Summary of NIH Workshop on Predictive Drug Toxicology

The NIH Summit Workshop on Predictive Drug Toxicology took place at the Natcher Conference Center in Bethesda, Maryland, on June 15 to 17, 2004. This workshop was an activity of the Molecular Libraries and Imaging initiative of the NIH Roadmap for Medical Research. Workshop participants discussed existing and developing technologies in the prediction of drug absorption, distribution, metabolism, excretion, and toxicology. They also identified advances that are needed to improve the preclinical testing of drug candidates.

In recent years, the completion of the Human Genome Project and the development of several "-omics" technologies and combinatorial chemical libraries have led to the identification of a number of new potential drug targets. However, there has not been a corresponding increase in the number of new drug classes developed and approved for human use, and there is considerable concern over the reduced number of drug candidates in the pipeline. A major reason for this lack of new therapeutic agents is the failure rate of drug candidates in clinical testing due to unexpected adverse effects and serious toxicities forcing their withdrawal. Compounding the problem is the high cost of developing a new drug, which reportedly ranges between $500 million and $900 million. Clearly, better tests are needed to identify drug-induced toxicities at the preclinical stage and screen out compounds that may fail later when tested in human subjects. It is far preferable for drug candidates "to fail early and fail often" than for compounds to enter testing in human subjects and then fail in late Phase II or Phase III clinical trials.

Workshop participants identified the following tools, technologies, and other needs that would improve the ability to predict toxicity at the preclinical stage:

- Greater understanding of species differences with respect to drug action and drug-induced toxicities to permit prediction of drug effects in humans;
- Transgenic mouse models to understand the function of gene targets in specific organs and to permit specific hypothesis testing for individual gene products in target organs which may help predict how drugs will affect humans;
- Improved primary cell culture procedures to create stable cell lines for drug metabolism and toxicity testing;
- Biomarkers of toxicity based on genomic, proteomic, and metabolomic studies to understand toxicity in humans;
- Improved proteomic platforms to permit the use of mice, dogs, and monkeys to improve predictive toxicity testing in humans;
- Improved tools to link mRNA levels, protein expression levels, protein activities, and metabolite profiles with chemical scaffolds to permit hypothesis development and testing;
- Quantitative *in vitro* assays to identify toxicities in target organs (brain, GI tract, heart, kidney, liver, lung, and other organs);
- Methods to determine the concentration of drugs or drug metabolites in target organs.
- Additional information on drug transport processes and metabolism in target organs, particularly the central nervous system;
- Highly specific inhibitors and specific, identified substrates to permit the characterization of various drug transporters;

- Greater information on reactive metabolite formation by drug-metabolizing enzymes in different species, especially humans, and the effects of these metabolites on cellular processes;
- *In silico* modeling data of drug transport and drug metabolism to understand drug action and drug-induced toxicities;
- Additional structure studies of different drug-metabolizing enzymes to permit more effective three-dimensional QSAR modeling;
- Improved data sets and QSAR modeling tools to improve predictive toxicity modeling;
- Accurate databases of structure-activity relationships that can be easily accessed, used, and updated by the scientific community; and
- Pharmacokinetic-pharmacodynamic models of drug toxicities in humans;

The following pages are summaries of the speakers' presentations and discussion groups

National Institutes of Health (NIH)
Natcher Auditorium

TUESDAY, JUNE 15, 2004

Richard Okita, Ph.D., program director for the National Institute of General Medical Sciences (NIGMS), convened the workshop at 8:05 a.m. The workshop was sponsored by the Molecular Libraries and Imaging Roadmap Initiative, Predictive Absorption, Distribution, Metabolism and Excretion (ADME) and Toxicology section. Dr. Okita thanked the planning committee and scientific organizing committee of this workshop.

Dr. Okita reviewed the upcoming agenda and noted that NIH has been interested in this area for several years. Four years ago the National Cancer Institute (NCI) issued a program announcement (PA) on predictive toxicology and has been soliciting applications. Two years ago NIGMS sponsored a conference that focused on toxicity as a major problem in drug development, and last year NIH Director Elias Zerhouni, M.D., decided that predictive toxicology should be part of the Roadmap Initiatives of the director's office.

Dr. Okita requested that speakers and participants in the workshop share ideas about how NIH can best promote research related to this topic.

Tuesday presentations:

- **Overview Talks**
  - *The Importance of Being Able to Predict Human Sensitivity:* Joseph E. Tomaszewski, Ph.D.

