in the U.S.). Information from QSARs has been used to decrease the uncertainty factor. In other cases, it has been found that a safety factor of 100 is not adequately protective (research by Jan Ahlers, German Environmental Protection Agency).

There have been attempts to build models for PCBs, which tend to bioaccumulate, and the model incorporates the accelerating effects of the chemical over time. These models have had mixed success.

Ultimately, extrapolating data may be more of a science policy decision. Health Canada will not use acute data to derive subchronic or chronic values.

- **There is a critical need to evaluate thousands of chemicals that have no toxicity data, and all options should be evaluated.**

Participants acknowledged that many approaches have been suggested for evaluation of chemicals that lack toxicity data. Some think it is possible to take LD_{50}s and divide by some number and use this derived dose as a substitute for chronic effects. Others assert that since the MOA for acute effect is generally different from that of chronic effects, it is inappropriate to extrapolate from acute to chronic effects for most chemicals. In addition, communicating risk to the public can be challenging when acute to chronic extrapolations are performed.

The process of determining which chemicals should be on the CCL requires consideration of not only potency, but also severity. Using LD_{50}s does not seem to fit well into this paradigm.

Correlations have been done using regression analysis, primarily as a first-tier approach. This must be followed by an assessment of what is known about the chemical of interest and whether there are characteristics that can be used to make predictions. In short, expert judgment is required.

Regardless of the methods used, transparency, communication of assumptions, domain of applicability, and communication of uncertainties are critical components of any risk characterization.

- **The probabilistic derivation of QSARs could make better use of the available data.**

The application of probabilistic techniques for QSARs is feasible, using a range of data for each input. This may actually represent the best use of the available data. Monte Carlo approaches can then be used to generate a range of QSARs.

Also, in some databases, there are a number of measurements for any given endpoint. To develop a QSAR, decisions must be made on the selection of input values. Some may take the most conservative value, while other approaches will consider using the average. Guidance is extremely limited, and transparency is critical.

6.2 QSAR Closing Remarks

It is important to have multiple tools to use for the evaluation of chemical toxicity. The characterization of MOA can provide critical information regarding chemical toxicity. For chemical reactivity or cytotoxicity, it can be useful to have a genomic fingerprint of the response to the chemical to determine whether the observed effects represent different

gene-based responses. Some technologies may be better for screening than for regulatory decision making because they may be more readily validated, accepted, etc. An iterative approach between QSARs and -omic technologies can be used to reduce the uncertainties in each. In effect, a validated QSAR can be used to reduce uncertainty in -omic approaches and vice versa.

In theory, it should be possible to use -omics research data as descriptors in QSARs. MOA descriptors may be informed by genomics and proteomics. Caution is needed in the use of genomics because upregulated genes may not be expressed or functional (e.g., post-translational modification, etc.). Proteomics may provide more relevant information, and metabolomics may fit more readily with QSARs, but this has not yet been attempted. In addition, participants pointed out a need to consider dosing issues.

The integration of QSARs with PBPK modeling, in which MOA is factored into the overall framework, is possible and useful. The use of shape signatures allows for the comparison of surface features and integrates MOA (e.g., receptor binding) into the methodology. Models are much more sophisticated, incorporating structural features and property features; therefore, they should allow for more flexibility in evaluating chemicals by adding MOA considerations. A large group of chemicals can be subdivided into clusters. QSAR models can be developed globally for all chemicals in all clusters, or they can be tailored via selection of descriptors for each cluster.

The QSAR equation describes a mathematical relationship that maps the health endpoint to descriptors. Descriptors may be circumstantial. They may not be related to MOA, but they may be able accurately describe the relationship between a chemical structure and health endpoint. The use of graph theory, which is not dependent on conformation or biological interactions, to define QSAR descriptors is an important alternative when MOA information about the chemical is lacking. Nonmechanistic descriptors allow the problem of data gaps to be bypassed. However, to derive MOA meaning from these types of descriptors is to risk over-interpreting the model.

In some instances, the toxicity of metabolites has been incorporated in the original development of the equation. Although certain descriptors are more commonly used, a mechanistic context, if known, must be used as a starting point for the selection of descriptors. Since the mechanistic context varies based on chemical class, it is not possible to make blanket statements regarding the selection of descriptors.

A panelist pointed out that a few descriptors could be selected that may be most relevant from the MOA standpoint. Tissue characteristics, essentially static or defined, can be built in as a constant. These descriptors should relate to the metabolite, and tissue characteristics can be an important factor, particularly in terms of the concentrations of metabolizing enzymes, etc. For most endpoints, descriptors are not yet available to build in tissue characteristics, although some approaches do implicitly include metabolism. The pharmacology industry must routinely make predictions about metabolism in order to predict toxicity and the possible cellular targets (i.e., DNA, protein, etc.) when selecting descriptors.

Qualitative applications of QSAR analysis are possible and can provide important information for hazard identification. One type of hybrid analysis could be the initial classification based on MOA, followed by the application of a QSAR model. Semiquantitative QSARs can take the form of regression trees, using decision logic to inform the interpretation, such as binning and identifying the threshold of concern.

7.0 Major Considerations and Recommendations

Discussion of the VFAR and QSAR charge questions gave rise to the following major considerations:

- Because technology allows for a very broad array of gene identification, there is no need to omit any classes of VFs from consideration in the initial development of VFAR methodology. Such elimination should occur only when the irrelevance of the VF can be demonstrated. In addition, the presence of a VF may be necessary but not sufficient for the development of pathogenicity. There are other factors, such as those that permit the expression of VFs, the survival and persistence of the microbes, or even a particular array of microbes in the environment, that permit the development of pathogenicity. There also is an urgent need to characterize background levels of VFs in the environment to better recognize a change in conditions that may pose a human health risk.
- The analysis of VFs can provide information regarding genetic engineering for changes occurring due to both bioweapons and naturally occurring genetic evolution. However, VFs may not be the focus of genetic engineering for the purpose of bioweapon development. There may be other characteristics that are altered to increase exposure and risk.
- There are many tools and technologies available for examining VFs, including genomics and gene arrays, PCR, and proteomics for the analysis of protein products. These technologies are all under development in terms of their applicability to VFARs, but there are limitations due to sample collection and processing issues that must be addressed before these technologies can be applied to surveillance in water or air.
- Genetic changes that occur naturally are an excellent example of the flexibility of the microbial genome. Most notably, microbes can transfer plasmids, resulting in the rapid exchange of genetic material. Increases in potency are not always understood. There is a need to look for unusual combinations of genes as well as other factors. In general, a change in potency is accompanied by a string of changes, not just a single change.
- For the purposes of public health protection, the goal is to be able to use VFARs to aid in the:
    - Identification of the presence of microbes of concern
    - Identification of accessory genes necessary for virulence
    - Identification of environmental conditions necessary for virulence
    - Extrapolation from virulence gene expression to virulence protein expression
    - Prediction of the magnitude of the health hazard represented
    - Determination of the infectivity or dose-response relationships to gauge the response needed to prevent or mitigate an outbreak

These characterizations and predictions would provide information critical to understanding the magnitude of the public health risk associated with a natural or intentional exposure event.

- The current state of knowledge is focused on the identification of VFs and how these VFs function in the microbe to express virulence. The scientific community does not yet have the capability to link VF information to health outcomes, though the potential exists. Because of the degeneracy of the genetic code, it is possible for there to be alterations in the gene while its activity is preserved. With constant changes in the microbial genome, it is necessary to maintain surveillance for these changes and evaluate how they affect virulence. It is possible to make primers for areas where changes cannot be made without changing function, thereby minimizing the chance of missing known VFs.
- For both chemical and biological threats to human health, the universe of microbes and chemicals needs to be characterized and narrowed for the purposes of regulatory prioritization and development of remedial action strategies. Also, for both approaches to be effective, either the MOA or mechanism of toxicity must be determined. This is an essential component of expert system based structure-activity relationships where the aspect of the structure of the chemical that results in a particular effect or outcome must be determined. This concept can greatly enhance QSAR model development and interpretation.
- In terms of the role of -omics and QSARs in EPA's framework for risk assessment, any useful and valid information will help decrease uncertainty in the context of the overall WOE. Some technologies may be better for screening than for regulatory decision making because they may not be fully validated or accepted. QSARs and -omics technologies fit into this category. Currently, genomics technologies primarily serve as hazard identification tools by providing insight into the potential MOA by which a chemical is acting. Such knowledge can inform the interpretation of QSARs. The integration of QSARs with -omics technologies may allow these complementary technologies to reinforce each other. Computational toxicologists are working on this integration.

- There are several opportunities to combine QSARs and MOA information to better inform risk assessment, and the panel noted that routine acceptance of QSAR predictions will likely require that they be derived with an underlying mechanistic understanding. As models become more sophisticated, they will further incorporate structural features and property features to allow for evaluating chemicals more fully through the consideration of MOA data. Several examples of developments in this area were described. Dr. Kannan Krishnan discussed the integration of QSARs with PBPK modeling, where MOA considerations (e.g., identification of appropriate dose metrics based on chemical metabolism prediction) are factored into the PBPK model. Dr. Welsh discussed the growing use of tools in bioinformatics (e.g., protein structure prediction and libraries). Such tools have allowed for the use of shape signatures based on the comparison of surface features to integrate MOA (e.g., receptor binding) into QSAR methodology. MOA data can be applied to larger groups of chemicals to identify clusters of more closely related chemicals — this is the conceptual basis for decision tree and regression tree approaches. QSAR models can be tailored via selection of descriptors for each cluster to provide more uniform training sets for QSAR development or aid in interpreting global QSAR predictions.
- The focus of QSAR is on describing the potential interaction between chemical and biological molecules. There are two basic types of chemical-biological interactions. Receptor-based interactions often are the basis of endocrine disruption effects, and covalent interactions occur with nonspecific macromolecular binding. Mechanistic QSARs for predicting receptor-based interactions are commonly used in drug development and are increasingly being used for toxicity prediction. Nevertheless, many chemicals act via the disruption of membranes. The latter are relatively nonspecific, but it is useful to focus on covalent interactions, which can be quite complex, even within a chemical class, as was highlighted in the context of phenolic electrophiles. To be most useful, QSARs need to account for this complexity more fully. While mechanistic QSARs are preferred, an intermediate step in this direction is to focus efforts on endpoint-specific QSARs since the specificity of target organs can arise based on adsorption and distribution (toxicokientics) or the nature of cell/tissue response (toxicodynamics).
- Although certain descriptors (i.e., molecular size and hydrophobicity) are more commonly used, the mechanistic context must be used as a starting point for the selection of descriptors. Since the mechanistic context varies based on chemical class, it is not possible to make blanket statements regarding the selection of descriptors. Examples of descriptors based on chemical mechanisms are those descriptors that describe accumulation at or penetration through the membrane, reactivity toward cellular macromolecules, or receptor binding with critical targets, and others.
- Several approaches for hybrid SAR/QSAR analyses were discussed. Approaches ranged from using MOA descriptors as a screening step for the initial classification of chemicals to help in interpreting global QSARs to direct use of MOA descriptors in developing quantitative endpoint-specific logistic regression models. Semiquantitative QSAR methods included decision trees or modifications of this concept that use parallel sets of decision trees to improve predictability. Binned chemicals identified through these tools could serve as endpoint-specific QSAR training sets or be used to identify characteristics associated with potency categories for risk assessment using threshold of concern approaches.
- For chemical risk assessment, there is often a need to extrapolate from dose-response data based on exposure durations of less than a lifetime to estimate the effects of lifelong exposure. Traditionally, for EPA risk assessments, a default factor of 10 is applied to adjust adverse effect levels from subchronic (i.e., exposure for roughly 10 percent of the lifetime) to chronic exposure conditions. This can be useful for the extrapolation of subchronic to chronic toxicity; however, it is inappropriate for the extrapolation from acute to chronic exposure because the critical endpoints are often different and the MOA is different between acute and chronic exposure. The panel noted that while several correlation approaches have been developed to address this situation, these are not QSARs *per se*. While QSARs may address this application directly, they can provide important insights. For example, QSARs are used to predict toxicokinetic parameters (e.g., partition coefficients or metabolism parameters) that impact decisions regarding the potential for increased body burden with longer-duration exposures. Furthermore, QSARs can provide information pertaining to both acute and chronic toxicity mechanisms, which impacts considerations of potential for accumulation of tissue damage with increased exposure duration.

From the discussion of these charge questions came the following major recommendations:

1. Several recommendations on near-term applications of VFAR/QSAR models were discussed. To advance the applicability of VFARs in real-world situations, it is critical to facilitate the analysis of samples collected during natural outbreaks of microbial diseases. This will permit the identification of background levels of VFs and advance the understanding of the natural evolution of VFs in addition to providing the framework to test predictions of VFARs. Another potential opportunity for the advancement of VFAR research involves the BioWatch Program, which consists of continuous sampling at locations across the country. This would be an opportunity for researchers to obtain material for the characterization of background levels of VFARs in urban environments, in addition to

testing hypotheses. The state of the science regarding QSAR modeling is considerably more advanced than that of VFARs, therefore the key recommendation for near-term applications focused on the integration of MOA and PBPK with QSAR models to enhance biological applicability.

2. For both VFARs and QSARs, host-specific factors alter the dose-response relationship (e.g., individual variability in metabolism, sensitive subpopulations, and host immune response); therefore, there will always be uncertainty in the ability to model host factors. These limitations should not be a deterrent from using these approaches in the evaluation of the universe of chemicals and microbes that require attention. For the initial prioritization of chemicals or microbes, when toxicological data are lacking, QSARs and VFARs can be particularly useful. Similarly, the databases and models under development could be critical to facilitating a rapid response in the event of an intentional attack. QSARs and VFARs can provide critical information regarding alerts to human health concerns, and chemical and biological plausibility in terms of potential human health effects, particularly as input to comprehensive WOE approaches.

3. For the initial prioritization of chemicals or microbes, when toxicological data are lacking, QSARs and VFARs can be extremely useful. Similarly, the databases and models under development could be critical to facilitating a rapid response in the event of an intentional attack. However, as noted by the expert speakers, these methods may not be sufficient for all chemicals or all microbes. One panelist charged that all the tools available should be used to begin to address these urgent public health concerns. QSARs and VFARs can be important tools in characterizing human health risks based on the weight of the evidence. Both QSARs and VFARs can be used to advance understanding of potential human health effects as well as in the regulatory context to help prioritize chemicals of concern. How EPA applies those concepts will likely vary by EPA division. A similar process is occurring in the EU.

4. Other panelists urged that to move this discipline forward, single QSAR or VFAR predictions should not be considered an answer. Rather, consensus or WOE approaches result in a more robust analysis. It is critical to be able to demonstrate how QSAR and VFAR tools can contribute to an understanding of health risks by providing information on hazard assessment and dose-response relationships.

5. As a result of the discussions, participants noted the creation of more questions. It is becoming more common to develop handbooks and guidelines to derive the necessary components. The field is very dynamic and needs virtual and enhanced screening in addition to genomics. QSARs and VFARS are, and will always be, two tools among many.

8.0 References

Ashby J, Tennant RW. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutat Res*. 1988 Jan; 204(1): 17–115. Review. http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&Cmd=ShowDetailView&TermToSearch=3277047&ordinalpos=382&itool=EntrezSystem2.PEntrez.Pubmed.Pubmed_ResultsPanel.Pubmed_RVDocSum

Béliveau M, Lipscomb J, Tardif R, Krishnan K. 2005. Quantitative structure-property relationships for interspecies extrapolation of the inhalation pharmacokinetics of organic chemicals. *Chem Res Toxicol*. 18(3): 475–85.

Béliveau M and Krishnan K. 2003. *In silico* approaches for developing physiologically based pharmacokinetic (PBPK) models. In: H Salem and S Katz, eds. *Alternative Toxicological Methods*. CRC press, NY, 479–532.

Blower PE, Yang C, Fligner MA, Verducci JS, Yu L, Richman S, Weinstein JN. 2002. Pharmacogenomic analysis: correlating molecular substructure classes with microarray gene expression data. *Pharmacogenomics J*. 2002; 2(4): 259–71.

Bussey KJ, Chin K, Lababidi S, Reimers M, Reinhold WC, Kuo WL, Gwadry F, Ajay, Kouros-Mehr H, Fridlyand J, Jain A, Collins C, Nishizuka S, Tonon G, Roschke A, Gehlhaus K, Kirsch I, Scudiero DA, Gray JW, Weinstein JN. 2006. Integrating data on DNA copy number with gene expression levels and drug sensitivities in the NCI-60 cell line panel. *Mol Cancer Ther*. 2006 Apr; 5(4): 853–67.

Cronin MTD, Aptula AO, Duffy JC, Netzeva TI, Rowe PH, Valkova IV, Schultz TW. 2002. Comparative assessment of methods to develop QSARs for the prediction of the toxicity of phenols to Tetrahymena pyriformis. *Chemosphere* 49: 1201–1221.

Ekins, Nikolsky and Nikolskaya in *Trends in Pharmacological Sciences* Vol. 26, No 4, April 2005.

Gray LE Jr. and Ostby J. 1993. The effect of prenatal administration of azo dyes on testicular development in the mouse: A structure activity profile of dyes derived from benzidine, dimethylbenzidine, or dimethoxybenzidine. *Fundamental and Applied Toxicology*: 20, 177–183.

Harada A, Hanzawa M, Saito J, Hashimoto K. 1992. Quantitative analysis of structure-toxicity relationships of substituted anilines by use of Balb/3T3 Cells. *Environmental Toxicology and Chemistry*, Vol. 11: 973–980.

Knaak JB, Dary CC, Power F, Thompson CB, Blancato JN. 2004. Physicochemical and biological data for the development of predictive organophosphorus pesticide QSARs and PBPK/PD models for human risk assessment. *Crit Rev Toxicol*. 34(2): 143–207.

Lewis DFV, Ioanides C, and Parke DV. 1993. Validation of a novel molecular orbit approach (COMPACT) for the prospective safety evaluation of chemicals, by comparison with rodent carcinogenicity and Salmonella mutagenicity data evaluated. *Mutat Res*. 291: 61–77.

Lowell HH, Maynard EL, Kier LB. 1989. Structure-activity relationship studies on the toxicity of benzene derivatives: III. Predictions and extension to new substituents. *Environmental Toxicology and Chemistry*, Vol. 8: 431–436.

JC Madden, MTD Cronin. 2006. Structure-based methods for the prediction of drug metabolism. *Expert Opinion on Drug Metabolism and Toxicology 2*: in press.

National Research Council (NRC). 1999. Identifying Future Drinking Water Contaminants Based on the 1998 Workshop on Emerging Drinking Water Contaminants. Water Science and Technology Board, Board on Environmental Studies and Toxicology. National Academy Press, Washington, DC.

National Research Council (NRC). 1999. Setting Priorities for Drinking Water Contaminants. Committee on Drinking Water Contaminants, Water Science and Technology Board, Board on Environmental Studies and Toxicology. National Academy Press, Washington, DC.

National Research Council (NRC). 2001. Classifying Drinking Water Contaminants for Regulatory Consideration. Committee on Drinking Water Contaminants, Water Science and Technology Board, Board on Environmental Studies and Toxicology. National Academy Press, Washington, DC.

Nishizuka S, Charboneau L, Young L, Major S, Reinhold WC, Waltham M, Kouros-Mehr H, Bussey KJ, Lee JK, Espina V, Munson PJ, Petricoin E 3rd, Liotta LA, Weinstein JN. 2003. Proteomic profiling of the NCI-60 cancer cell lines using new high-density reverse-phase lysate microarrays. Proc Natl Acad Sci USA. 2003 Nov 25; 100(24): 14229–34.

Rosenkranz HS and Klopman G. 1989. Structural basis of the mutagenicity of phenyazoaniline dyes. *Mutat Res*. 221: 217–234.

Schultz TW, Sinks GD, Cronin MTD. 1997. Identification of mechanisms of toxic action of phenols to Tetrahymena pyriformis from molecular descriptors. In: Chen F and Schuurmann G (Eds) *Quantitative Structure-Activity Relationships in Environmental Sciences* - VII. SETAC Press, Pensacola, USA, 329–342.

Shea, DA and Lister S. 2003. The BioWatch Program: Detection of Bioterrorism. Congressional Research Service Report No. RL 32152, November 19, 2003. http://www.fas.org/sgp/crs/terror/RL32152.html

U.S. Environmental Protection Agency (EPA). 1994. Assessment Tools for the Evaluation of Risk (ASTER). On-line Database. Environmental Research Laboratory-Duluth.

U.S. Environmental Protection Agency (EPA). 1992. Provisional Guidance for the Qualitative Risk Assessment of Polycyclic Aromatic Hydrocarbons. Prepared by the Environmental Criteria and Assessment Office, Office of Health and Environmental Assessment, Cincinnati, OH, for the Office of Research and Development, Cincinnati, OH.

U.S. Environmental Protection Agency (EPA). 1989 Update to the Interim Procedures for Estimating Risk Associated with Exposures to Mixtures of Chlorinated Dibenzo-*p*-Dioxins and -Dibenzofurans (CDDs and CDFs). Risk Assessment Forum, Washington, DC.

Waller CL, Evans MV, McKinney JD. 1996. Modeling the cytochrome P450-mediated metabolism of chlorinated volatile organic compounds. *Drug Metab Dispos*. 24(2): 203–10.