- *Toxicology Liabilities and their Effect on Success Rates in the Pharmaceutical Industry: How Can We Sharpen the Tools that We Have or Force New Ones?:* Richard Robertson, Ph.D.
- *Drug Metabolism/Pharmacokinetics (DMPK)—Friend or Foe of Drug Development?:* Geoffrey Tucker, Ph.D.
- *Development and Application of Computational Toxicology and Informatics Resources at the FDA Center for Drug Evaluation and Research (CDER):* Joseph Contrera, Ph.D.

❑ ***In Vitro* And *In Vivo* Drug Toxicity Models — Presentations**

- *Screening the Receptorome to Discover Molecular Targets Responsible for Drug Toxicities:* Bryan Roth, M.D., Ph.D.
- *Identification of Serotonin 5-HT_{2B} Receptor Cardiovascular Functions, or How Combined Mouse Molecular Genetics and Pharmacology Reinterpret the Toxicology of Fenfluramine (Amphetamine Derived) Anorexigens:* Luc Maroteaux, Ph.D.
- *A Genetic Roadmap for Predictive Toxicology:* Guochun Liao, Ph.D.
- *Genetic and Phenotypic Engineering of Mouse Models of Common Human Disease:* Joseph Nadeau, Ph.D.

❑ **Metabonomics — Presentations**

- *Introduction:* Lois Lehman-McKeeman, Ph.D.
- *Metabonomics And Drug Safety Evaluation:* Donald Robertson, Ph.D.
- *Mass Spectrometry-Based Metabonomics; Data Analysis and Pathway Mapping:* Susan Sumner, Ph.D.
- *Systems Toxicology and the Chemical Effects in Biological Systems Knowledgebase*: Michael Waters, Ph.D.

OVERVIEW TALKS

THE IMPORTANCE OF BEING ABLE TO PREDICT HUMAN SENSITIVITY

Joseph E. Tomaszewski, Ph.D.
National Cancer Institute

In oncology, where many patients are treated with toxic agents, the ability to predict human sensitivity is of paramount importance. With terminal cancer patients as subjects in phase I clinical trials, it is important to have a good paradigm to predict toxicity. This presentation examined time and financial considerations, toxicity considerations, predictability of animal studies, predictability of *in vitro* studies, and a wish list for future research.

The road from drug discovery to approval is long and arduous and can take 15–16 years from discovery to marketing. Of 5,000–10,000 or more compounds that are initially screened, 2,500 enter preclinical testing, five enter clinical testing, and one is ultimately approved by the Food and Drug Administration (FDA). Typically, it costs $800 million to bring a new drug to market. Including anti-infectives, 46 percent of drug development is terminated because of lack of efficacy, 17 percent because of animal toxicity, 16 percent because of adverse reactions in humans, and seven percent for either pharmakinetic, commercial, or miscellaneous reasons.

Serious liver injury is the primary reason for drug removals or restrictions. A recent examination of drugs withdrawn from the market for unexpected adverse events in a 12–month period raised questions about the FDA's drug review process. Some (for example, fenfluramine) had been on the market for years with millions of patient exposures before being withdrawn. Some had relatively few patients exposed to the product during testing, and it took the many post-market cases to detect problems.

Toxicities, such as seen with troglitazone, an oral hypoglycemic agent, and Baycol, a statin, probably could not have been predicted through traditional toxicity studies, although drug/drug interaction studies might have predicted some. Animal tests increase the predictability of human response and show bone marrow and gastrointestinal toxicities, but other toxicities are not that well predicted. The mouse is the poorest predictor of human toxicity; dog and nonhuman primates yield better results. For example, dogs were better than mice and mice better than rats in predicting maximum tolerated dose. However, dogs and nonprimate humans are difficult to get and expensive.

Looking at the value of *in vitro* data, in studies by NCI and others, bone marrow assay results found that mouse data alone accurately predicted human maximum dose for 38 of 48 drugs, and mouse and dog data accurately predicted for 43 of 48 drugs. There is disparity between species in bone marrow data with much more sensitivity seen in humans, but researchers have determined that in general *in vitro* bone marrow assays are highly predictive of human sensitivity.

For future research, scientists should aim not for assays that merely rank new chemical entities or assays with concentrations that are not physiologically relevant, but rather for quantitative *in vitro* assays that are predictive of human sensitivity in relation to animal sensitivity, targeting the brain, GI tract, heart, kidney, liver, and lung. Ideally, *in vivo* and *in vitro* animal and human data

from tissue banks will be available to generate accurate predictions for starting doses, maximum tolerated dose, dose limiting toxicity, and pharmacokinetics.