Weisburger EK. 1979. *N*-Substituted aryl compounds in carcinogenesis and mutagenesis. Presented at the International Conference on Carcinogenic and Mutagenic *N*-substituted Aryl Compounds, Rockville, Maryland.

Weisburger JH and Fiala ES. 1979. Mechanisms of species, strain, and dose effects in arylamine carcinogenesis. Presented at the International Conference on Carcinogenic and Mutagenic N-substituted Aryl Compounds, Rockville, Maryland.

Appendix A
List of Speakers

Andy Avel
Assistant Center Director
U.S. EPA, Office of Research and Development
National Homeland Security Research Center
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7951
Email: avel.andy@epa.gov

Mark Cronin, Ph.D.
School of Pharmacy and Chemistry
Liverpool John Moores University
Byrom Street
Liverpool, England L3 3AF
Phone (from UK): 0151 231 2402
Phone (from outside UK): + 44 151 231 2402
FAX (from UK): 0151 231 2170
FAX (from outside UK): + 44 151 231 2170
Email: m.t.cronin@ljmu.ac.uk

Syed A. Hashsham, Ph.D.
Edwin Willits Associate Professor
Department of Civil and Environmental
Engineering Center for Microbial Ecology
Michigan State University
A126 Research Complex-Engineering
East Lansing, MI 48824
Phone: 517-355-8241
FAX: 517-355-0250
Email: hashsham@egr.msu.edu

Jonathan Herrmann, P.E., DEE
Center Director
U.S. EPA, Office of Research and Development
National Homeland Security Research Center
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7839
Email: herrmann.jonathan@epa.gov

Kannan Krishnan, Ph.D.
Professeur titulaire et Directeur TOXHUM
Université de Montréal
2375 Cote Ste. Catherine, Room 4105
Montreal, PQ, Canada, H3T 1A8
Phone: 514-343-6581
FAX: 514-343-2200
Email: kannan.krishnan@umontreal.ca

Andrew Maier, Ph.D., CIH, DABT
Associate Director
Toxicology Excellence for Risk Assessment
2300 Montana Avenue, Suite 409
Cincinnati, OH 45211
Phone: 513-542-7475 x23
FAX: 513-542-7487
Email: maier@tera.org

Chandrika Moudgal, M.S.
U.S. EPA, Office of Research and Development
National Homeland Security Research Center
Threat and Consequence Assessment Division
1001 SW 5th Avenue, Suite 1510
Portland, OR 97204
Phone: 503-326-3541
FAX: 503-326-4005
Email: moudgal.chandrika@epa.gov

Joan B. Rose, Ph.D.
Homer Nowlin Chair in Water Research
Department of Fisheries and Wildlife
Michigan State University
13 Natural Resources
East Lansing, MI 48824
Phone: 517-432-4412
Fax: 517-432-1699
Email: rosejo@msu.edu

R. Paul Schaudies, Ph.D.
Assistant Vice President
Science Applications International Corporation
Biological and Chemical Defense
9700 Great Seneca Highway, Suite 220
Rockville, MD 20850
Phone: 240-453-6312
FAX: 240-453-6208
Email: schaudiesr@saic.com

Subhas Sikdar, Ph.D.
Acting Associate Director for Health
U.S. EPA, Office of Research and Development
National Risk Management Research Laboratory
26 West Martin Luther King Drive (MS 235)
Cincinnati, OH 45268
Phone: 513-569-7528
Email: sikdar.subhas@epa.gov

Cindy Sonich-Mullin, M. En.
Division Director
U.S. EPA, Office of Research and Development
National Homeland Security Research Center
Threat and Consequence Assessment Division
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7923
Email: sonich-mullin.cynthia@epa.gov

Gerard Stelma, Ph.D.
U.S. EPA, Office of Research and Development
National Exposure Research Laboratory
26 West Martin Luther King Drive (MS 593)
Cincinnati, OH 45268-1320
Phone: 513-569-7384
Email: stelma.gerard@epa.gov

William J. Welsh, Ph.D.
Norman H. Edelman Professor in Bioinformatics
Department of Pharmacology Robert Wood Johnson Medical School University of Medicine & Dentistry of New Jersey (UMDNJ)
Director
UMDNJ Informatics Institute
Director
UMDNJ Environmental Bioinformatics & Computational Toxicology Center
675 Hoes Lane
Piscataway, NJ 08854
Phone: 732-235-3234
FAX: 732-235-3475
Email: welshwj@umdnj.edu

Andrew Worth, Ph.D.
Scientific Officer
European Commission – Joint Research Centre
Via Enrico Fermi 1
21020 Ispra (VA), Italy
Phone: +39 0332 789566
FAX: +39 0332 786717
Email: andrew.worth@jrc.it

Douglas Young, Ph.D.
Branch Chief
U.S. EPA, Office of Research and Development
National Risk Management Research Laboratory
Clean Processes Branch
26 West Martin Luther King Drive
Cincinnati, OH 45268-1320
Phone: 513-569-7624

Appendix B
Biosketches of Speakers and Panelist

Mr. Andy Avel started his career in 1972 as an engineering geologist for the U.S. Tennessee Valley Authority and over the following ten years was assigned in Chattanooga, Kingsport, and Knoxville, TN. He joined the Clinch River Breeder Reactor Plant Project in Oak Ridge, TN, in 1982, as a geotechnical engineer. Upon cancellation of the Breeder Reactor, Andy moved to the Department of Energy's Office of Civilian Radioactive Waste Management in Columbus, OH, where he served as a licensing engineer. He returned to Oak Ridge as a project manager in the Formerly Utilized Sites Remedial Action Program and then moved to the Fernald Feed Materials Production Plant, near Cincinnati, where he managed the CERCLA cleanup program.

Mr. Avel joined ORD in 1991 as the Director of the Office of the Senior Official in Cincinnati. Following the reorganization of 1996, he was assigned as special assistant to the Director of NRMRL and then as the Acting Lab Director (ALD) for Pesticides and Toxic Substances. In November of 2002, Andy joined the National Homeland Security Research Center as the Deputy Director for Management. In January 2005, he was named Acting Director of NHSRC.

Dr. Mark Cronin is Professor of Predictive Toxicology in the School of Pharmacy and Chemistry at Liverpool John Moores University, England. He was previously a lecturer (from 1994) and reader (from 2001) in that department. In addition to a full teaching load on the Master of Pharmacy degree course, he maintains an active research focus on the development of computational methods to predict toxicity. Particular emphasis at the moment is on the prediction of reactive toxicity (e.g., skin sensitization) and the use of quantitative structure-activity relationships (QSARs) for regulatory purposes. He has over 150 publications in these areas and has co-organized a number of conferences in predictive toxicology. Mark obtained his degree in Biology and Ph.D. in ecotoxicological QSAR from Liverpool Polytechnic.

Dr. Syed A. Hashsham is Edwin Willits Associate Professor of Civil and Environmental Engineering at Michigan State University (MSU). He is also a Co-Principal Investigator (PI) in the Center for Microbial Ecology and CAMRA, the U.S. EPA/DHS Center for Advancing Microbial Risks Assessment. Syed's expertise is in the area of environmental genomics and modeling of molecular data with a focus on microbial issues related to drinking water and wastewater. His research work is sponsored by the NIH, EPA, DHS, DoD, NSF, and state agencies. He has published on DNA biochip-based parallel microbial detection (*Biosensors & Bioelectronics*, 2004), VFAR (*Water Science and Technology*, 2004), microbial community dynamics (*Applied and Environmental Microbiology*, 2000), probe design (*Nucleic Acids Research*, 2006) and dehalo-respiration (*Science*, 2002). Syed earned his Ph.D. in Environmental Engineering and Science from the University of Illinois at Urbana-Champaign and conducted post-doctoral research at the Center for Microbial Ecology at MSU and Stanford University.

Mr. Jonathan Herrmann has been with EPA since 1975. He first worked in the Agency's Region VIII office in Denver, Colorado. He came to the EPA's Office of Research and Development (ORD) in 1978 and, except for a brief time in the private sector in the early 1980s, has been with ORD in Cincinnati, OH. Mr. Herrmann holds a Bachelor's Degree in Civil Engineering from Youngstown State University and a Master's Degree in Business Administration from Xavier University. He is a Registered Professional in Engineering in the State of Ohio. He is a member of the American Society of Civil Engineers, the American Academy of Environmental Engineers, and the American Water Works Association.

Mr. Herrmann's career has spanned many areas. He has worked in mined land reclamation, Superfund site remediation, land disposal of hazardous and household wastes, and environmental technology testing and evaluation. In the mid-1990s he was a strategic planner for the National Risk Management Research Laboratory and lead the development of ORD's Pollution Prevention Research Strategy and Mercury Research Strategy.

Mr. Herrmann joined NHSRC in September 2002 as the Water Security Team Leader and with a group of scientists and engineers developed the Water Security Research and Technical Support Action Plan in cooperation with the Agency's Office of Water (OW). He is currently serving as the Center Director for NHSRC. As such, he is responsible for the day-to-day personnel, funding, and product delivery aspects of the Center.

Dr. Kannan Krishnan received his Ph.D. in Public Health from Université de Montréal, Canada, and postdoctoral training from the Chemical Industry Institute of Toxicology (CIIT), Research Triangle Park, North Carolina. He is currently Professor of Occupational and Environmental Health and Director of the Human Toxicology Research Group (TOXHUM) at Université de Montréal. He has been the leader of the risk assessment methodologies theme team of the Canadian Network of Toxicology Centers (1994–2001), and Vice President of the Biological Modeling Specialty Section of the Society of Toxicology (2001–2002). He has also been a member of the U.S. National Academy of Sciences (NAS) Sub-committee on Acute Exposure Guideline Levels (2001–2004), member of the U.S. EPA's Human Studies Review Board (2006–), president of the Risk Assessment Specialty Section of the Society of Toxicology (2005–2006), and a temporary advisor for the World Health Organization for developing a scientific document on the principles for

evaluating health risks in children associated with chemical exposures. His expertise is in the areas of mixture toxicology, health risk assessment methods, and the development of quantitative structure-pharmacokinetic relationships. He has been a peer reviewer of several IRIS updates, risk assessments, mixture risk assessment supplemental guidance, and efforts on interactions for U.S. EPA. He has also been actively involved as a reviewer of ATSDR documents on toxicological profiles and interaction profiles. He has been on the editorial boards of *Toxicological Sciences*, the *International Journal of Toxicology*, the *Journal of Applied Toxicology* and the *Journal of Child Health*. An author of a textbook on environmental pollution, Dr. Krishnan has authored or coauthored over 100 full-length publications and 250 abstracts in the general areas of toxicology, PBPK modeling, QSARs, and risk assessment. His research team received the *Best paper award* (2003) from the Board of Publications of the Society of Toxicology (U.S.A.) for a publication on a novel risk assessment methodology (Haddad S, Béliveau M, Tardif R, and Krishnan K. [2001]) and a PBPK model-based approach for the risk assessment of chemical mixtures (*Toxicological Sciences* 63: 125–135) and more recently received recognition for a publication on QSAR modeling (Béliveau M, Lipscomb J, Tardif R, and Krishnan K [2005]), Quantitative structure-property relationships for interspecies extrapolation of the inhalation pharmacokinetics of organic chemicals (*Chemical Research in Toxicology* 18: 475–485) was part of a "top ten" list of publications advancing the Science of Risk Assessment. Dr. Krishnan was honored with the *Veylian Henderson Award* in 2000 by the Society of Toxicology of Canada for significant contributions to the field of toxicology.

Dr. Andrew Maier currently serves as the Associate Director for the nonprofit organization Toxicology Excellence for Risk Assessment (*TERA*). In his capacity as a toxicologist and risk assessor, he has coauthored technical reports, human health risk assessment documents, and toxicity summaries covering more than 100 individual substances for government and private sponsors. He has led a variety of efforts for developing and applying methods in preventive toxicology and hazard screening that make use of QSAR approaches. Dr. Maier completed his M.S. in industrial health at the University of Michigan and his Ph.D. in toxicology at the University of Cincinnati. He has research interests in the molecular mechanisms of toxicity and has conducted basic research in the areas of metal and polycyclic aromatic hydrocarbon mixtures, environmentally relevant genetic polymorphisms, and risk assessment methods. His recent research efforts have focused on using early biological effect markers and MOA information to reduce uncertainties in chemical risk assessment. Dr. Maier remains active in communicating his findings through participation in professional societies such as the Society of Toxicology. He is a Diplomate of the American Board of Toxicology.

Ms. Chandrika Moudgal is currently PI and technical lead on four projects related to the development of end point-specific QSAR models, PI and technical lead to explore the state of VFAR science and develop a case study using cyanotoxins, and PI and technical lead on a project to develop a Web-based "Data Dictionary" for agents of concern to NHSRC. In addition, she supports TCAD's Provisional Advisory Level (PAL) guidance documents. She also lends support to other NHSRC divisions by reviewing technical documents.

Chandrika earned her M.S. in Toxicology from the University of Cincinnati and her B.S. in Chemistry from the University of Gujarat in Ahmadabad, India. She has also completed course work for an M.S. in Environmental Science at the Ohio State University. Prior to joining NHSRC, Ms. Moudgal served as an environmental health scientist at the National Center for Environmental Assessment (NCEA), ORD, U.S. EPA for approximately seven years. In this position she gained experience and expertise in the development and application of QSARs to fill experimental data gaps and expertise in the application of the Agency's risk assessment methodology. She also served as chemical manager and reviewer of documents for the IRIS database. Additional previous experience includes serving as an Organic Chemistry Section Supervisor with R.D. Zande & Associates in Columbus for seven years; working as a water research analyst for the City of Columbus, drinking water treatment plant for one year; and working as a laboratory scientist for the State of New Hampshire for three years. Chandrika has published several papers related to QSAR research and has presented various papers both nationally and internationally on the topic.

Dr. Joan Rose serves as the Homer Nowlin Chair in Water Research at Michigan State University and is currently Director of the Center for Water Sciences. Dr. Rose received her Ph.D. in Microbiology from the University of Arizona in 1985. She served as a Professor in the College of Marine Science, University of South Florida (USF) from 1998 to 2002.

Dr. Rose's professional experience includes environmental virology, environmental parasitology, drinking water treatment and disinfection, microbial risk assessment, wastewater treatment and reuse, water pollution microbiology, mycology, and food microbiology. Dr. Rose is an international expert in water microbiology, water quality, and public health safety, publishing more than 200 manuscripts. She has been involved in the investigation of numerous waterborne outbreaks worldwide. Her work has examined new molecular methods for waterborne pathogens and zoonotic agents such as *Cryptosporidium* and enteric viruses and source tracking techniques. She has been involved in the study of water supplies, water used for food production, and coastal environments as well as water treatment, wastewater treatment, reclaimed water and water reuse, and quantitative microbial risk assessment. She is specifically interested in microbial pathogen transport in coastal systems and has studied the impact of wastewater discharges and climate on water quality. She was named as one of the 21 most influential people in water in the 21st Century by *Water Technology Magazine* (2000) and won the Clarke Water Prize (one of five international awards for contributions to water science and technology).

Current service on advisory committees includes Chair of the Drinking Water Committee for the Science Advisory Board for the U.S. Environmental Protection Agency; the Science Advisory Board of the International Commission of the Great Lakes, 2003–08; Vice-Chair of USA National Committee for the International Water Association (IWA), 2002–06, Member of the Strategic Council for IWA 2005–08, Chair of the Specialist Group Health-Related Water Microbiology (IWA) 2004–07; Research Advisory Board, National Water Research Institute, 2002–06, and Council Policy Committee for the American Society of Microbiology, 2001–06.

Sources of recent grant and/or contract support include NOAA, U.S. EPA, Water Environmental Research Foundation, NSF, and AWWARF. She was recently awarded as PI as $10 million grant for directing the *Center for Advancing Microbial Risk Assessment* funded by EPA and the U.S. Department of Homeland Security.

Dr. R. Paul Schaudies, Assistant VP at Science Applications International corporation (SAIC), heads a diverse team of technologists who conduct contract biomedical research, scientific analyses, and technical support. Dr. Schaudies is an internationally recognized expert in the fields of biological and chemical warfare defense. He served as a primary Science and Technology Consultant to the Incident Commander, Sergeants-at-Arms for the House and Senate, and U.S. EPA On-Scene Coordinator in response to the October 2001 anthrax incident in Washington D.C. He has served on five National Academy committees in the areas of biological defense and nanotechnology. He has served on numerous national level advisory panels for the Defense Intelligence Agency, the Defense Advanced Research Projects Agency, and the Department of Energy. Dr. Schaudies served 12 years as a U.S. Army officer. While on active duty, Dr. Schaudies served as Chief of the General Support Laboratory in the Department of Clinical Investigation at Walter Reed Army Medical Center, a Senior Researcher at the Walter Reed Army Institute for Research, and a Program Manager for Biological and Chemical Defense Research at the Central Measurement and Signature Intelligence Office at the Defense Intelligence Agency. Dr. Schaudies received his Bachelor's degree in Chemistry from Wake Forest University and his doctoral degree in Biochemistry from Temple University School of Medicine.

Ms. Cynthia Sonich-Mullin is the Director of the Threat and Consequence Assessment Division (TCAD) at the National Homeland Security Research Center. She has provided program leadership, focusing on rapid risk assessment and support to the entire NHSRC team and ORD, since March 2003.

Prior to this assignment, Ms. Sonich-Mullin worked in the National Center for Environmental Assessment in a number of capacities. Most recently, she served concurrent details as the Acting Deputy Director, Cincinnati Division, and the Acting Center Director for Human Health Research.

Since 1993, Ms. Sonich-Mullin has worked on behalf of the International Programme on Chemical Safety (IPCS), a joint program of the World Health Organization (WHO), United Nations Environment Programme, and the International Labour Organization. In October 1993, Ms. Sonich-Mullin worked with IPCS to initiate the IPCS Project: Harmonization of Approaches to the Assessment of Risk from Exposure to Chemicals, on behalf of the WHO. In this capacity, she worked as an IPCS/WHO staff member in Geneva, Switzerland for three years. Upon returning to the U.S., Ms. Sonich-Mullin (as part of U.S. EPA's contribution to the WHO) has continued to work on various aspects of the Harmonization Project.

Prior to the detail, Ms. Sonich-Mullin was a scientist with U.S. EPA's Environmental Criteria and Assessment Office (now the National Center for Environmental Assessment), serving a number of roles including:

- Acting Deputy Director, Cincinnati Division
- Chief, Chemical Mixtures Assessment Branch
- Chief, Systemic Toxicants Assessment Branch

In these capacities, she led and participated in projects related to the assessment of chemicals in air, drinking water and ambient water, municipal solid waste disposal options, and on issues related to Superfund sites. She worked on the development of Agency risk assessment guidelines and served on numerous task groups and research committees including the Agency's Water Research Committee, Air Risk Information Support Center, and as Director of the Superfund Technical Support Center, a center designed to provide risk assessment support, guidance, and advice on issues specifically pertaining to Superfund sites. In a concurrent assignment, she was selected to serve on Vice President Al Gore's Commission to Reinvent Government.

Ms. Sonich-Mullin began her career at EPA working as an environmental health scientist with the Health Effects Research Laboratory. In this capacity, she designed and conducted epidemiological studies related to water contamination and has published in this area. Some specific issues studied included the health effects associated with drinking water chlorination, health effects of sodium in drinking water, and the health effects of the carbon tetrachloride spill into the Ohio River in the late 1970s.

Ms. Sonich-Mullin holds a Master of Environmental Sciences degree, specializing in Applied Biology/Zoology from the Institute of Environmental Sciences, Miami University, Oxford, Ohio. She has also completed doctoral course work in Epidemiology and Biostatistics at the University of Cincinnati, College of Medicine, Cincinnati, Ohio.

Dr. Gerard N. Stelma Jr. is a Senior Science Advisor for the Microbiological and Chemical Exposure Assessment Research Division (MCEARD) of the National Exposure Research Laboratory (NERL), which is part of U.S. EPA's Office of Research and Development. In this role, he provides expert advice regarding microbiological issues, principally

those pertaining to bacterial pathogens, to the division's microbiologists and to various EPA program offices. Dr. Stelma served as the Chief of MCEARD's Microbial Exposure Research Branch for 13 years and as Acting Director of MCEARD for nearly 3 years. Prior to his arrival at EPA, Dr. Stelma was a research microbiologist for the FDA. He holds a B.S. in Biology from the University of Michigan and a Ph.D. in Microbiology from Michigan State University.

Dr. Subhas K. Sikdar is the Acting Associate Director for Health for NRMRL. As the Director of the Sustainable Technology Division until Jan 9, 2004, he was the primary spokesman for U.S. EPA's R&D on clean technologies and pollution prevention. He directed research, both intramural and extramural, on tools and methods for pollution prevention, cleaner process technologies, and demonstration and verification of cleaner technologies. Before joining EPA in 1990, Dr. Sikdar held managerial positions at the National Institute of Standards and Technology in Boulder, Colorado, and General Electric Corporate Research & Development Center in Schenectady, New York. He began his professional career as a Senior Research Engineer with Occidental Research Corporation in Irvine, California, in 1975. Dr. Sikdar earned his B.S. in Chemistry, a B.Tech in Chemical Engineering, and an M.Tech in Polymer Science from Calcutta University in India. He received his M.S. and Ph.D. in Chemical Engineering from the University of Arizona. Dr. Sikdar is a Fellow of the American Association for the Advancement of Science (AAAS), Fellow of the American Institute of Chemical Engineers, Honorary Fellow of the Indian Institute of Chemical Engineers, winner of three EPA bronze medals, an R&D 100 award (1990), AIChE's Larry Cecil Award for Environmental Chemical Engineering (2002), and University of Arizona's Distinguished Engineering Alumnus Award (2003). In the past he was a member of the Vision 2020 Steering Committee for the chemical industry, an action network leader of the Council for Chemical Research. He is a member of the Board of Governors of the Council for Chemical Research (CCR) and of the Green Chemistry Institute, a member of AIChE's Research and New Technology Committee, and the Chair of the Sustainable Engineering Forum. For some years he has been championing the concepts and methods for clean products and processes through a NATO pilot project, two NATO workshops, and an Engineering Foundation conference. He is a current member of the Industrial Advisory Board of the University of Arizona's College of Engineering and of the Department of Chemical and Environmental Engineering, and of the Department of Chemical and Environmental Engineering of the Illinois Institute of Technology. Dr. Sikdar is the leader of the technical expert group for the Center of Excellence on Environmental Engineering and Hazardous Wastes, which is composed of several universities in Thailand. He is the founder and co-Editor-in-Chief of the international journal, *Clean Technologies and Environmental Policy*, published quarterly by Springer Verlag of Germany. Dr. Sikdar has published more than 60 technical papers in reputed journals, holds 22 U.S. patents, and has edited 13 books.

Dr. Sikdar was instrumental in developing the highly successful Occidental Hemihydrate process for phosphoric acid manufacture. His other technical achievements include developing several membrane processes for pervaporative separation of VOCs from aqueous effluents and for highly selective sorption of heavy metals, masterminding the development of a waste reduction algorithm for process design (the WAR algorithm), a solvent design algorithm (PARIS II), and a data portal for life cycle assessment (LCAccess).

Dr. William J. (Bill) Welsh holds the Norman H. Edelman Professorship in Bioinformatics and Computer-Aided Molecular Design in the Department of Pharmacology at the Robert Wood Johnson Medical School (RWJMS) in Piscataway NJ, University of Medicine and Dentistry of New Jersey (UMDNJ). Concurrently, he serves as Director of the *UMDNJ Informatics Institute* (http://informatics.umdnj.edu) that coordinates university-wide initiatives in bioinformatics, clinical informatics, and computer-aided molecular design. He is also PI and Director of the EPA-supported *New Jersey Research Center for Environmental Bioinformatics and Computational Predictive Toxicology*, the first of its kind in the nation. He is a member of various centers and institutes of excellence at UMDNJ and Rutgers University, including the Cancer Institute of New Jersey, the New Jersey Center for Biomaterials, Rutgers University School of Pharmacy, and the Environmental & Occupational Health Sciences Institute (EOHSI).

Dr. Welsh earned a B.S. degree (*magna cum laude*) in Chemistry from St. Joseph's University (Philadelphia, PA) and a Ph.D. degree in Theoretical Physical Chemistry from the University of Pennsylvania (Philadelphia, PA). He conducted postdoctoral research in the laboratory of Dr. James E. Mark, Distinguished Professor of Polymer Science at the University of Cincinnati (Cincinnati, OH). In 1985, Dr. Welsh joined the University of Missouri (St. Louis) as an Associate Professor of Chemistry and rose through the ranks to Distinguished Professor in 1998. During this period he was appointed Director, Laboratory for Computer-Aided Molecular Design, at the University of Missouri. In 2001, Dr. Welsh joined UMDNJ-Robert Wood Johnson Medical School to assume his present role.

Dr. Welsh's laboratory specializes in the development and application of computational tools for drug discovery. Promising candidates emanating from these rational design approaches are synthesized and tested as potential therapeutic or diagnostic agents. His laboratory is widely reputed for its innovation, such as the development of the *Shape Signatures* tool and the discovery of potential drug candidates for the treatment of cancer, severe and chronic pain, neurodegenerative diseases, and heart conditions. Dr. Welsh's publication record includes over 350 articles in peer-reviewed books and journals, 600 abstracts from presentations at professional scientific meetings, and several patents and patent applications. He is the recipient of numerous awards and honors, including the *Teacher of the Year Award* (1983 and 1985), the *St. Louis Research Award*

(1998), the *University of Missouri-St. Louis Chancellor's Research and Creativity Award* (2001), the *University of Missouri Entrepreneur of the Year Award* (2001), the *Norman H. Edelman Endowed Professorship in Bioinformatics* at UMDNJ-RWJMS (2003), and most recently the *John C. Krantz, Jr. Award* (2004). He serves on the advisory boards of several scientific journals. Spanning the last twenty years, over 125 students (postgraduate and graduate students, undergraduates, and research associates) have trained in his laboratory.

Dr. Andrew Worth works at the European Chemicals Bureau (ECB) within the European Commission's Joint Research Centre (JRC) in Italy. He joined the JRC with degrees in Physiological Sciences and Linguistics from the University of Oxford (UK) and with post-graduate experience in the fields of biochemistry and toxicology. He subsequently gained a Ph.D in Computational Toxicology from Liverpool John Moores University (UK). His research interests have focused on the development of QSAR models and methods and on the development of Integrated Testing Strategies for chemical toxicity based on the use of physicochemical and *in vitro* data. Since 2003 he has been leading the JRC Project on Computational Toxicology. In addition to coordinating QSAR-related work within the JRC, Dr. Worth also chairs the EU Working Group on QSARs and represents the European Commission in several OECD working groups.

Dr. Douglas Young leads the Clean Processes Branch (CPB) that resides in the Sustainable Technology Division (STD) in the Office of Research and Development within EPA. STD is home to EPA's in-house research in the areas of Green Chemistry and Sustainability. Dr. Young's research is in the areas of environmental impact assessment as it pertains to the chemical processing industry and the estimation of acute toxicity measurements. He has been intimately involved in the creation of the Computational Toxicology Research Program and the National Center for Computational Toxicology within the EPA. He was instrumental in the development and commercialization of the generalized Waste Reduction (WAR) algorithm. Dr. Young received his Ph.D. from the University of Arizona where his dissertation focused on the bioremediation of high-energy explosive waste generated at the Los Alamos National Laboratory. He received his M.S. from the University of Notre Dame and his B.S. from the University of Michigan. All three of Dr. Young's degrees are in Chemical Engineering.

Appendix C
Workshop Agenda

Tuesday, June 20, 2006

8:00 am Welcome (Chandrika Moudgal, NHSRC)

8:10 am Opening remarks (Andy Avel, NHSRC; Jonathan Herrmann, NHSRC; Subhas Sikdar, NRMRL)

8:30 am Background on NHSRC and NRMRL (Cindy Sonich-Mullin, NHSRC; Douglas Young, NRMRL) and introduction of expert panel members

8:50 am QSAR/VFAR program and charge to expert panel members (Chandrika Moudgal, NHSRC)

9:00 am Introduction to the VFAR concept (Dr. Gerald Stelma, NERL)

9:20 am Using VFAR in the Risk Assessment Framework (Dr. Joan Rose, MSU)

9:50 am VFAR: factors related to genomic variabilities (Dr. Syed Hashsham, MSU)

10:20 am Break

10:50 am A bioinformatic approach to VFAR analysis and characterization (Dr. Paul Schaudies, SAIC)

11:20 am VFAR charge questions 1 and 2 (Discussion)

12:00 pm Lunch

1:00 pm VFAR charge questions 3 and 4 (Discussion)

3:00 pm Break

3:30 pm VFAR charge questions 5, 6, and 7 (Discussion)

5:00 pm VFAR closing remarks from panel and EPA

Wednesday, June 21, 2006

8:00 am From reactivity to regulation: integrating alternative techniques to predict toxicity (Dr. Mark Cronin, TOXHUM)

8:20 am Integrated QSAR-PBPK modeling for risk assessment applications (Dr. Kannan Krishnan, TOXHUM)

8:40 am Integration of MOA and WOE concepts in predictive toxicology (Dr. Andrew Maier, TERA)

9:00 am Activities at the new UMDNJ Computational Toxicology Center: advanced QSAR-based methods of rapid hazard identification, prediction, and characterization (Dr. William Welsh, RWJMS)

9:20 am The role of the European Chemicals Bureau in promoting the regulatory implementation of estimation methods (Dr. Andrew Worth, JRC)

9:40 am Break

10:10 am QSAR charge questions 1, 2, and 3 (Discussion)

12:00 pm Lunch

1:00 pm QSAR charge questions 4 and 5 (Discussion)

3:00 pm Break

3:30 pm QSAR charge questions 6 and 7 (Discussion)

4:30 pm Workshop closing remarks (Panel, NHSRC and NRMRL management, Douglas Young, Chandrika Moudgal)

5:00 pm Adjourn

Appendix D
List of Attendees

Femi Adeshina, Ph.D.
1200 Pennsylvania Avenue, NW (8801R)
Washington, DC 20460
Phone: 202-564-1539
Email: adeshina.femi@epa.gov

Caroline Baier-Anderson, Ph.D.
EnDyna, Inc.
7925 Jones Branch Drive, Suite 5300
McLean, VA 22102
Phone: 410-610-1737
FAX: 703-873-4372
Email: canderson@endyna.com

Irv Baumel, Ph.D.
USEPA/NHSRC/TCAD
Phone: 202-564-2338
Email: baumel.irwin@epa.gov

Dominic L. Boccelli, Ph.D.
Environmental Engineer
USEPA/ORD/NHSRC/WIPD
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7654
FAX: 513-487-2555
Email: boccelli.dominic@epa.gov

Kathyrn Boyle
Chemist
USEPA/OPP
1200 Pennsylvania Avenue, NW (7506P)
Washington, DC 20460
Phone: 703-305-6304
Email: boyle.kathryn@epa.gov

Nichole Brinkman
Biologist
USEPA/NERL
26 West Martin Luther King Drive (MS 320)
Cincinnati, OH 45268
Phone: 513-569-7315
FAX: 513-569-7117
Email: brinkman.nichole@epa.gov

Karen Burgan
Senior Policy Advisory
USEPA/OSWER/OEM/NPPD
1200 Pennsylvania Avenue, NW (5104A)
Washington, DC 20460
Phone: 202-564-1978
FAX: 202-564-2620
Email: burgan.karen@epa.gov

Dan Chappie
Battelle
10300 Alliance Road, Suite 155
Cincinnati, OH 45242
Phone: 513-362-2600
FAX: 513-362-2610

Kathy Clayton
USEPA/ORD/NHSRC
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7046
Email: clayton.kathy-ci@epa.gov

Maura J. Donohue, Ph.D.
Chemist
USEPA/ORD/NERL/MCEARD/CERB
26 West Martin Luther King Drive
Cincinnati, OH 45268
Phone: 513-569-7634
FAX: 513-569-7757
Email: donohue.maura@epa.gov

Anthony Fristachi, M.S.
Exposure Analyst
USEPA/ORD/NCEA
26 West Martin Luther King Drive (MS A110)
Cincinnati, OH 45268
Phone: 513-569-7144
FAX: 513-487-2539
Email: fristachi.anthony@epa.gov

Bernard Gadagbui, Ph.D.
Toxicology Excellence for Risk Assessment
2300 Montana Avenue, Suite 409
Cincinnati, OH 45211
Phone: 513-542-7475 ext. 27
FAX: 513-542-7487
Email: bgadagbui@tera.org

Robert Goble, Ph.D.
Research Professor and Director
George Perkins Marsh Institute, Clark University
950 Main Street
Worcester, MA 01610
Phone: 508-751-4612
FAX: 508-751-4600
Email: rgoble@clarku.edu

Paul Harten, Ph.D.
Physical Scientist
USEPA/ORD/NRMRL
26 West Martin Luther King Drive
Cincinnati, OH 45268
Phone: 513-569-7045
Email: harten.paul@epa.gov

Stephanie Hines
OSU Extension - Clermont County
1000 Locust Street, P.O. Box 670
Owensville, OH 45160
Phone: 513-732-7070
Email: hines.180@osu.edu

Sheldon Jobe
EnDyna, Inc.
7925 Jones Branch Drive
Suite 5300
McLean, VA 22102
Phone: 703-873-4367
FAX: 703-873-4372
Email: sjobe@endyna.com

Barbara Klieforth
Biologist
USEPA/ORD/OSA
1300 Pennsylvania Avenue, NW RM B26J
Washington, DC 20004
Phone: 202-564-6787
FAX: 202-565-2431
Email: klieforth.barbara@epa.gov

Steven S. Kueberuwa, Ph.D.
Toxicologist
USEPA/OW/OST/HECD
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Phone: 202-566-0233
FAX: 202-566-1139
Email: kueberuwa.steven@epa.gov

Jason C. Lambert, Ph.D.
ORISE Fellow
USEPA/ORD/NCEA
26 West Martin Luther King Drive (MS A110)
Cincinnati, OH 45268-1320
Phone: 513-569-7078
Email: lambert.jason@epa.gov

Josh Larson
Biosecurity Analyst
Sandia National Laboratories
P.O. Box 5800 MS 1371
Albuquerque, NM 87185
Phone: 505-844-0357
FAX: 505-284-8870
Email: jjlarso@sandia.gov

Todd Martin, Ph.D.
Research Chemical Engineer
USEPA/NRMRL/CPB
26 West Martin Luther King Drive (MS 443)
Cincinnati, OH 45268
Phone: 513-569-7682
Email: martin.todd@epa.gov

Deborah McKean, Ph.D.
Toxicologist
USEPA/OSWER/OEM/NDT
26 West Martin Luther King Drive
Cincinnati, OH 45268
Phone: 513-487-2435
FAX: 513-487-2537
Email: mckean.deborah@epa.gov

Leroy Michelsen
Engineer
OSWER/OEM/NDT
26 West Martin Luther King Drive
Cincinnati, OH 45268-1320
Phone: 513-487-2431
FAX: 513-487-2537
Email: mickelsen.leroy@epa.gov

Matthew D. Miller, Ph.D.
Post-doctoral research associate
University of Missouri – Kansas City
7543 Terrace Street
Kansas City, MO 64114-1637
Phone: 816-277-8264
FAX: 816-235-6543
Email: mdma95@umkc.edu

H.A. Minnigh
RCAP Solutions, Inc./CECIA, UIPR
P.O. Box 48
Lajas, PR 00667
Phone: 787-392-7186
FAX: 787-892-2089
Email: hminnigh@compuserve.com

Vlasta Molak, Ph.D.
President and CEO
GAIA Foundation, Inc.
8987 Cotillion Drive
Cincinnati, OH 45231
Phone: 513-521-9321
Email: drmolak@gmail.com

Tonya Nichols, Ph.D.
USEPA
26 West Martin Luther King Drive (MS 163)
Cincinnati, OH 45268-1320
Phone: 513-569-7805
Email: nichols.tonya@epa.gov

Appendix E
Workshop Presentation Materials

Introduction to The VFARs Concept

Jerry Stelma
June 20, 2006

Although this work was reviewed by EPA and approved for presentation, it may not necessarily reflect official Agency policy

Research and Development at EPA

- 1,950 employees
- $700 million budget
- $100 million extramural research grant program
- 13 lab or research facilities across the U.S.
- Credible, relevant and timely research results and technical support that inform EPA policy decisions

Making decisions with sound science requires..

- Relevant, high quality, cutting-edge research in human health, ecology, pollution control and prevention, economics and decision sciences
- Proper characterization of scientific findings
- Appropriate use of science in the decision process

Research and development contribute uniquely to..

- Health and ecological research, as well as research in pollution prevention and new technology
- In-house research and an external grants program
- Problem-driven and core research

High Priority Research Areas

- Human Health
- Particulate Matter
- Drinking Water
- Clean Water
- Global Change
- Endocrine Disruptors
- Ecological Risk
- Pollution Prevention
- Homeland Security

The Contaminant Candidate List(CCL)

- Developed as a result of the 1996 amendments to SDWA
  - EPA must periodically develop a list of currently unregulated contaminants
  - EPA must select 5 contaminants for regulatory decisions per 5 years

The Contaminant Candidate List(CCL)

- Method for developing the lists not specified by SDWA
- Methods for selecting the five or more contaminants not specified by SDWA

NRC Workshop Results

- "*Identifying future Drinking Water Contaminants*" Recommendations
  - A process was developed to narrow, focus and prioritize contaminants.
  - Explore the feasibility of using virulence factor activity relationships (VFARs) for microbial contaminants

Origin of the concept

- Structural activity relationships (SARs) found in chemicals
  - Compares newly identified chemical structures to known chemical structures
  - Toxicity is predicted by the comparisons

Premise

- Architectural and biochemical components of microorganisms that cause disease are also structurally related

Central Concept

- Ability to predict virulence by microbial characteristics
  - Microbial VFARs should function much the same as QSARs do in chemistry
  - Research has shown certain common characteristics among pathogens
  - "Descriptors" have been tied to specific genes

Why would we expect structural relationships among genes?

- Parallel evolution
- Horizontal gene exchange
  - Common occurrence within a genus
  - Has been observed beyond genus boundaries
- Genetic engineering

Examples of Descriptors

- Genetic elements
- Surface proteins
- Toxins
- Attachment Factors
- Metabolic pathways
- Invasion factors

Current Challenges to use of VFARs/Microarrays

- QSARs vs VFARs: Does the biological universe parallel the chemical universe?
  - Chemicals are static
  - Microbes are dynamic
  - Examples of parallel VFs
    - Cholera toxin and *E. coli* LT
    - Pyrogenic toxins of *Strep*. and *Staph*.

Current Challenges to use of VFARs/Microarrays

- Examples of parallel VFs
  - Cholera toxin and E. coli LT
  - Pyrogenic toxins of Strep. and Staph.
- Examples of unique VFs
  - Salmonella *inv*A gene
  - Legionella *mip* gene

Current Challenges to use of VFARs/Microarrays

- Too many unknown virulence genes
  - Individual virulence genes are necessary for virulence
  - Individual virulence genes are not sufficient for virulence
  - Entire arrays of virulence genes are needed

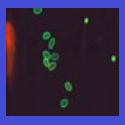
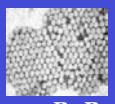
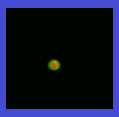
Current Challenges to use of VFARs/Microarrays

- Host susceptibility factors and dosages
- DNA Variability among structural genes
- Effect of unexpressed virulence genes?
  - Genes can be present but not expressed
- Are VFARs valid for viruses and protozoa?
  - All are obligate parasites
  - Factors leading to species specificity?
- Effect of DNA from dead cells

Current Challenges to use of VFARs/Microarrays

- "The message is that there are known knowns - there are things that we know that we know. There are known unknowns - that is to say, there are things that we now know we don't know. But there are also unknown unknowns - there are things we do not know we don't know. And each day we discover a few more of those unknown unknowns".
- Rumsfeld 2003

Using VFAR in a Risk Assessment Framework

Joan B. Rose
rosejo@msu.edu
Homer Nowlin Endowed Chair for Water Research
MICHIGAN STATE UNIVERSITY

Definitions used in risk analysis

Risk assessment	The qualitative or quantitative characterization and estimation of potential adverse health effects associated with exposure of individuals or populations to hazards (materials or situations, physical, chemical and or microbial agents.)
Risk management	The process for controlling risks, weighing alternatives, selecting appropriate action, taking into account risk assessment, values, engineering, economics, legal and political issues.
Risk communication	The communication of risks to managers, stakeholders, public officials, and the public, includes public perception and ability to exchange scientific information.

Risk assessment is a method to examine qualitatively or quantitatively the potential for harm from exposure to contaminants or specific hazards.

- Monitoring and data are some of the keys to establishing risks and therefore safety goals.

Quantitative Risk Assessment QRA

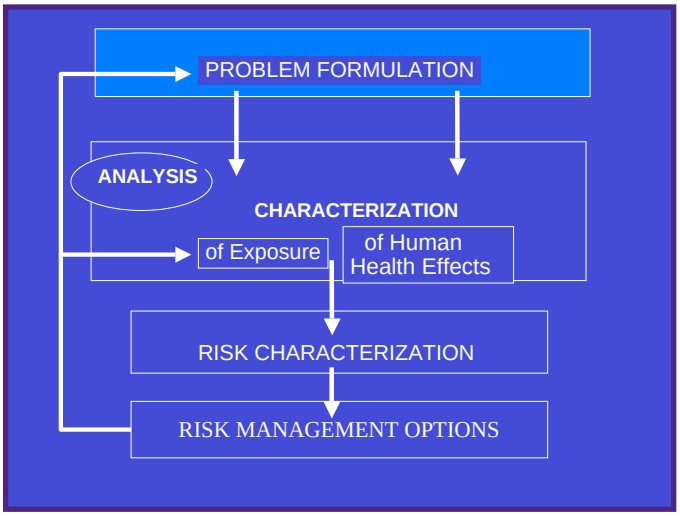
- Tool used to estimate adverse health effects associated with specific hazards.
- Elicits a statistical estimate or probability of harm.
- Used for risk management decisions.