TOXICOLOGY LIABILITIES AND THEIR EFFECT ON SUCCESS RATES IN THE PHARMACEUTICAL INDUSTRY: HOW CAN WE SHARPEN THE TOOLS THAT WE HAVE OR FORCE NEW ONES?
Richard Robertson, Ph.D.
Bristol Meyer Squibb Company

The industry dilemma is that as research and development (R&D) investment has increased, the number of approvals has declined. The problem is real—while the mid–1980s to the mid–1990s are referred to as the golden age of pharmaceuticals, the situation is now a perilous morass, with companies consuming each other and a resultant decreasing number of companies actually engaged in pharmaceutical research.

While the success rate from submission to market had an upturn in the past ten years, with an approximate 90 percent success rate, the success rates from first pivotal dose have dropped from 70–50 percent. Large declines are also seen in first patient dose to market (30–12 percent) and first human dose to market, which is down to ten percent, a rate that is not sustainable.

Large differences are seen in success rates in the various therapeutic classes. There are much higher rates in the anti-infectives because mechanisms are understood, with surrogate markers from *in vitro* data. But cardiovascular and anti-cancer drugs have much lower success rates, and success for nervous system drugs is worst of all, with an abysmal one percent of nervous system drugs that enter human trials making it to the market. There are many reasons for this, including lack of good animal drugs and the difficulty in designing good clinical protocols. This is an area of huge unmet medical need.

The declining success rates result from a confluence of factors: the demand for larger safety bases; the rapidly escalating cost of clinical development programs with downward pressures on drug pricing; the vastly expanded knowledge of the human genome, which allows identification of targets to outstrip the ability to understand and develop them; and the inability to transfer surrogate markers to clinical use and to find the right target and understand its biology.

Several measures can address the problems. Scientists need to find the right target and understand its biology. Then the task is to find the right molecule—its pharmacology, ADME, toxicologic liabilities, and pharmaceutical properties. It is necessary to develop a system to fail early and often to allow as many shots as possible at the goal—predictive, high throughput, integrated, interactive screening across multiple systems pharmacology, ADME, toxicology, and pharmaceutics. The tools to do this are not yet in place, but it is where the industry needs to go.

Specifically, this means improved target organ-specific knockout technologies in mice and rats, precise examination of the role of biotransformation in target organ drug toxicity, and hypothesis testing of the role of specific gene expression pathways to organ toxicity. It means major improvements in the use of primary cell cultures stable cell lines for predictive toxicology assays, with the ability to maintain cellular phenotypes *in vitro* to allow direct comparison of possible target organ toxicities and mechanisms between laboratory animals and humans. Biology systems development is needed to develop pathway tools for proteomics—tools linking mRNA, protein, and intermediary metabolic changes with chemical scaffolds would allow toxicology hypothesis development and testing. Improved proteomics platforms are also

needed, as are increased understandings of drug absorption processes and effects of various foods. Finally, functional alliances are necessary with chemists, and a persistent, risk-taking approach that takes into consideration all existing ADME, toxicology, pharmacology, and pharmaceutics knowledge. Increased NIH funding is needed to establish core university training programs in whole animal pharmacology and for physician/scientists in clinical pharmacology, with an emphasis on understanding how to apply cutting edge technologies to the clinical setting.

DISCUSSION

In answer to a question about what is meant by pathway tools, Dr. Robertson said integrated systems, computer systems, and software that will allow chemists to look at the effects of a universe of changes, not one change at a time.

Naresh Chand, National Institute on Drug Abuse, asked the opinion of industry on the use of microassays to determine the effect of one compound on millions of genes. Dr. Robertson responded that in most cases, understandings and software tools are lacking, and the tendency is to look at what is already known about a gene. Some private industry is focusing on array work, he added, and some technology has been developed that allows hypothesis testing, but researchers are still in the observational stage and haven't yet moved into an interpretive phase.

In response to a question whether the FDA will eventually replace animal testing with *in vitro* testing, Dr. Robertson declined to speak for the FDA but commented that it would be foolish to replace animal testing because so much remains unknown about the translation of gene pathway into biology that could not be predicted by *in vitro* systems. He added that the key is to not repeat animal studies.

In response to a question about the reason for declining drug approval times, Dr. Robertson said there are multiple reasons, including downward pressure on pricing leading to smaller return on investments with a smaller margin for error. This means a relatively minor imperfection in a drug may prompt a decision to not go forward. Examples include dosage, food interactions, drug/drug interactions, and disappointing work so far to develop neuroscience drugs.

James Halpert, University of Texas, noted a de-emphasis of pharmacology at academic institutions and his concern that future physicians will not be adequately qualified to prescribe drugs, and asked how that can be counteracted. It must be, Dr. Robertson said, adding that the pharmaceutical industry has a declining number of drug representatives and increasing reliance on advertising. The discipline is being given short shrift and NIH needs to target funding to address the problem. Another participant commented that students are more interested in work at molecular and cellular levels, and not drug delivery. Dr. Robertson said that at present we do not fundamentally understand how drugs are absorbed.