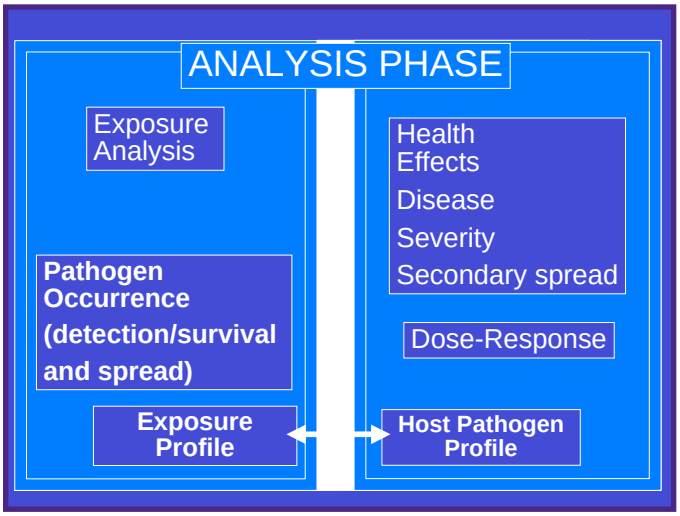
NATIONAL ACADEMY OF SCIENCES RISK ASSESSMENT PARADIGM

- **HAZARD IDENTIFICATION**
  Types of microorganisms and disease end-points
- **DOSE-RESPONSE**
  Human feeding studies, clinical studies, less virulent microbes and health adults
- **EXPOSURE**
  Monitoring data, indicators and modeling used to address exposure
- **RISK CHARACTERIZATION**
  Magnitude of the risk, uncertainty and variability

Evolution of QMRA

- **<1980** — Indicator approaches used
  Suggesting that some level of contamination below which one is safe
- **1980's** — Initial Dose Response Concepts
  Application in development of EPA Rules
- **1988** — Dose-response for Giardia, viruses in Water.
- **1990** — Adoption for food safety
  WHO food and water consultations
  Dynamic model applications
  ILSI framework documents
- **2000's** — Air and Home Land Security applications
  Reg framework development
  Population sensitivities



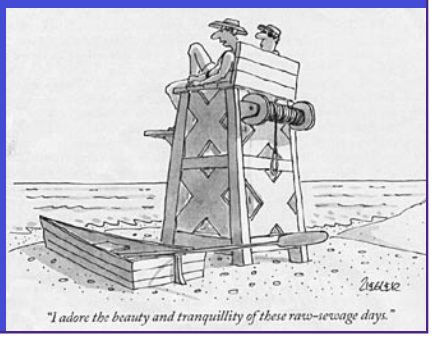


Relationship of MRA to VFAR

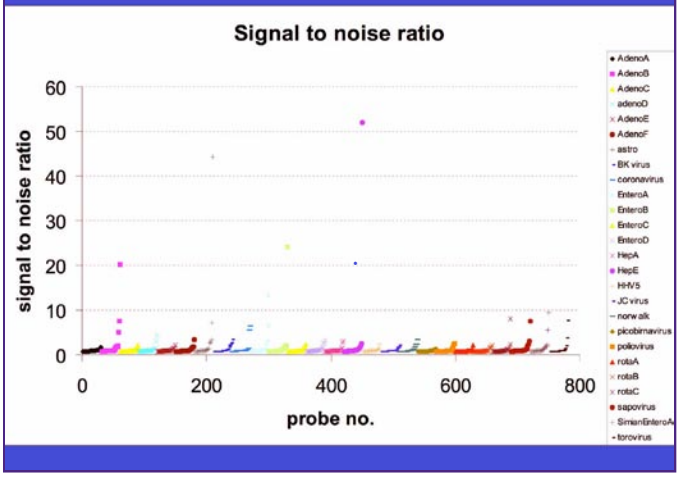
Hazard ID	Source; Identification; virulence; potential for severe outcomes
Dose-response	Potency
Exposure	Source; persistence (in nature, during disinfection)
Characterization	Sensitive populations (receptors); Evolution of Pathogens

The HAZARDS

Pathogen Discovery in Intestinal Tract and in Sewage Through Genomics



Microarray for Virus Discovery from Sewage (Mark Wong, MSU)



Biological Hazards

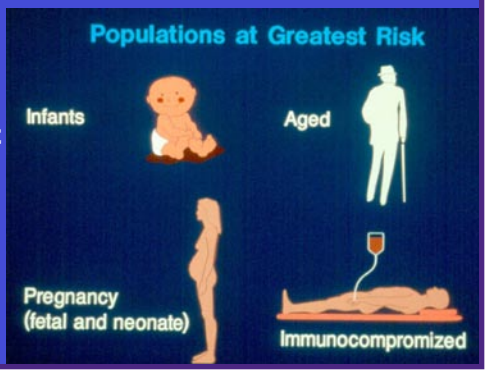
- Viruses, prions, bacteria, and protozoa are more likely than fungi or helminths to be associated with emerging infections.
- Zoonotic pathogens comprise 75% of emerging infectious diseases.
- Pathogens which are subject to relatively frequent mutation or genomic reassortment events (e.g. RNA viruses and viruses with segmented genomes) are more likely to emerge.
- Pathogens which infect multiple hosts or pathogens that infect species that can harbour multiply closely related agents providing an opportunity for reassortment or recombination (e.g. SARS in cats) are likely to emerge.
- Agents transmissible by more than one route or by indirect contact, e.g. water, food, environmental contamination, vectors, etc, are likely to emerge.

Acute and Chronic Outcome Associated with Microbial infections

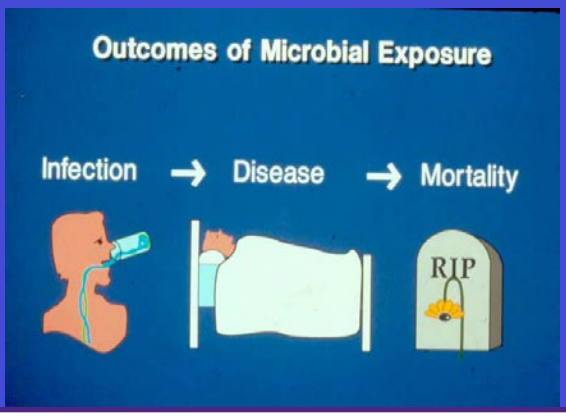
	Acute disease	**Chronic disease**
Microorganism	**Outcomes**	**Outcomes**
Campylobacter	Diarrhea	Gullain-Barre' syndrome
E. Coli 015H7	Diarrhea	Hemolytic uremic syndrome
Helicobacter	Gastritis	Ulcers and stomach cancer
Salmonella, Shigella, & Yersinia	Diarrhea	Reactive arthritis
Coxsackievirus B *Adenoviruses*	Encephalitis, aseptic Meningitis, diarrhea, respiratory disease	Diabetes Myocarditis Obesity
Giardia	Diarrhea	Failure to thrive, lactose intolerance, chronic joint pain
Toxoplama	Newborn syndrome, hearing and visual loss	Mental retardation, dementia, seizures

Morbidity and Mortality greater in the Sensitive Populations
30% of our populations Fall into one of the Sensitive Populations at any one time.

ZOONOTIC
AGENTS
OPPORTUNISTIC
AGENTS
EFFECT
THIS
GROUP



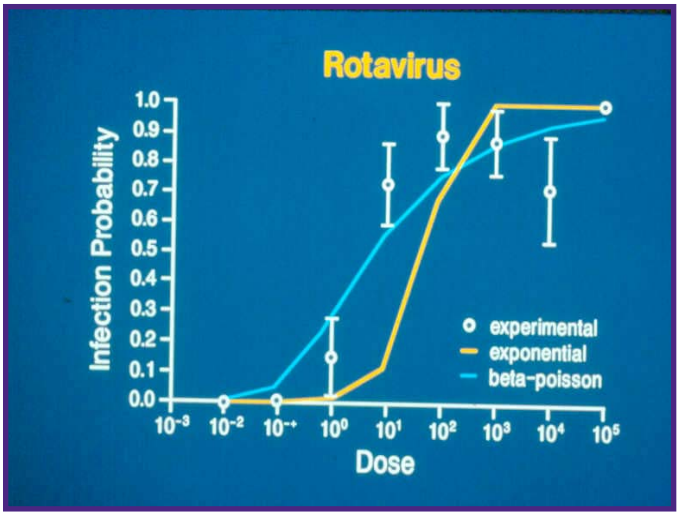
DOSE-RESPONSE

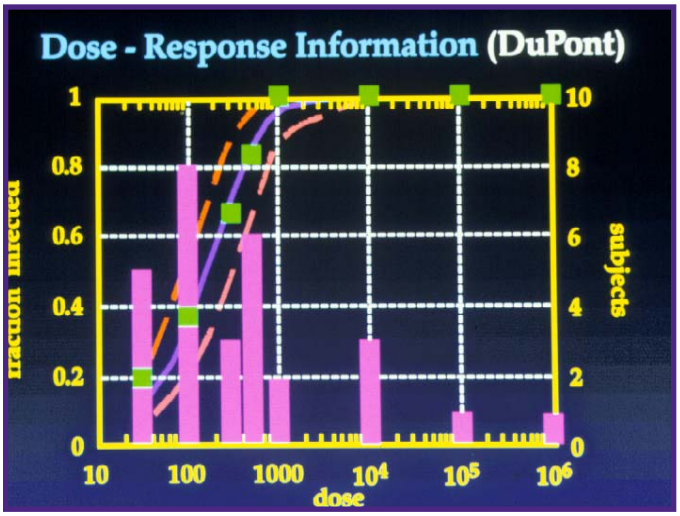


Dose-Response

Dose-response data sets have been developed in human feeding studies for

- Dose measurements were by PFU/or by infectious titer, CFU or cysts or oocysts.
- End points of measurements were excretion of the pathogen and/or antibody response, rarely disease.
- Mathematically address the shape of the ratio of those affected/exposed.
- Need minimum of three doses. Must have doses which elicit effects different from 0% and 100%





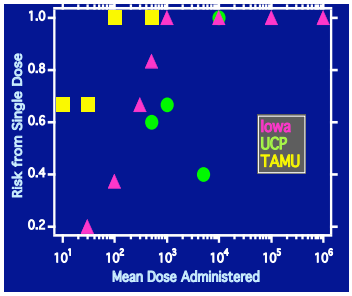
Strain Differences

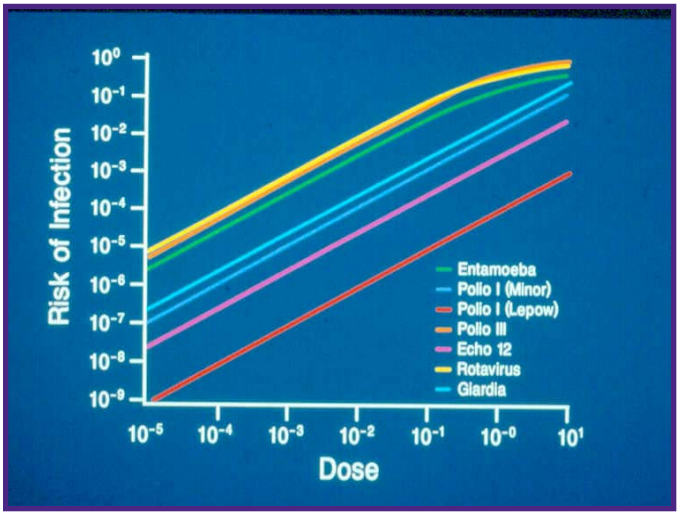
Human volunteers, *C. parvum*,

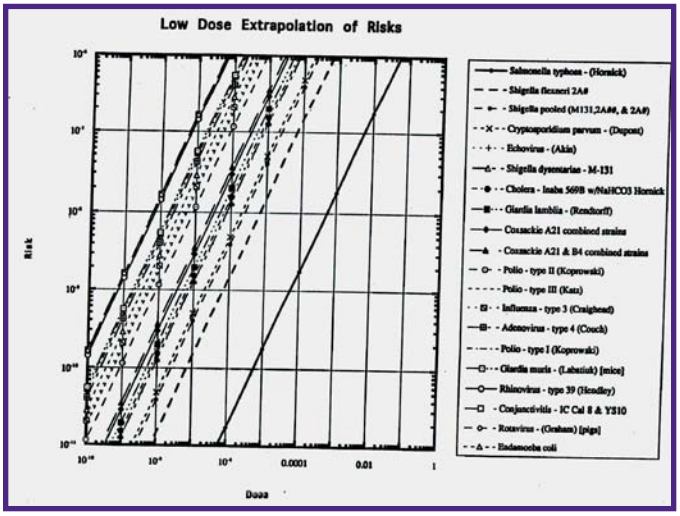
DuPont et al. (1995)

Okhuysen et al. (1999)

Potential for probabilistic modeling of inter-strain variability (Teunis and colleagues)





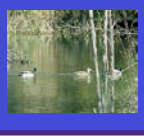


EXPOSURE ASSESSMENT

- Route of Exposure
- Duration of exposure
  - Seconds, hours, minutes
- Number of exposures
  - How many times in a day, month, year
- Degree of exposure
  - Liters of water ingested
  - Liters of air inhaled
  - Grams of food ingested

Microbial Source Tracking

•Tools are now available to determine the molecular fingerprint of the fecal pollution.
•Health risks
•Remediation
•Prioritization
•Responsibility

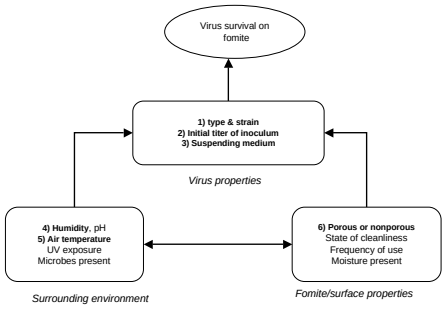


Host Specific Markers are Key to Source Tracking Future

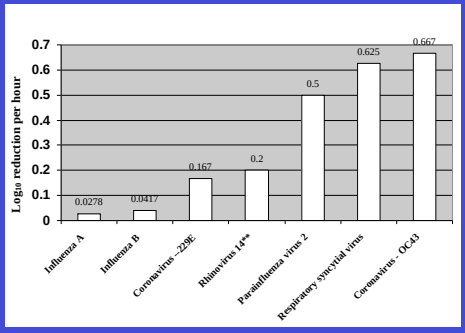
- Bacteroides (genetic approaches PCR)
- 4/4 sewage; 4/4 human; 4/5 cow (lowest concentration missed) 4/4 dogs however no marker for Birds: Missed 2 samples with dog and 2 with cow that were mixed.
- E.coli Toxin genes able to detect sewage (4/4).
- Enteroviruses and Adenoviruses found in 3 of 4 sewage samples.
- Enterococci ESP marker found in 109 human sewage water samples and zero of 80 animal samples.

EXPOSURE ASSESSMENT

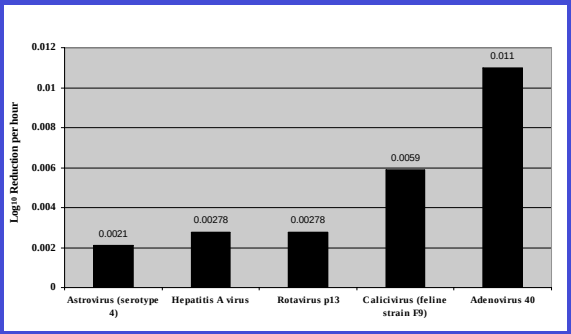
- Occurrence
- Survival
- Regrowth
- Accumulation
- Transport

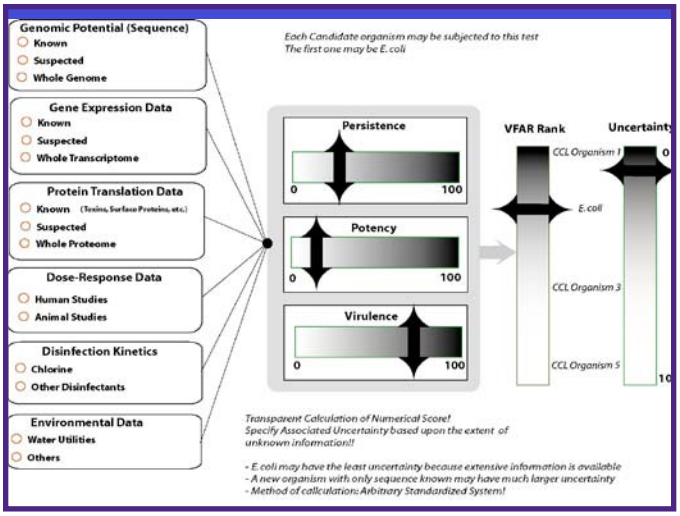


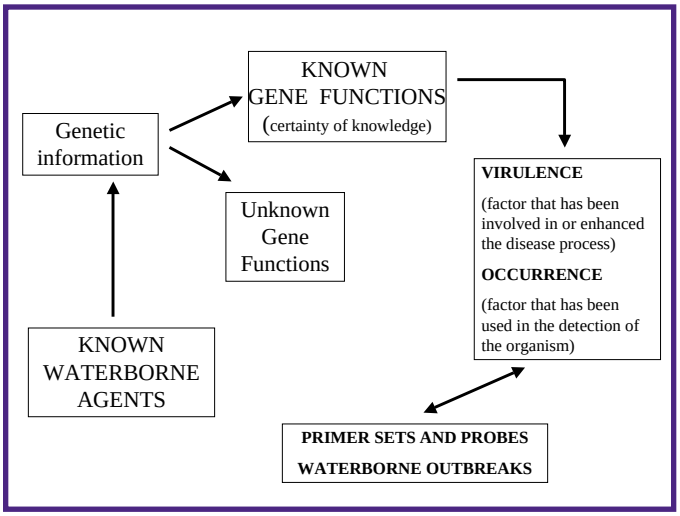
Inactivation Rates on Fomites (Gerba and Boone, Univ Arizona)



Inactivation on Fomites (Gerba and Boone, Univ. of Arizona





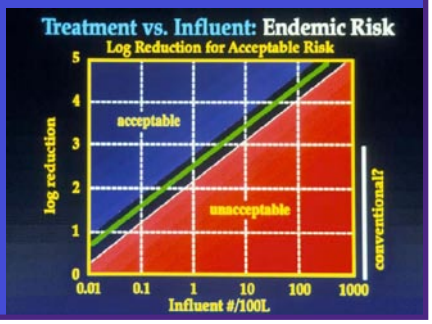


New Tools and Data bases for Assessing Occurrence and Safety

Understanding Genetic Detection in Water

Which sequence For disinfection? Removal capabilities

QPCR



QSARs

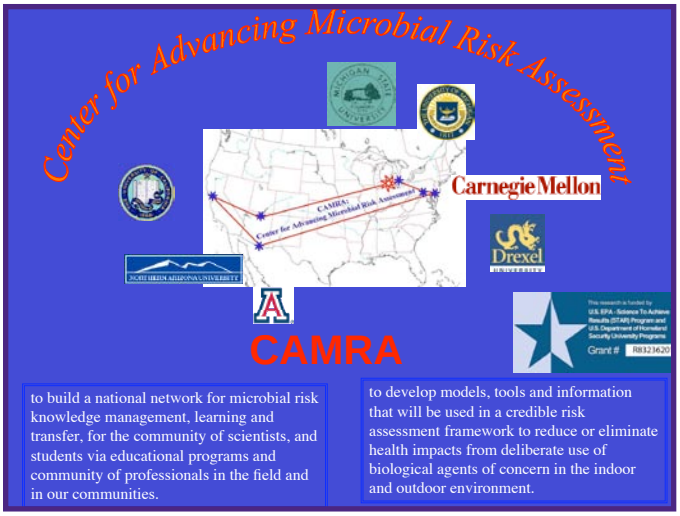
- Quantitative Structure-Activity Relationship used by EPA for over 13 years for hazard risk evaluation of chemicals part of the new chemicals program.
- First explored in 1950s by Hansch to correlate the molecular structure to biological activity.
- 4500 citations
- Software program (PBT profiler) developed just released 10 years in the making. (enter by drawing the structure, entering the identifying # or written chemical linear structure.
- Persistence (1/2 lives predicted ambient conditions.
- Bioaccumulation
- Toxicity (acute and chronic fish toxicity)
- Predictive, some uncertainty, limitations (does not do metals, endocrines).
- Can place them into chemical categories.
- Defines high, medium and low risk.

WATERBORNE DISEASE GENOMICS PROGRAM

A Long-term Commitment to Developing the Data, the Technology, Supporting Analyses, Algorithms and Research projects including a Program in Functional Genomics is necessary.

Recognition that this is a predictive approach to examining risk and uncertainty will be part of the program.

May not work for all classes of Microbes equally.



THANK YOU

Science for

Societal Benefits.

MICHIGAN STATE UNIVERSITY

VFAR: Factors Related to Genomic Variabilities

US EPA QSAR/VFAR Workshop
Cincinnati, OH

June 20, 2006
9:50 AM

Syed A. Hashsham

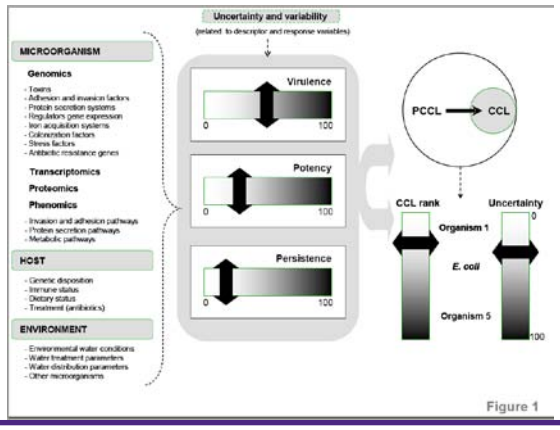
Associate Professor

Department of Civil and Environmental Engineering and
Center for Microbial Ecology

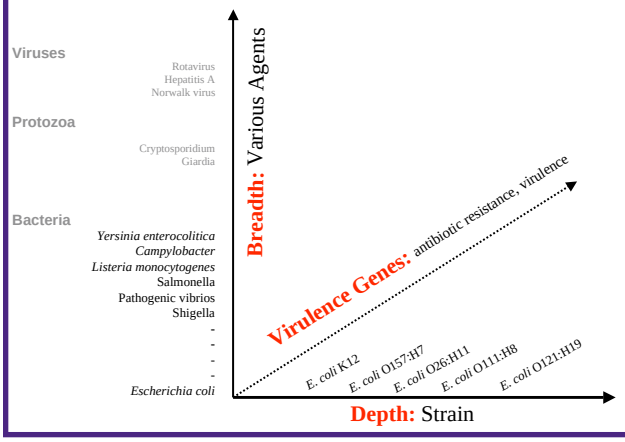
FACTORS RELATED TO DEVELOPMENT

The Overall Concept

Tourlousse et al., Water Environment Research, ***79*** *(2007)*



Assumptions!



Variable Virulence Factors- Variable Effects

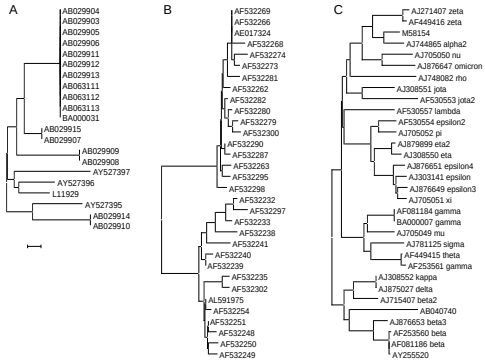
Strain	Pathotype	eae	stx1	stx2	set1A	SenA	ehxA	hylA	espP	sat	iucC	shuS/chu	ipaH
K12	Not pathogenic	-	-	-	-	-	-	-	-	-	-	-	-
W3110	Not pathogenic	-	-	-	-	-	-	-	-	-	-	-	-
EDL 933	EHEC	●	●	●	-	-	●	-	●	-	-	●	-
Sakai	EHEC	●	●	●	-	-	●	-	●	-	-	●	-
CFT073	UPEC	-	-	-	●	-	-	●	-	●	●	●	-
S. flexneri 2a str. 301	EIEC	-	-	-	●	●	-	-	-	-	●	-	●
S. sonnei Ss046	EIEC	-	-	-	-	●	-	-	-	-	●	-	●
S. dysenteriae 1 str. 197	EIEC	-	●	-	-	●	-	-	-	-	-	●	●
S. boydii 4 str. 227	EIEC	-	-	-	-	●	-	-	-	-	●	-	●

EHEC: Enterohemorrhagic *E. coli*; UPEC: Uropathogenic *E. coli*; EIEC: Enteroinvasive *E. coli*

Tourlousse et al., 2006. *Water Environment Research*, Special Issue (*Accepted*)

Specific Examples

toxR of *V. parahaemolyticus* | *iap* of *L. monocytogenes* | *eae* of *E. coli*



Tourlousse et al., 2006. *Water Environment Research*, Special Issue (*Accepted*)

Wide (and Dynamic) Range of Genetic Variability

Pathogen	Gene	Sequences analyzed	Gene length[a]	Average sequence diversity		Maximum sequence variability	
		(No.)	(bp)	(bp)	(%)	(bp)	(%)
E. coli	*eaeA*	76	2822	356.4	12.6	564	20.0
H. pylori	*vacA*	77	3896	306.6	7.9	602	15.5
H. pylori	*cagA*	89	3585	218.9	6.1	388	10.8
E. coli	*stx1A*	41	948	22.7	2.4	77	8.1
H. pylori	*ureA*	36	717	17.0	2.4	37	5.2
L. monocytogenes	*iap*	42	1440	29.5	2.0	67	4.7
L. monocytogenes	*plcB*	116	870	20.2	2.3	37	4.3
V. parahaemolyticus	*tdh*	20	570	10.8	1.9	20	3.5
C. perfringens	*plc*	18	1197	14.9	1.2	28	2.3
V. cholerae	*ctxA*	30	777	1.7	0.2	17	2.2
V. cholerae	*ctxB*	33	375	2.6	0.7	7	1.9
V. parahaemolyticus	*toxR*	20	879	6.1	0.7	14	1.6

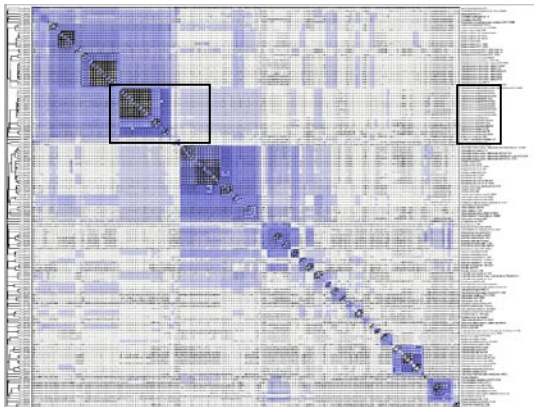
Dynamic: Changes as the database grows!

Tourlousse et al., 2006. *Water Environment Research*, Special Issue (*Accepted*)

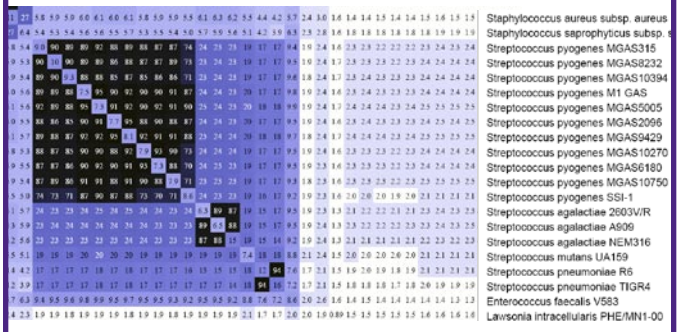
Ranking the Marker Genes Specificity: FunGene Pipeline

Potential marker	Major organism containing the potential marker	Bit score for training sequences[1]	No. of sequences given by FOFE[2]	% for genes associated with target organism[3]	Potential as marker 10= high 1= low
AcdT	*Lactobacillus acidophilus*	139	2	100%	2
ace	*Enterococcus faecalis*	1928	16	100%	8
ace-48	*Enterococcus faecalis*	300	4	100%	4
bft	*Bacteroides fragilis*	1243	6	100%	6
bvt79	*Butyrivibrio fibrisolvens*	143	11	27%	3
celA	*Ruminococcus albus*	1177	4	50%	5
cfiA	*Bacteroides fragilis*	798	15	100%	7
cfxA	*Bacteroides-Prevotella*	1003	19	90%	7
cpb	*Clostridium perfringens*	1067	32	100%	8
cpe	*Clostridium perfringens*	999	5	100%	6
esp	*Enterococcus faecium*	5525	4	100%	8
etx	*Clostridium perfringens*	999	5	100%	6
gelE	*Enterococcus faecalis*	1600	3	100%	6
gtfB	*Streptococcus spp., Leuconostoc spp.*	4692	32	88%	10
mutB	*Streptococcus mutans*	2971	3	100%	6
nisB	*Lactococcus lactis*	3108	5	100%	7
ply	*Streptococcus spp., Clostridium spp.*	1496	24	79%	8
rumA	*Ruminococcus gnavus, Butyrivibrio, etc.*	142	10	20%	2
sda	*Streptococcus pyogenes*	1184	4	100%	6
stxA2	*Escherichia coli, Shigella*	1005	102	90%	9
uidA	*E. coli, Shigella*	1825	66	79%	9

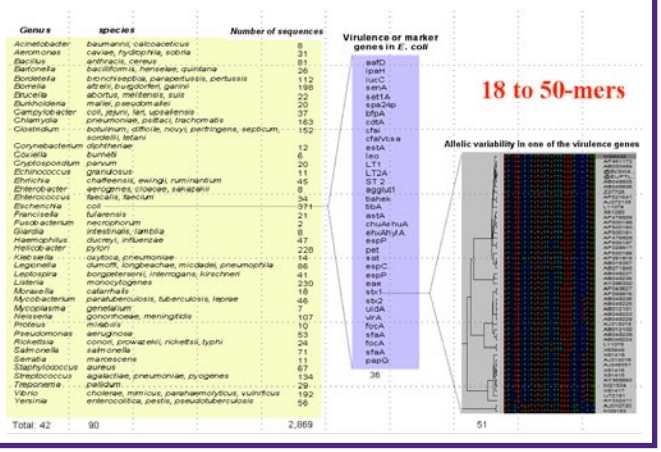
Hundreds of genomes are now available!



Activities (as in VFAR) are not always available!

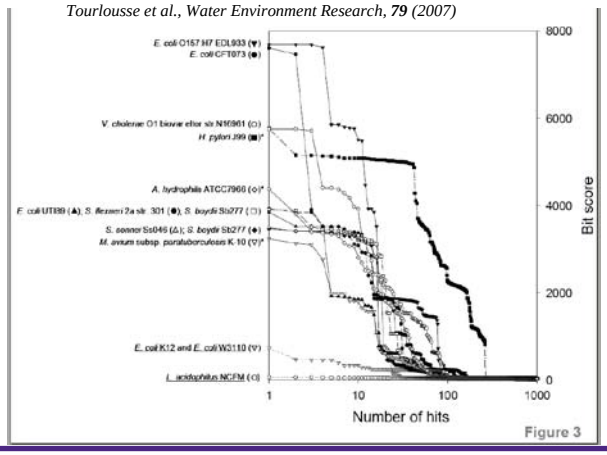


Virulence and Marker Genes (VMG) Database: Dynamic



CCL Organisms Compared to the VMG Database

*Tourlousse et al., Water Environment Research, **79** (2007)*

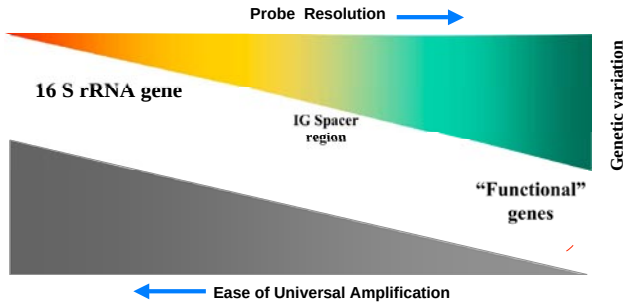


FACTORS RELATED TO MONITORING

Micro-fuidic Chip for 20 Waterborne Pathogens

1. *Aeromonas hydrophila*
2. *Burkholderia pseudomallei, mallei*
3. *Campylobacter jejuni*
4. *Clostridium perfringens*
5. *Enterococcus faecalis, faecium*
6. *Escherichia coli, Shigella*
7. *Helicobacter pylori*
8. *Klebsiella pneumoniae*
9. *Legionella pneumophila*
10. *Leptospira interrogans*
11. *Listeria monocytogenes*
12. *Mycobacterium avium, paratuberculosis, tuberculosis, leprae*
13. *Pseudomonas aeruginosa*
14. *Salmonella typhimurium* DT104
15. *Staphylococcus aureus*
16. *Vibrio cholerae, mimicus, vulnificus*
17. *Vibrio parahaemolyticus*
18. *Yersinia enterocolitica, pestis, pseudotuberculosis*
19. *Cryptosporidium parvum, hominis*
20. *Giardia lamblia, intestinalis*

16S & 23S rRNAs vs. VMGs



Multiplex Amplification- A Must!

Multiplex PCR-amplification followed by **DNAchip-based amplicon identification**

Without Multiplex Amplification

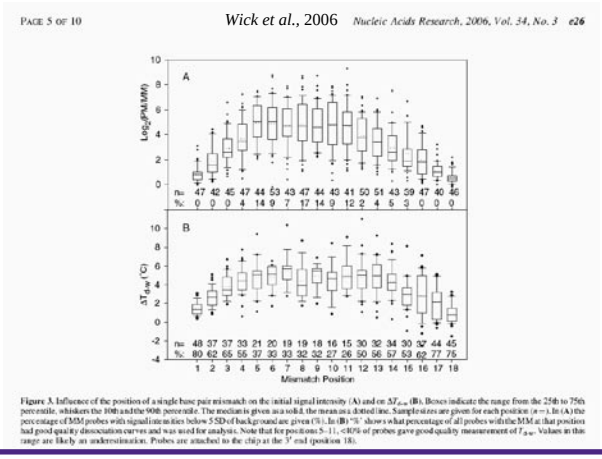
~1 % of the population

With Multiplex Amplification

DNA
multiplex PCR-amplification
up to 30 amplicons
in 1 reaction
mixture of amplicons
labeling of amplicons
identification of amplicons using DNAchip

0.01 to 0.0001%

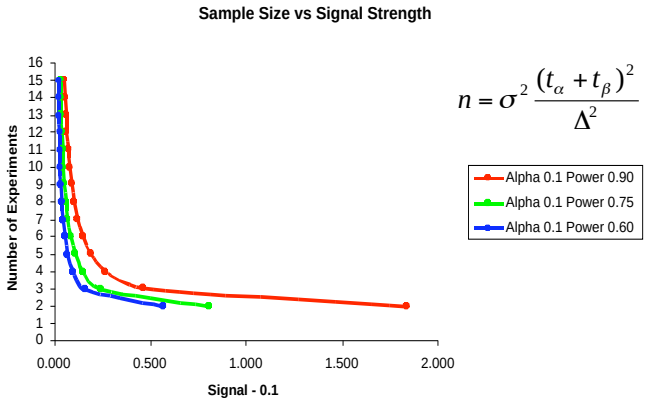
Limits of Hybridization: 18-mer Probe



PAGE 5 OF 10 — *Wick et al.*, 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 3* e26

Figure 3. Influence of the position of a single base pair mismatch on the initial signal intensity (A) and on ΔT_{dup} (B). Boxes indicate the range from the 25th to 75th percentile, whiskers the 10th and the 90th percentile. The median is given as a solid, the mean as a dotted line. Sample sizes are given for each position ($n = 1$ in (A) the percentage of MM probes with signal intensities below 5 SD of background are given (%). In (B) '%' shows what percentage of all probes with the MM at that position had good quality dissociation curves and was used for analysis. Note that for positions 5–11, <30% of probes gave good quality measurement of T_{dup}. Values in this range are likely an underestimation. Probes are attached to the chip at the 3′ end (position 18).

Replication of Complex Target Mixtures



$$n = \sigma^2 \frac{(t_\alpha + t_\beta)^2}{\Delta^2}$$

Screening for All Known VMGs

*Tourlousse et al., Water Environment Research, **79** (2007)*

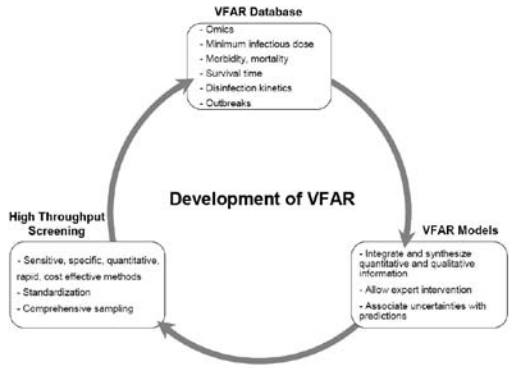


Figure 4

Post-doctoral Associates:

Lukas Wick	*E. coli* chip, Protocols
Jean Marie Rouillard	Probe design
Yongmei Xia	Target synthesis
Trinh Pham	Goal Labeling

Doctoral candidates:

Robert Stedtfeld	Pathogen chip/
Sam Baushke	Bioinformatics/PCR-chip
Dieter Tourlousse	Functional genes
Ruifang Xu	Protocol optimization
Yu Yang	Target gene
Munir Ahsan	Chip modeling

Research Associates/MS:

Sarah Miller	Time optimization
Vidya Srinivasan	Gold laebeling

Undergraduates:

Amanda Herzog	Hybridization

PIs:

James Tiedje	Syed Hashsham	Erdogan Gulari
James Cole	Joan Rose	Thomas Whittam

MICHIGAN STATE UNIVERSITY

CHIP TEAM

Funding

National Institutes of Health-NCRR
Michigan economic Development Corporation
MSU Foundation
Department of Defense

A Bioinformatic Approach to VFAR Analysis and Characterization

EPA QSAR/VFAR Workshop
20-21 June 2006

R. Paul Schaudies, Ph.D.
schaudiesr@saic.com

Molecular Radar™ Biological Identification Technology

- Highly multiplexed nucleic acid hybridization based approach
- Target unique and virulence related genetic regions
- Microarray format allows for identification of tens of thousands of individual sequences in parallel
- "Complete" genetic characterization within 4-24 hours
- Customizable levels of resolution

2

Capabilities Offered

- Simultaneously identify multiple pathogens
- Strain level resolution
- Identify signs of genetic engineering
- Characterize unknown organisms
- Virulence factors and antibiotic resistance characterization
- Functional equivalent of a 10,000-fold multiplexed PCR reaction
- Technology is adaptable to multiple platforms and applications

3

System Concept

Assay Development

- Computer identification of informative DNA/RNA sequences
- Identification of candidate oligonucleotides
- On-chip synthesis of oligos

Routine Sample Analysis

- Whole-genome amplification with label incorporation
- Hybridization on chip
- Spot profile identifies sequences present in original sample

4

Unique Sequences Generated by FIGUR Software

Accession Number	Organism	Distribution of Unique Sequences	Unique Bases	% Unique
NC_003997	*Bacillus anthracis* chromosomal		232331/5227293	4.44%
NC_001496	pXO1 plasmid		62372/181654	34.34%
NC_007323	pXO2 plasmid		48345/94829	50.98%
NC_003909	*Bacillus cereus* chromosomal		584913/5224283	11.20%
NC_005707	pBC10987 plasmid		91907/208369	44.11%
NC_004721	pClin15 plasmid		12095/15100	80.10%
NC_005957	*Bacillus thuringiensis* chromosomal		159941/5237680	3.05%
AL_731825	pBToxis plasmid		48871/127923	38.20%
NC_005567	pGI3 plasmid		8365/11365	73.60%
NC_003143	*Yersinia pestis* chromosomal		96032/4653728	2.06%
NC_002144	pYC plasmid		4475/5919	75.60%
NC_003132	pPCP1 plasmid		4764/9612	49.56%
NC_004835	pMT1 plasmid		26492/100984	26.23%
NC_006155	*Yersinia pseudotuberculosis* chromosomal		261660/4744671	5.51%
*	pYPtb32953 plasmid		22572/27702	81.48%
*	*Yersinia enterocolitica* chromosomal		108867/4615899	2.36%
NC_005017	pYVe8081 plasmid		8802/67720	13.00%
NC_006570	Francisella tularensis		1304572/1895998	68.81%
			1370214/3886914	35.25%

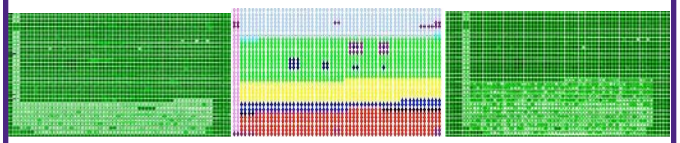
5

SAIC VER 1 Pathogen Array

- Sequences selected following initial screening arrays with SAIC funding
- Organisms arrayed in groups to aid rapid visual analysis
- Bioinformatics required for detailed strain level analysis

6

Bacillus anthracis Ames vs Sterne on SAIC VER 1 Array



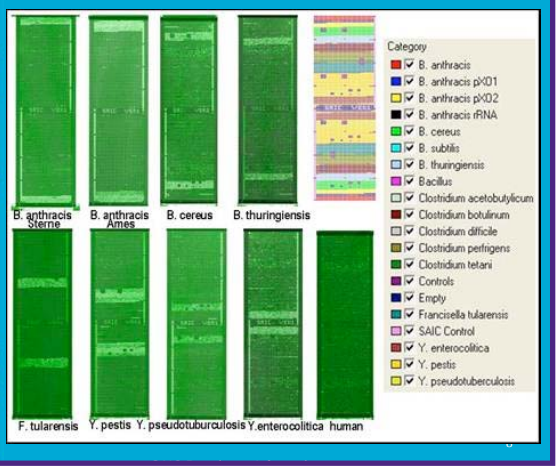
B anthracis Sterne

B anthracis Ames

ID	Spot Color
pXO1	Blue
pXO2	Yellow
Genomic	Red
Controls	Purple

7

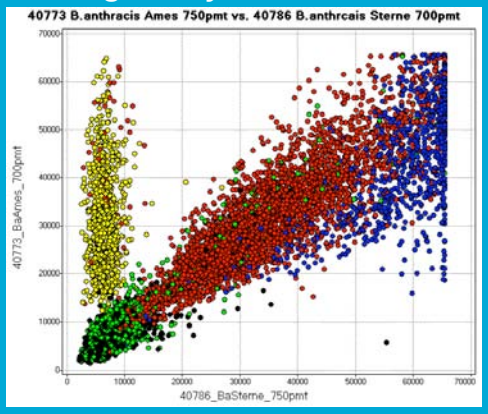
SAIC VER1 Array Hybridizations



Examples of Available Arrays

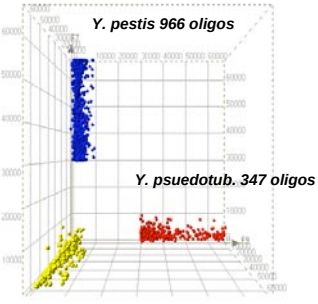
Array Design	Array Name	Array Content	# oligos
1672	Bacillus Array	*B. anthracis*	2000
		B. thuringiensis	1123
		B. cereus	2000
		Bacillus Plasmids	621
		Virulence Genes	60
		Antibiotic Resistance Genes	196
1683/84/85	Mixed Bacterial Array	*F. tularensis*	4544
		Brucella suis	144
		Brucella melitensis	167
		Brucella abortus	176
		Burkholderia mallei	4122
		Burkholderia pseudomallei	4537
		Escherichia coli k12	343
		Escherichia coli O157:h7	421
		Escherichia coli plasmids	46
1692/93	Clostridium Array	*Clostridium botulinum*	4256
		Clostridium tetani	1226
		Clostridium perfringens	4532
		Clostridium plasmids	167
1644/45	Yersinia Array	*Yersinia pestis*	1822
		Yersinia pseudotuberculosis	4127
		Yersinia enterocolitica	2127
		Yersinia plasmids	2473
1682	Mixed Virus Array	Variola and 4 related species	41
		Ebola	1198
		Dengue Fever	762
		Marburg	678
		Lassa Fever	504
		Rift Valley Fever	718
		Machupo	502
		CCHF	974
		HIV	878
		West Nile	558
		Adenovirus	1012
		Japanese Encephalitis	716

Screening Array Ames vs Sterne



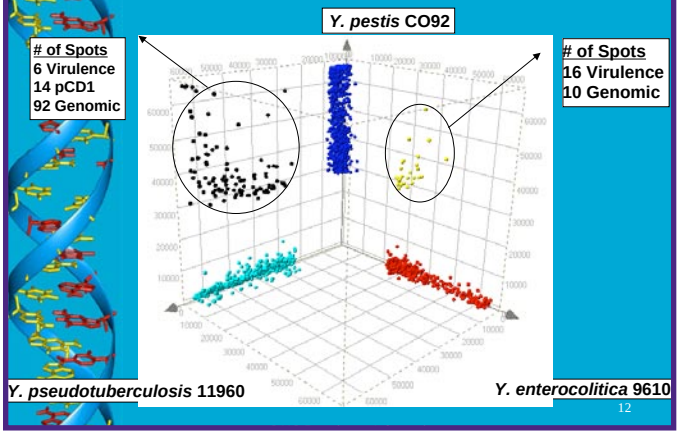
10

Species Level Differentiation for *Yersinia*



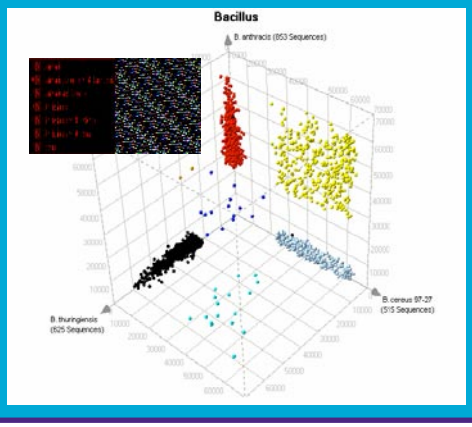
11

Virulence Factors for *Yersinia*



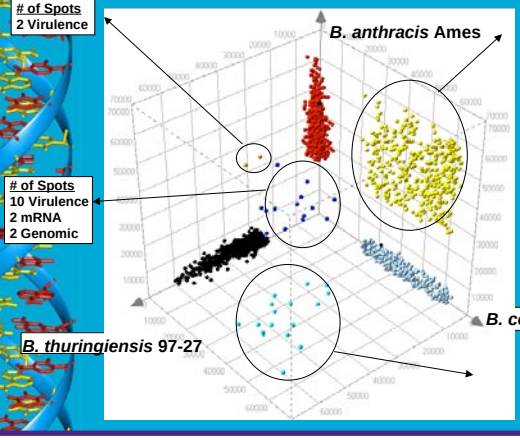
12

Bacillus Differentiation Presented in 3-D Format



13

Virulence Factors for *Bacillus*



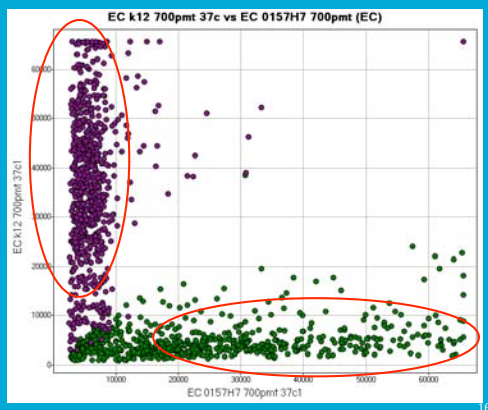
14

Microarray Design For Various Pathogens

Food/Water Testing Array Content	# oligos
Aeromonas hydrophila	560
Aeromonas punctata plasmid pFBAOT6	467
Brucella abortus	125
Brucella melitensis	500
Burkholderia mallei	750
Burkholderia pseudomallei	750
Calicivirus	148
Camphylobacter jejuni	750
Clostridium botulinum	750
Coxiella burnetii	750
E. coli K12	734
E. coli O157:H7	750
Helicobater pylori	750
Hepatitis D	25
Listeria monocytogenes	750
Norwalk Virus	40
Pseudomonas aeruginosa	850
Rickettsia conorii	850
Salmonella enterica	850
Shigella flexneri	850

15

E. coli K12 vs *E. coli* O157:H7



16

Summary

- Molecular Radar™ provides high fidelity identification and virulence factor characterization of microorganisms
- We have achieved resolution down to the level of strain for pathogens and near-neighbor organisms
- We can design and validate arrays for any DNA or RNA containing organism at desired level of resolution
- Array can be tailored to different levels of fidelity
- Capability exists today to analyze samples

17

From Reactivity to Regulation: Integrating Alternative Techniques to Predict Toxicity

Mark Cronin

School of Pharmacy and Chemistry
Liverpool John Moores University
England

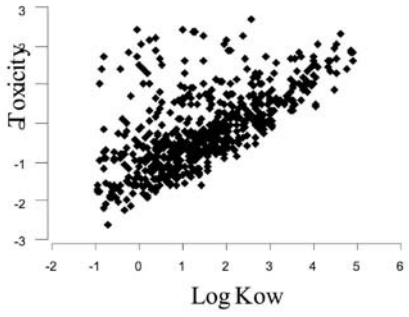
Models

Mechanisms

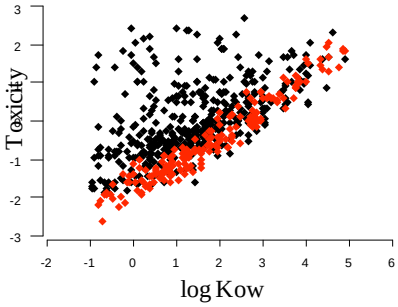
Modes

Madness

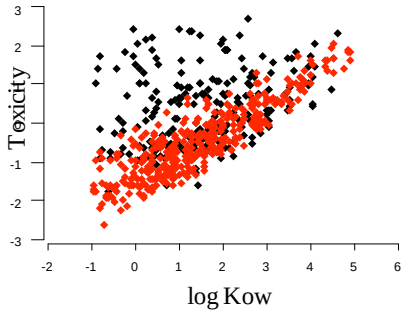
Cytotoxicity vs Hydrophobicity for Approximately 500 Chemicals



An Unspecific Mechanism (Non-Polar Narcosis) is Easily Predicted

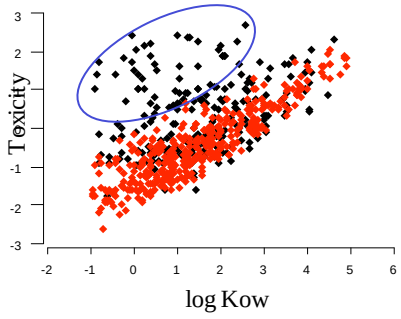


Unspecific Bioreactive Compounds



Toxicity = 0.65 log K_{ow} – 0.34 E_{lumo} – 1.11
n = 353 r^2 = 0.86 s = 0.35

Toxicity of Specifically Acting Electrophiles is Underpredicted



The Toxicity of Specifically Acting Electrophiles is Poorly Modelled by QSAR Approaches

we are not very good at parametrising reactivity

quantification of reactivity is difficult

reactivity is poorly quantified

reactivity is not well parametrised

θυαντιφιχατιον οφ ρεαχτ ιωιτψ ισ διφφιχυλτ

Quantitative Assessment of Reactivity: Glutathione Reactivity Assay

- An olefin conjugated to a carbonyl group, is inherently electrophilic
- Potential to act by Michael-type nucleophilic addition to macromolecules
- Measured GSH reactivity is related directly to cytotoxicity

$\log \text{Toxicity} = 0.95 \log \text{GSH}_{\text{reactivity}} + 0.54$

$n = 46 \quad r^2 = 0.91 \quad s = 0.27 \quad F = 460$

Schultz TW et al (2005) *SAR QSAR Environ. Res.* 16: 313–322

Reactive Mechanistic Domains: Electrophiles in Toxicology

- Michael acceptor
- S_NAr
- S_N2
- Schiff base
- Acyl transfer
- Metabolically activated compounds

In Chemico Assays for Reactivity: Spanning the Electrophilic Mechanisms

Other Toxicity Endpoints with Electrophilic Mechanisms

- Skin sensitisation
- Respiratory sensitisation
- Carcinogenicity/ mutagenicity
- Skin irritation
- Inhalation irritation
- Liver toxicity
- Idiosyncratic drug toxicity

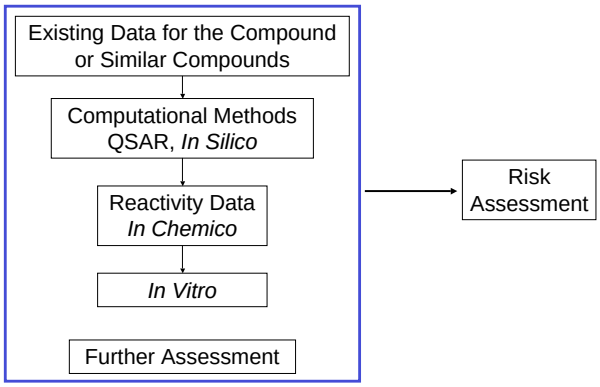
Are they the Same Mechanisms?

- Chemically the mechanisms are the same, the target and endpoint differ
- Useful information may be obtained if we can extrapolate this information

Application to Regulatory Problems

- **New chemicals legislation will require**
  - **Increased risk assessment**
  - **Potential increase in animal testing**
  - **Increase in cost**
- **There is an incentive for the greater use of alternative methods**
- **We know we have a problem predicting "reactive toxicity"**
- **How can we implement our knowledge of reactive toxicity across endpoints**

Alternative Methods: Integrated Testing Strategy



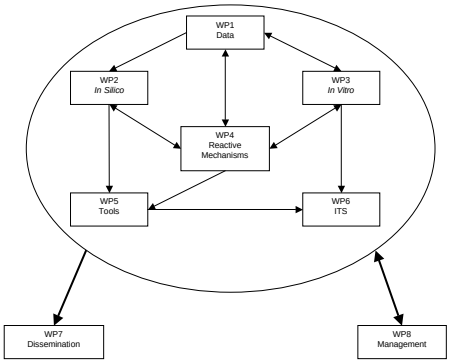
Prediction Models for Reactive Toxicity: Application of *in Chemico* Measurement

Using gutathione reactivity as a model soft nucleophile:

- If Michael addend, $pIGC_{50} = 1.01\ pEC_{50}(GSH) + 0.57$
  Schultz TW et al (2005) *SAR QSAR Environ. Res.* 16: 313-322
- If R(GSH) > -0.55, chemicals are Skin Sensitisers
  Aptula AO et al (2006) *Toxicol. in Vitro* 20: 239-247

Can we go *in chemico* to *in silico*?

Defra LINK Project: Work Plan



Conclusions

- **Specific reactivity is poorly parametrised in toxicology, but underpins many endpoints**
- **Measuring reactivity *in chemico* has been shown to assist in predicting reactive toxicity better**
- **Needs for more reactivity data, computational capability and strategies for implementation**

Integrated QSAR-PBPK modeling for risk assessment

Kannan Krishnan
Université de Montréal, Canada

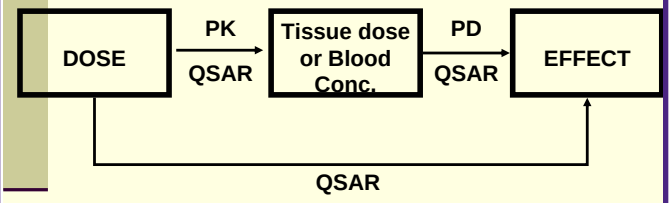
Outline

- Introduction
- QSAR-PBPK: Development
- Risk assessment applications
- Conclusion

QSARs – Current Paradigm

- NOAELs vs chemical structure or props.
- Context-specific QSAR
- *Duration of exposure (short-term)*
- *Oral route*
- *Species of interest (Rat)*
- For a different route, species & duration
  - Develop new sets of QSARs
  - Develop `extrapolable` QSARs

QSARs – An alternative paradigm

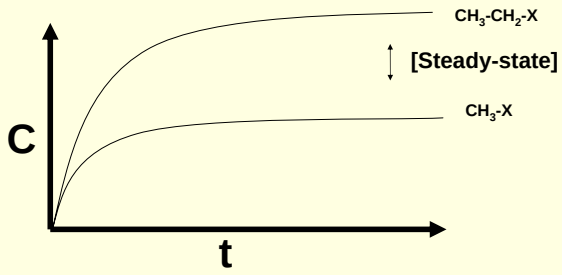


- Relative contribution of the TK and TD processes
- Extrapolations based on TK determinants

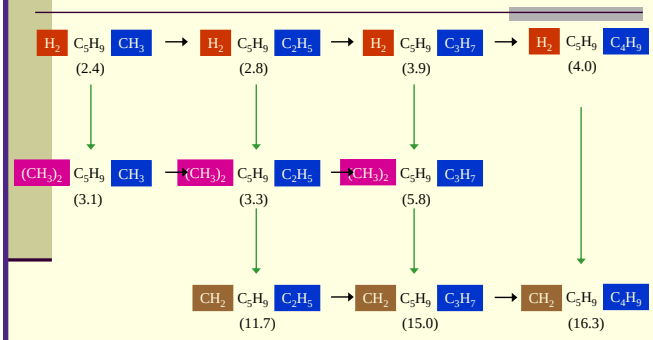
QSAR: PK-TK

- QSAR models are based on response-specific dose level for each species
- No efforts on the relationship between structure and internal dose
- Can we develop QSARs for pharmacokinetic profiles ? (changing as a function of route, dose and species)
- Inhalation, steady-state, rats…

Blood Concentration at Steady-state



Blood concentration *vs* structure



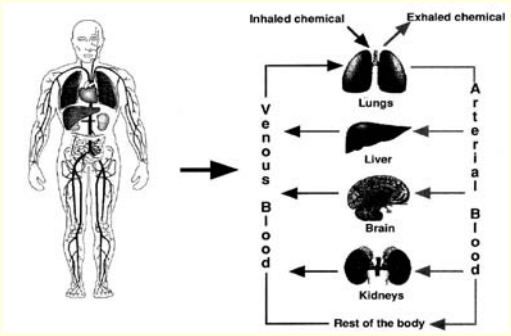
Structure vs Blood concentration

$H_2\ C_5H_9\ CH_3$

C = [2 x –3.25] + [1 x 6.8] + [1 x –1.2]

= 3.25 μM

PBPK Models

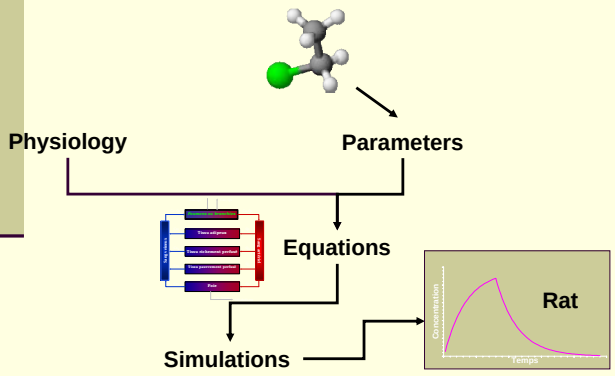


Physiology, partition coefficients, metabolic clearance

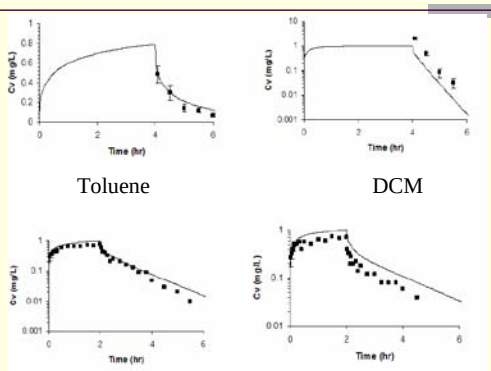
QSARs for PBPK Parameters

- Fragment constant approach
  - $P_{pbpk} = \Sigma\ n_f \cdot C_f$
- Multilinear regression (SPSS®)
- 46 VOCs, Fragments: CH_3, CH_2, CH, C, C=C, H, Cl, Br, F, B-ring, 2 E1 substrates
- Cross-validation, external validation

QSAR-PBPK Modeling

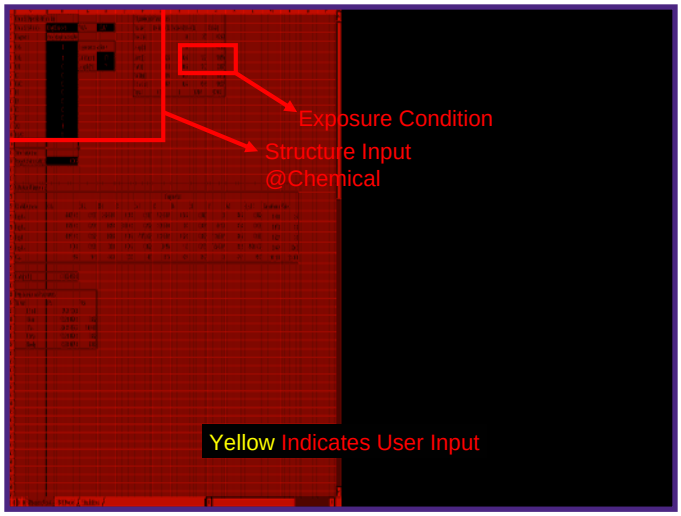


QSAR/PBPK modeling - Rat

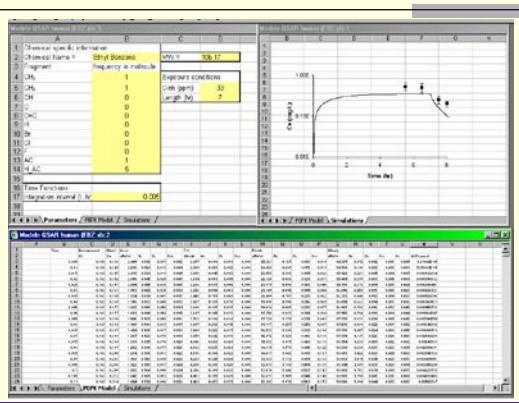


Chemicals in the application domain

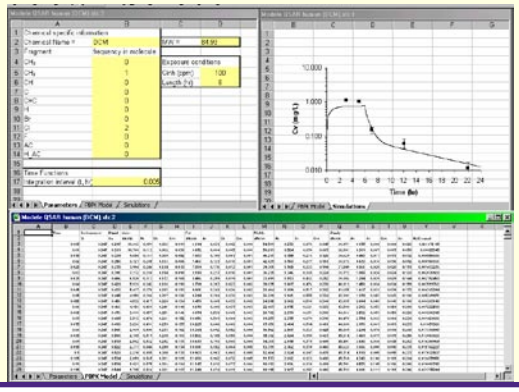
- Trifluoromethane
- Dichlorofluoromethane
- Bromodichloromethane
- Bromoform
- Dibromofluoromethane
- Bromoethane
- 1,1,1-Tribromoethane
- **2,2-Dichloro-1,1,1-trifluoroethane**
- 1,2-Dibromo-1,1,2-trifluoroethane
- 1-Chloropropene
- 1,2-Dichloropropene
- **1,3-Dichloropropene**
- 1,1-Dibromopropene
- 1-Bromo-2-chloropropene
- Pentane
- Tribromoethylene
- Tetrabromoethylene
- 1-Bromo-2-chloroethylene
- m-Dichlorobenzene
- Propylbenzene
- **1,2,4-trimethylbenzene**
- m-chloromethylbenzene
- **Ethyl benzene**



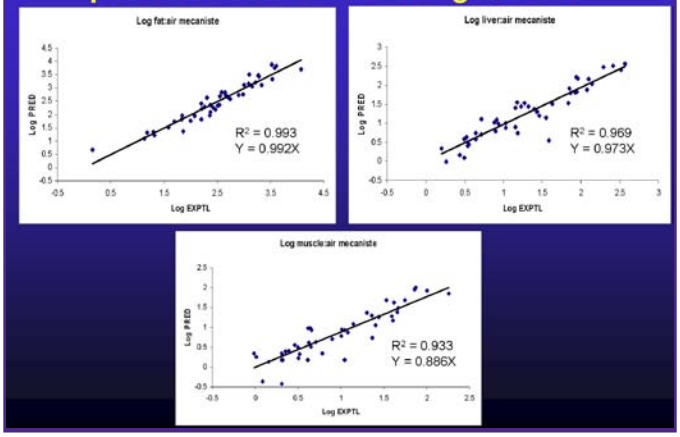
QSAR/PBPK model – Ethyl benzene



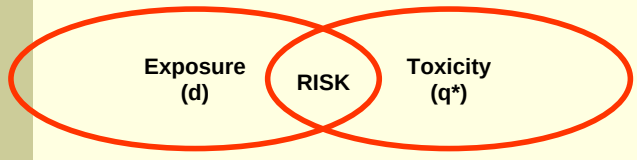
QSAR/PBPK model:Dichloromethane



Interspecies extrapolation of tissue:air partition coefficients using QSARs

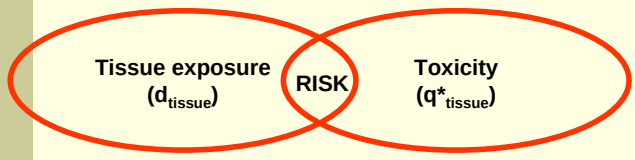


Risk Assessment



- Risk = $q^* \cdot d$

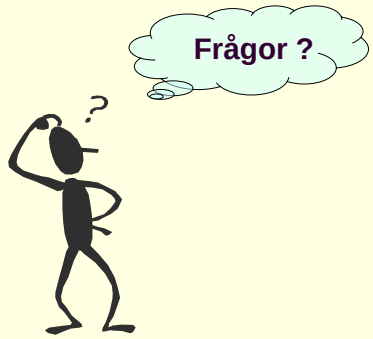
Risk Assessment



- d_{tissue} = Human PBPK model
- q* = Animal PBPK model

QSAR-PBPK Models in risk assessment

- QSAR-PBPK models facilitate internal dose based risk assessment (lethal and non-lethal effects)
- Influence of exposure concentrations, routes and scenarios can be examined
- Effects on specific sub-populations can be evaluated
- Modeling of multiroute exposures for risk assessment applications



Weight of Evidence and Mode of Action in Predictive Toxicology

Dr. Andrew Maier
and
Dr. Raghu Venkatapathy

June 21, 2006

Weight of Evidence (WOE) in Risk Assessment

- Risk Assessment Initiatives
  - U.S. EPA Cancer risk assessment – requires addition of a "weight of evidence narrative"
  - Increasingly used in Hazard Screening Algorithms (e.g., Health Canada ComHaz tool)
  - Weight of evidence characterized by use of "totality of the evidence" in making decisions about causality
  - **Emphasis on "Totality" has opened door for predictive toxicity tools**
- Evolving concept driven by
  - Improved biology understanding (understanding of the mode of action or MOA)
  - Increased sophistication and validation of alternative study designs and consideration of study design (e.g., gene knock-outs)
  - Improved quantitative tools (including toxicogenomics and QSAR)

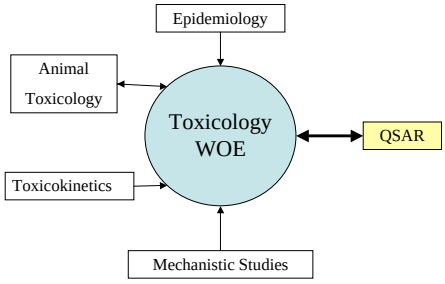
Role of Predictive Toxicity

Empirical Data Confidence

Consistency	Inadequate	Adequate
Yes	Collect Data	Characterize Risk
No	Collect Data	Resolve Conflict

= Candidates for Predictive Toxicity

WOE and QSAR



Tools for Evaluating WOE

- Hill criteria for causality
- Expert judgment
  - Peer review/consultation
  - Expert elicitation techniques
    - Survey approaches
    - Software tools
- Quantitative tools
  - Decision and Uncertainty Analysis
  - Bayesian Analysis

Biology understanding is needed for interpreting results

	Genotoxicity QSAR Modules		
Type of Damage	**Mouse Lymphoma**	**Chromosome Aberrations in CHO cells**	**Ames Bacterial Mutagenicity**
Point mutation	Yes	No	Yes
Oligonucleotide insertion or deletion	Yes	No	Yes
Allele Loss	Yes	No	No
Small Chromosome alteration	Yes	?	No
Large Chromosome alteration	Yes	Yes	No
Aneuploidy	?	Yes	No

Adapted from M. Moore (2004)

Biological Understanding

- Our level of understanding of the underlying biological basis of toxic responses represents a continuum.
- For risk assessment we often distinguish between knowing the mechanism of toxicity versus the mode of action.
- Mechanism of toxicity refers to a detailed understanding to the cellular and subcellular level of the basis for toxicity.
- Mode of action refers to a less detailed level of understanding, but ability to identify key precursor steps in the pathway to toxic response.

Defining Mode of Action

- A critical challenge in integrating mode of action data in *global* QSARs is defining appropriate predictors:
- What does mode of action mean?
  - Target organ? (liver toxicity)
  - General cellular response (necrosis)
  - Subcellular target (ATP synthesis disruption)
  - Potential presence of reactive moiety (electrophiles, oxygen radicals)

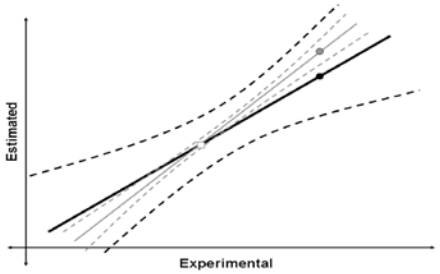
Problem Statement

- Currently SARs and QSARs are often used as independent tools, a practice that does not optimize what can be learned when the varying approaches are used in a coordinated manner.
- Approaches for developing consensus modeling approaches that use biology understanding (MOA) for integrating SAR and QSAR models are needed.

Goal - Maximizing Use of Biology

- Mechanism of action known. Develop mechanistic QSARs – excellent predictivity – but limited applicability
- Mode of action data available. develop hybrid or MOA-informed QSAR – balance of predictivity with applicability
- Biology unknown – use global statistical QSAR – decreased predictivity – but broad applicability

Using MOA to Refine Statistical QSARs



Basic principles would indicate that correlations of similar chemicals would improve prediction.

MOA in Logistic Regression

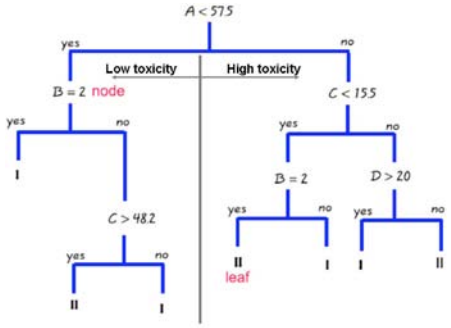
- Endpoint specific SAR models are often designed as either expert system-based models or statistical models
- A hybrid approach that uses logistic regression analysis with a dummy dependent variable coded 0 and 1 (for negatives and positives, respectively) can allow the input of key data derived from MOA decision rules.
- Probability of end point toxicity is:

$$\log\left(\frac{\pi}{1-\pi}\right) = b_o + b_1x_1 + b_2x_2 + \ldots. + b_mx_m$$

Using "Omics" for Binning

- Toxicogenomics, proteomics, metabonomics already in use for hazard identification
- Used to identify MOA for hypothesis testing
- Public databases will be increasingly populated for data mining
- These data can serve as sorting variables to enhance QSAR development

Regression Tree Approach



Role of MOA based hybrids

- Inform the interpretation of global QSARs - e.g., identifying critical endpoints.
- Serve as sorting variables to bin chemicals for development endpoint or MOA-specific QSARs.
- If MOA biomarkers are used as dependent variable, then serve as QSAR endpoint verifiable by relatively non-invasive tests.
- Binned (or nodes) can be used for assigning potency using group average or "Threshold of Concern" approach.

Conclusions

- WOE evaluation represents a maturation in chemical risk assessment
- Critical use in resolving conflicting data – are assays or predictive tools testing the same thing? Can differences be explained by the MOA understanding?
- Advances in basic biology (molecular and cellular biology), chemistry (computational chemistry), and mathematics (better statistical and dose-response tools) should be used by the risk assessment community
- Tool developers should make full use of our mode of action understanding

Novel Approaches to QSAR & VFAR Modeling

William (Bill) Welsh

UMDNJ CompTox Center

welshwj@umdnj.edu

New Jersey Environmental Bioinformatics & Computational Toxicology Center

ebCTC

http://www.ebCTC.org

Funded with support from the
U.S. EPA, National Center for Environmental Research
Science to Achieve Results (STAR) Program

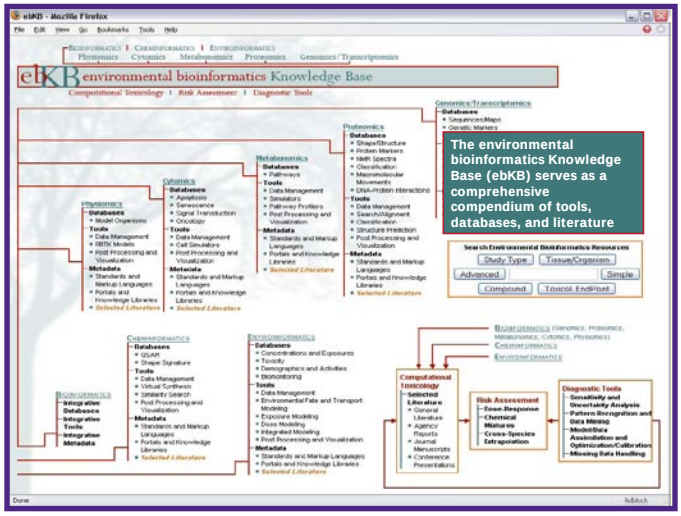
William Welsh, Center Director

Consortium Members

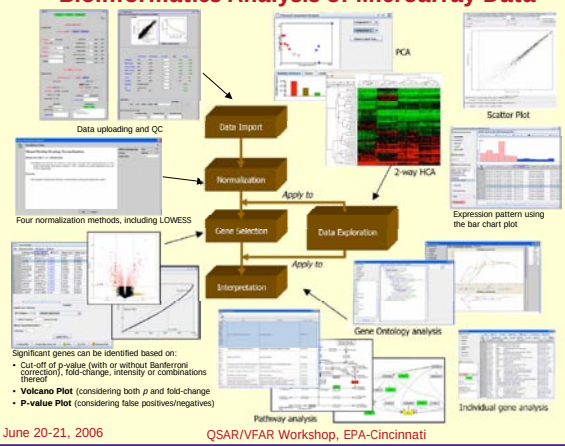
Center for Toxicoinformatics, NCTR

Major Research Thrusts

- DORIAN Computational Toxicology System that spans the Source->Dose->Outcome continuum
- The Environmental Bioinformatics Knowledge Base (ebKB)
- *ebTrack*, a toxicological bioinformatics platform to process genomics, proteomics and metabonomics data
- Hepatocyte Metabolic Model for Xenobiotics
- **ChemTox, a suite of chem-informatics tools for toxicant identification, prioritization, characterization**



ebTrack System: Extension of *Array Track*
- Bioinformatics Analysis of Microarray Data -

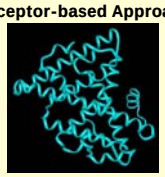


Overview of QSAR-based Approaches

- **Decision Forest (DF)**
  - fast consensus modeling technique that quantifies prediction confidence
- ***Shape Signatures***
  - enables fast large-scale screening of query chemicals against databases based on similarity in shape and other biorelevant molecular features
- ***Polynomial Neural Network* (*PNN*)**
  - generates optimal linear or nonlinear QSAR models in parametric form
- **Virtual High-Throughput Screening (vHTS)**
  - predict & quantify ligand binding affinity to proteins
  - provide insights into mechanism of action (toxicity pathways)
  - assess validity of cross-species extrapolation (e.g., rat vs. human)

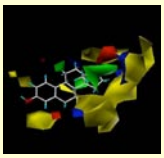
Integrated Approach

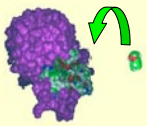
Receptor-based Approaches



Predictive Molecular Toxicology

Ligand-based Approaches

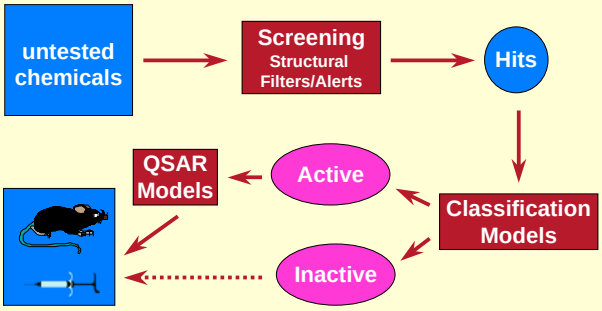




Virtual Screening

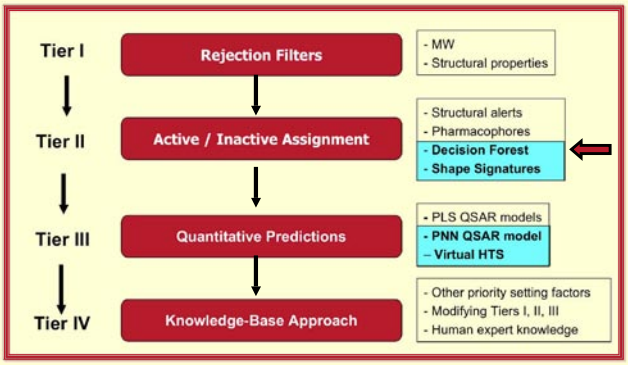
Computational Screening Paradigm

- Priority Setting -



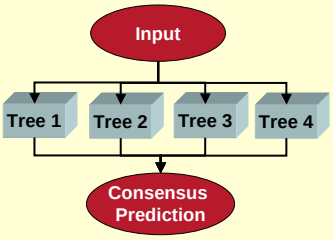
Schematic of Hierarchical Screening Framework

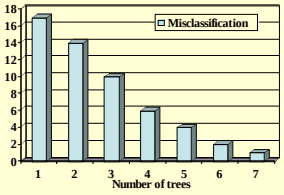
- addresses the need to minimize *false negatives* and *uncertainties* -



Decision Forest (DF)

- improve classification by combining individual models -



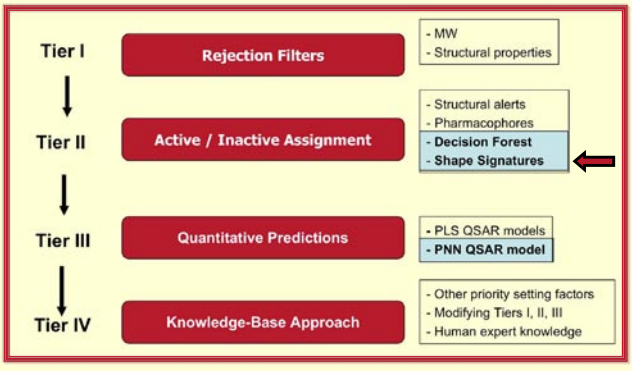


Key Features

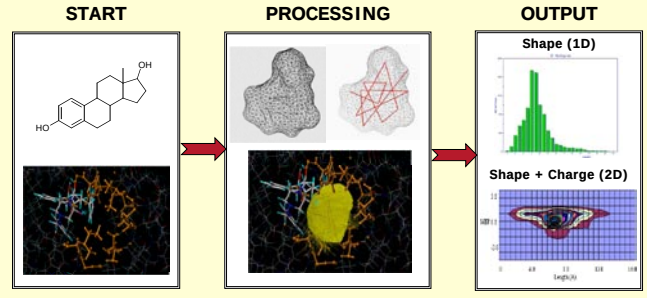
- Combining several independent yet predictive trees improves performance
- DF structure permits assessment of prediction confidence, reduces uncertainty
- Each tree consists of simple 'If-Then' branches, hence the DF is extremely fast

Schematic of Hierarchical Framework

- based on USFDA's EDKB -



Shape Signatures Tool



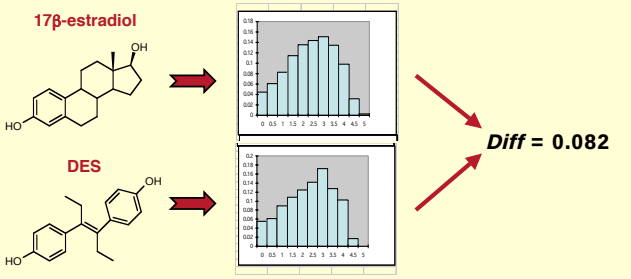
Small molecule or Protein binding pocket

Ray tracing to generate the raw data

1D and 2D Shape Signatures

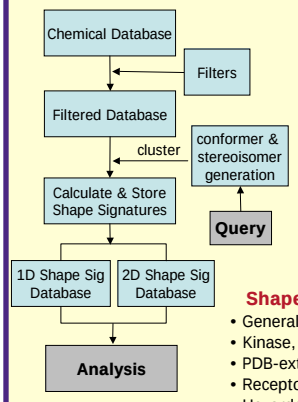
Shape Signatures Tool

molecules are compared by subtracting their histograms

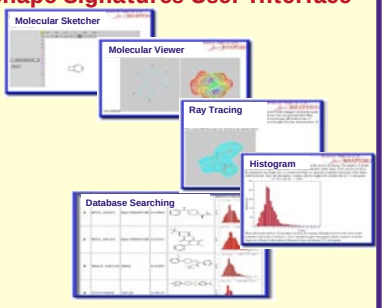


Small *Diff* value means that two molecules have similar shape and polarity

Flowchart



Shape Signatures User Interface

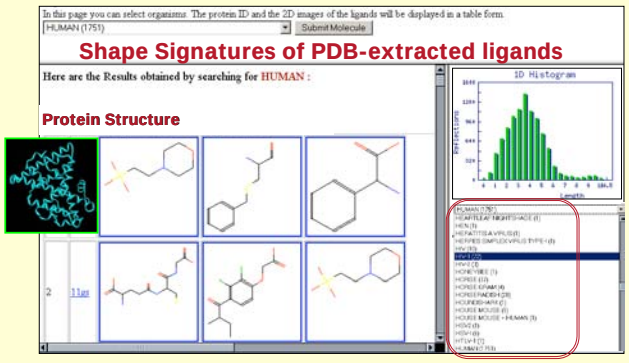


Shape Signature Databases

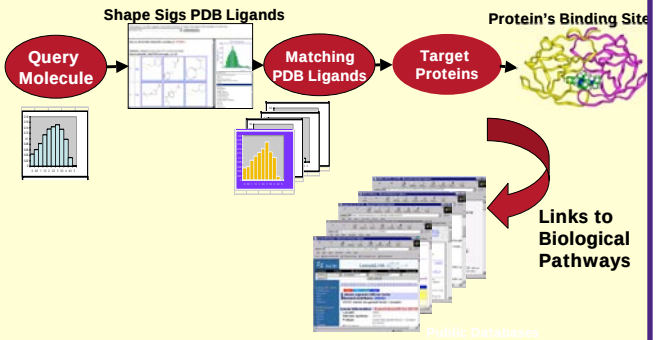
- General Database >4 million chemicals
- Kinase, GPCR, NR ligand databases
- PDB-extracted ligand database
- Receptor binding sites of 30,000 proteins (BWAs)
- Hazardous Chemicals (EDCs, H_2O CCLs, DSSTox, CWAs)

Chemical → Target Protein → Mechanisms

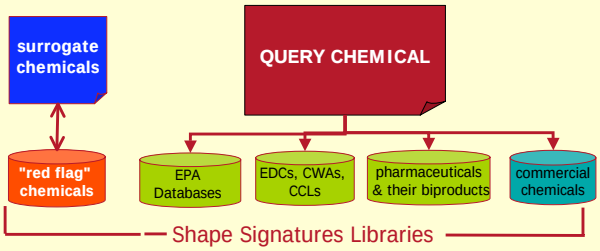
Protein Data Bank (PDB): World Repository of ~35,000
Protein-Ligand Crystal Structures (http://www.rcsb.org/pdb/)



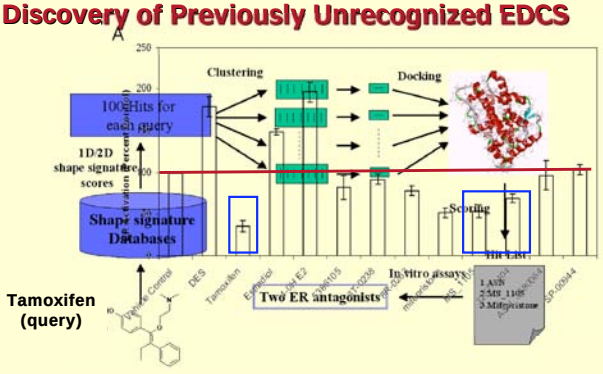
Molecules → Target Protein → Mechanism



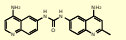
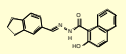
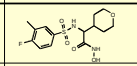
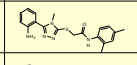
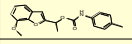
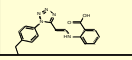
Identifying Problem Chemicals & Possible Surrogates

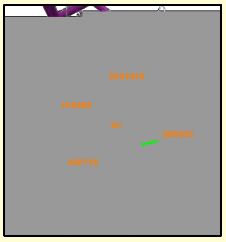


Discovery of Previously Unrecognized EDCS



Shape Signatures: Discovery of Anthrax LF Inhibitors

ID	Structure	Source	Docking Score	% Inhibition
QUERY NSC 12155 (known inhibitor)		NIH-NCI	33	95
QUERY LFI (known inhibitor)		Merck	39	-
--		Aldrich	42	98+
--		Asinex	38	98+
--		Bionet	36	98
--		Maybridge	36	97



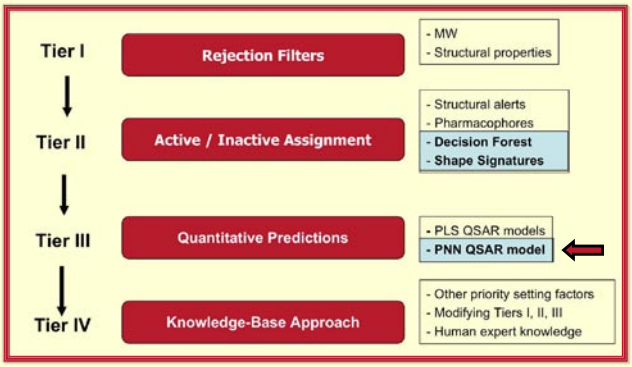
S503428 docked in the ligand binding pocket of anthrax LF.

Key Features of *Shape Signatures*

- Innovative: Encodes molecular shape and other biorelevant features in a single entity
- Non-congeneric: Finds hits missed by techniques that search on chemical (sub)structure
- User Oriented: fast, simple, expandable
- Versatile: works for any number or type of molecular species (organics, organometallics, ions, etc.)
- Applicable in ligand-based mode (ligand-ligand *similarity*) and receptor-based mode (ligand-receptor *complementarity*)

Schematic of Hierarchical Framework

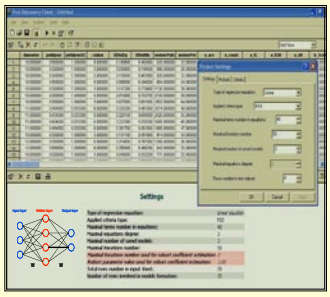
- based on USFDA's EDKB -



Polynomial Neural Network (PNN)

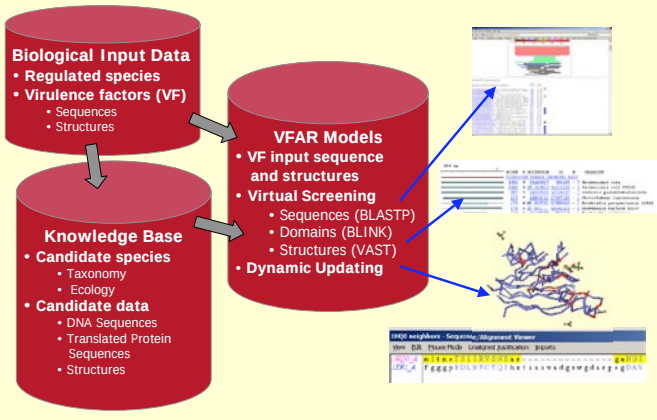
- combines the parametric form of PLS and the nonlinearity of ANNs -

Polynomial Neural Network

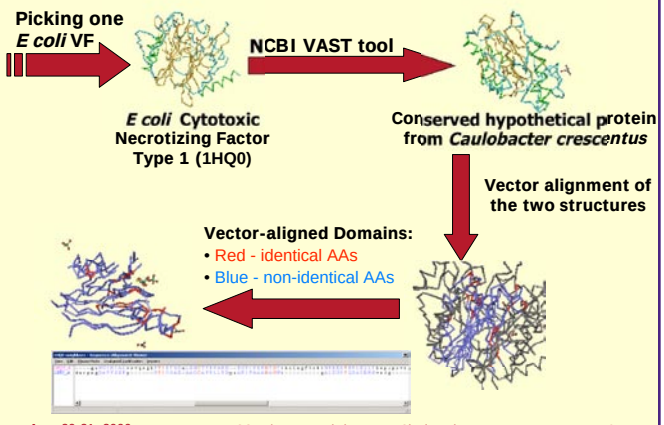


- Produces linear or non-linear QSAR models in parametric form
- User control of model complexity
- Insensitive to irrelevant variables and outliers
- Yields predictive models, even for sparse or noisy data sets
- Trains rapidly, thus amenable to large data sets
- Automatically selects best indep. variables; no preprocessing required
- Customizable to fit user's needs

Biological Data → VFAR Databases & Models



Bacterium → VF → VF Structure → Candidate Structure



Thank You!

welshwj@umdnj.edu

EUROPEAN COMMISSION
Joint Research Centre

ECB

Joint Research Centre

Role of the European Chemicals Bureau in Promoting the Regulatory Implementation of Estimation Methods

US EPA QSAR / VFAR Workshop, 21 June 2006

Andrew Worth

European Chemicals Bureau
Institute for Health & Consumer Protection (IHCP)
Joint Research Centre (JRC), European Commission
21020 Ispra (Va), Italy

http://ecb.jrc.it/QSAR E-mail: andrew.worth@jrc.it

ihcp

EUROPEAN COMMISSION
Joint Research Centre

ECB

Joint Research Centre

Outline

1. The Joint Research Centre (JRC) & the European Chemicals Bureau (ECB)
2. Use of estimation methods under REACH
3. ECB research in computational toxicology
4. ECB assessment of methods and models
5. Promoting (regulatory) acceptance and implementation
6. Training & capacity building

ihcp

EUROPEAN COMMISSION
Joint Research Centre

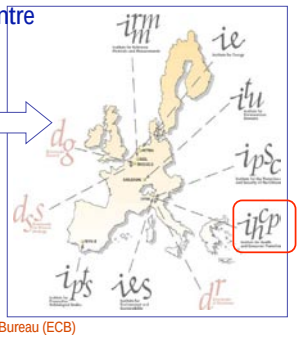
ECB

Joint Research Centre

The European Commission's Joint Research Centre

European Commission
Directorates-General
Directorates or Institutes
Units

JRC



European Chemicals Bureau (ECB)
European Centre for the Validation of Alternative Methods (ECVAM)
Physical & Chemical Exposure (PCE) Unit

ihcp

EUROPEAN COMMISSION
Joint Research Centre

ECB

Joint Research Centre

The European Chemicals Bureau: http://ecb.jrc.it

Assessment of chemicals

Existing Substances
New Substances
Biocides
Export / Import

IUCLID 5

REACH Support

Guidance & tools for industry & authorities

ECB

REACH IT & Informatics

REACH-IT for Chemicals Agency
IUCLID 5

Computational Toxicology

Development, validation, acceptance and implementation of estimation methods

EUROPEAN COMMISSION
Joint Research Centre

ECB

Joint Research Centre

Information requirements under REACH

Standard information requirements for chemicals are largely tonnage dependent, however:

- Annex VI Specific requirements are context-dependent
- Annexes VII-X Standard information requirements
- Annex XI "Adaptation" of standard information requirements:
  - replacing traditional test data with predictions or equivalent data
  - providing standard information at lower or higher tonnages
  - exposure-based waiving (Annexes VII & VIII, ≥100 tonnes)
  - providing additional information (if necessary)

⇒"Intelligent" rather than box-ticking approach to information gathering

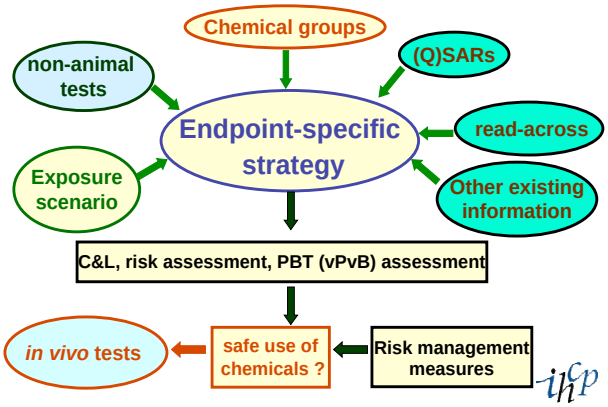
ihcp

EUROPEAN COMMISSION
Joint Research Centre

ECB

Joint Research Centre

Integrated Testing Strategies (ITS)



ihcp

EUROPEAN COMMISSION
Joint Research Centre

Annex XI of REACH – (Q)SARs

Results obtained from valid qualitative or quantitative structure-activity relationship models ((Q)SARs) may indicate the presence or absence of a certain dangerous property. Results of (Q)SARs may be used instead of testing when the following conditions are met:

- results are derived from a (Q)SAR model whose scientific validity has been established
- the substance falls within the applicability domain of the (Q)SAR model
- results are adequate for the purpose of classification and labelling and/or risk assessment, and
- adequate and reliable documentation of the applied method is provided

EUROPEAN COMMISSION
Joint Research Centre

Annex XI of REACH – Categories (1)

Substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group or "category" of substances.

Application of the group concept requires that physicochemical properties, human health effects and environmental effects or environmental fate may be predicted from data for a reference substance within the group by interpolation to other substances in the group (read-across approach). This avoids the need to test every substance for every endpoint.

... If the group concept is applied, substances shall be classified and labelled on this basis.

EUROPEAN COMMISSION
Joint Research Centre

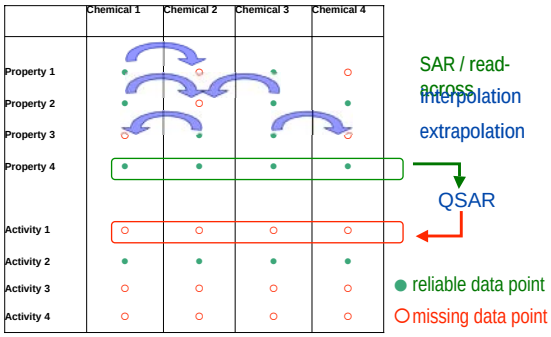
Annex XI of REACH – Categories (2)

In all cases results should:

- be adequate for the purpose of classification and labeling and/or risk assessment
- have adequate and reliable coverage of the key parameters addressed in the corresponding test method referred to in Article 12(2)
- cover an exposure duration comparable to or longer than the corresponding test method referred to in Article 12(2) if exposure duration is a relevant parameter, and
- adequate and reliable documentation of the applied method shall be provided

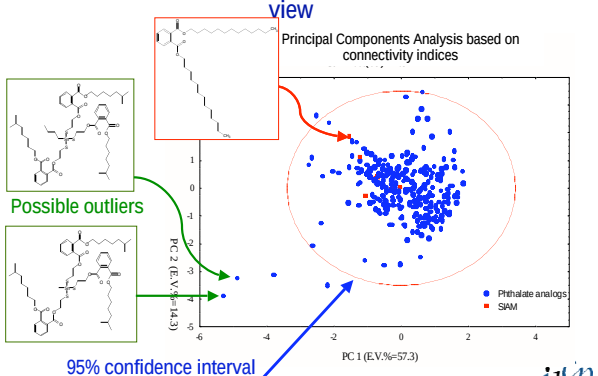
EUROPEAN COMMISSION
Joint Research Centre

Chemical category – administrative view



EUROPEAN COMMISSION
Joint Research Centre

Chemical category – QSAR view



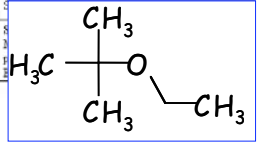
EUROPEAN COMMISSION
Joint Research Centre

Read-across assessment of ETBE

Table 1 Chemical structure of candidate analogs for ETBE

MTBE	Methyl trimethylacetate	TAME	2-Methyl-2-propanethiol
CAS No. 1634-04-4	CAS No. 598-98-1	CAS No. 994-05-8	CAS No. 75-66-1
MW: 88.15	MW: 116.16	MW: 102.18	MW: 90.18
MF: $C_5H_{12}O$	MF: $C_6H_{12}O_2$	MF: $C_6H_{14}O$	
SMILES: O(C(C)(C)C)C	SMILES: O=C(C(C)(C)C)OC	SMILES: COC(C)(C)CC	
Synonyms: Propane, 2-methoxy-2-methyl-, methyl t-butyl ether, Methyl	Synonyms: 2,2-dimethyl-propanoic acid, methyl ester, Methyl pivalate,	Synonyms: Methyl tert-amyl ether-, 2-Methyl-2-methoxybutane,	

2-ethoxy-2-methylpropane



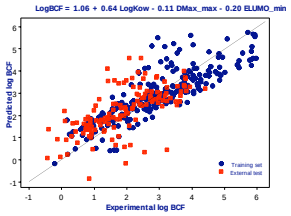
Physicochemical properties, mutagenicity, sensitisation, aquatic toxicity

EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

ECB research on computational toxicology (1)

1. QSARs for aquatic toxicity (& modes of action)
2. QSARs for bioaccumulation
3. QSARs for sensitisation
4. QSARs for endocrine disruption



EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

ECB research on computational toxicology (2)

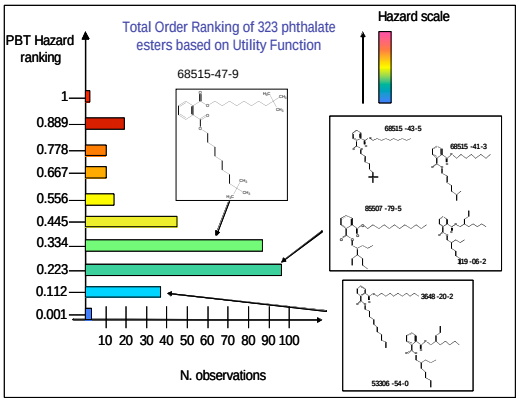
5. Methods for chemical similarity analysis and grouping
   - ECB workshop on TTC and grouping methods (Nov 05)
   - EU / OECD Guidance on grouping
6. Methods for descriptor-based ranking of chemicals
   - Workshop on Ranking Methods with Italian Chemometrics Society & Milan Bicocca University (2-4 October 06)
7. Methods for defining (Q)SAR applicability domains
8. Weight-of-evidence in hazard & risk assessment
   - ECB workshop on consensus modeling (Sept 05)
9. Computational nanotoxicology

EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

Chemometric Ranking Tools (1)

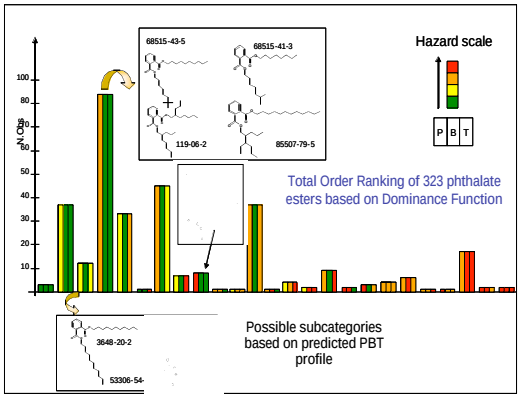
Possible subcategories based on PBT ranking



EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

Chemometric Ranking Tools (2)

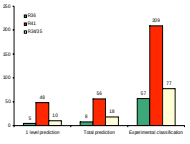


EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

ECB assessment of methods and models

1. QSAR validation studies (2005-2006)
   - Acute fish toxicity, skin penetration, skin sensitisation, steroid hormone receptor binding
2. Validation of BfR rulebases
   - Skin and eye irritation / corrosion
   - Eye irritation
3. Validation of TerraQSAR™ FHM model
4. Beta testing of AIM
5. Beta testing of AMBIT software for QSAR applications
   - http://ambit.acad.bg



EUROPEAN COMMISSION
Joint Research Centre

Joint Research Centre

Promoting (regulatory) acceptance

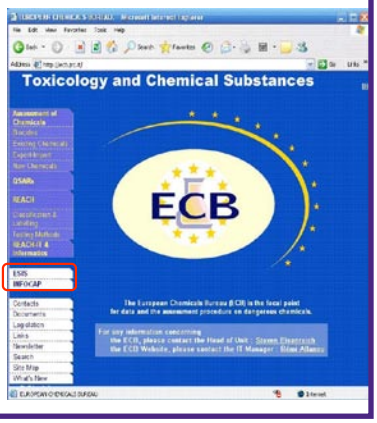
1. EU Working Group on QSARs
   - Capacity building among regulators and industry
   - Scientific & technical preparations for REACH
2. OECD *ad hoc* Group on QSARs
   - Principles for QSAR validation (adopted)
   - Practical guidance on QSAR validation (under review)
   - Case studies on regulatory acceptance (completed)
   - ECB hosted meeting on 8-9 June 06, Stresa, Italy
3. OECD Validation Management Group for Non-Animal Methods
   - ECB coordinates QSAR Task Group

EUROPEAN COMMISSION
Joint Research Centre

Promoting implementation: ECB QSAR Inventory

Joint Research Centre

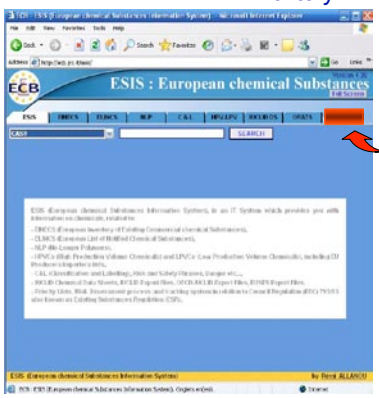
- ECB is designing a QSAR Inventory
- Oracle database
- Will be available via ECB Website
- Integration with ESIS (European chemical Substances Information System)
- Uploading of models via ECB website (QSAR Reporting Formats)
- Quality check of models by ECB

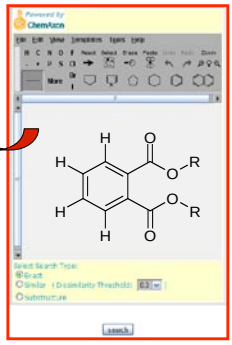


EUROPEAN COMMISSION
Joint Research Centre

Structure-searchable Interface to ESIS and QSAR Inventory

Joint Research Centre



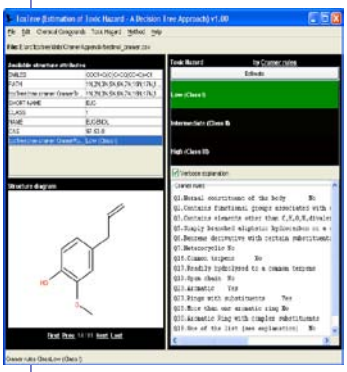


Structure- searchable interface

EUROPEAN COMMISSION
Joint Research Centre

Promoting implementation: Estimation tools

Joint Research Centre

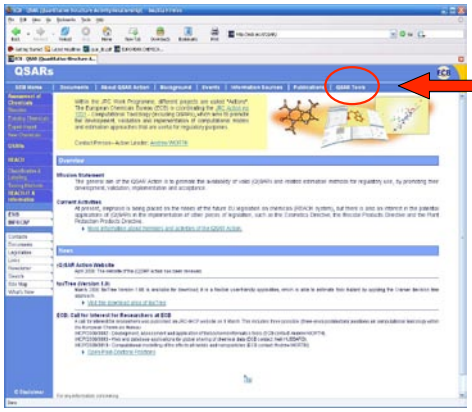


- ToxTree estimates toxic hazard by applying the Cramer classification scheme (33 structural rules)
- Groups chemicals according to structure for Threshold of Toxicological Concern estimation
- Developed by Nina Jeliazkova (Ideaconsult Ltd, Sofia, Bulgaria) (http://ambit.acad.bg) under ECB contract
- Flexible – can be adapted to encode different structural rules
- Freely available from ECB website

EUROPEAN COMMISSION
Joint Research Centre

Where to find ECB QSAR tools: http://ecb.jrc.it/QSAR

Joint Research Centre



EUROPEAN COMMISSION
Joint Research Centre

Capacity building

Joint Research Centre

1. Training on (Q)SARs for regulatory and industry end-users
   1st ECB course: 19-21 October 2005. Sofia, Bulgaria
   2nd ECB course: 24-25 July. Ispra, Italy
2. Training on decision analysis
   ECB / INERIS workshop planned (Nov-Dec 06)
3. Information tools via ECB website
   Danish QSAR database

EUROPEAN COMMISSION
Joint Research Centre

Challenges for the future

Joint Research Centre

1. Need for capacity building (stepping stone to acceptance)
   Training courses, workshops, learning by doing
2. Establishing the basis for REACH-implementation
   Chemical databases and tools for property estimation
   Guidance and criteria for use of (Q)SARs and grouping methods: "Manual of Experience"
3. Research to fill the information gaps
   ITS and its component parts, e.g. new, tailor-made (Q)SARs
   New methods for applicability domain assessment and chemical similarity analysis


EUROPEAN COMMISSION
Joint Research Centre
ECB
The ECB QSAR Team
Arianna Bassan
Chemoinformatics, QSAR tools, computational nanotoxicology
Ana Gallegos Saliner
Chemical similarity, skin irritation
Tatiana Netzeva
Environmental QSAR, consensus modeling, training and enlargement
Grace Patlewicz
Human health QSAR, decision analysis
Manuela Pavan
Ranking methods, environmental QSAR
Ivanka Tsakovska
3D QSAR modeling, eye irritation
Andrew Worth
Human health QSAR, Integrated Testing, regulatory applications
Joint Research Centre
ihcp

EPA/600/R-07/095 | August 2007 | www.epa.gov/ord

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

Office of Research and Development
National Homeland Security Research Center
Cincinnati, OH 45268

Official Business
Penalty for Private Use
$300

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