DRUG METABOLISM/PHARMACOKINETICS (DMPK)—FRIEND OR FOE OF DRUG DEVELOPMENT?

Geoffrey Tucker, Ph.D.
University of Sheffield

A successful drug must meet a medical need, demonstrate superior safety and efficacy to previous treatments, and have pharmocokinetic and metabolic advantages. In the past ten years, industry has given increased consideration to DMPK, in part because of pressure from academia and regulations. DMPK is about risk management. It has been driven by molecular biology, including *in vitro* technology and human systems; the realization that more than potency/selectivity is needed; and its usefulness to ascertain structural problems and drug interactions.

In the past, DMPK studies were added at the end of drug studies in a descriptive and reactive way. A 1988 study of the reasons for drug withdrawal during development determined that between 1964 and 1985, 90 percent of drug withdrawal was because of pharmacokinetics or related factors. That is now down to seven percent.

Two examples of drugs that slipped through without recognition of major DMPK deficiencies are benoxaprofen and terfenadine. Benoxaprofen (Opren) killed a significant number of patients, with different reactions from healthy and frail subjects, and it marks an important milestone in the development of regulatory requirements for DMPK profiling and the need of special consideration in prescribing drugs for the elderly. Terfenadine (Seldane), a non-sedating antihistamine, proved to be cardiotoxic, particularly in combination with some other drugs, although a metabolite, fexofenadine, is not cardiotoxic. It is one of more than a dozen drugs that have been either withdrawn from the market, terminated from clinical development, refused approval, or been approved only with severe prescribing restrictions because of interactions with other drugs.

The more modern use of DMPK incorporates it into drug design and synthesis in a predictive and proactive way with complete integration, feedback, and feed-forward. Many in industry are working on DMPK, which clearly can prevent drug disasters and contribute to selection of the successful drug. The drug with an ideal DMPK signature is soluble in 250 ml water, has a dose size of less than 500 mg, has complete and passive absorption, bioavailability greater than 20 percent, linear PK, half life of 8–12 hours, balanced clearance, low CYP interactions, and no active or reactive metabolites, P-glycoprotein, 2D6, or acyl glucuronide.

A more negative view of DMPK contends that it represents obsession with process, information overload, filters that are too low, and unbendable rules that could slow drug development and stall some compounds. However, the new DMPK is based on the hope that the development of "omics" beyond genomics will provide greater understandings, as will imaging, biochemical pathways of microdosing, and modeling and simulation. Models can be molecular, physico-chemical, physiologically based pharmocokinetic (PBPK), pharmacokinetic-pharmacodynamic (PKPD), clinical trials simulation, or pharmacoeconomic. Linking to toxicodynamics is more difficult than linking to pharmacodynamics. The need is to increase the ability to integrate information and understand and predict variability—specifically, who is at extreme risk, which is currently not well understood.

The interface between late preclinical and early clinical application can conflict, and variability, again, is a big issue. The system must be stressed using *in silico* plus *in cerebro* modeling to

tease out surprises. But simulation doesn't necessarily equal prediction—it intergrades and synthesizes existing information to take researchers to the next experiment, but may not predict outcome. Dr. Tucker and his colleagues have developed a basic algorithm to develop estimates of kinetics of drugs, but models are not always predictable. The formula begins with characteristics of the patient population—for example, height, weight, age, gender, ethnicity, liver size and blood flow, gut size and blood flow, genetic differences in enzymes and receptors—to ultimately predict estimated ranges of drug clearance. Simulation might also help in difficult decisions regarding drug/drug interactions

Dr. Tucker concluded by emphasizing that the challenge lies in producing the next generation of teachers, who must be the drug metabolism leaders of the future. He underscored the importance of this to industry as well as academia and expressed concern that science education is being "dumbed down."

DEVELOPMENT AND APPLICATION OF COMPUTATIONAL TOXICOLOGY AND INFORMATICS RESOURCES AT THE FDA CENTER FOR DRUG EVALUATION AND RESEARCH (CDER)
Joseph Contrera, Ph.D.
U.S. Food and Drug Administration

The Informatics and Computational Safety Analysis Staff (ICSAS) is a small group within the Center for Drug Evaluation and Research that is responsible for developing toxicology databases. The role of this applied regulatory research unit is to make better use of on-going work and develop databases of animal toxicology and clinical endpoints. It develops rules for quantifying toxicological and clinical adverse effects data and is also involved in evaluating quantitative structure-activity relationships (QSAR) and mining software using ICSAS databases. ICSAS is also working with software developers to develop QSAR prediction programs.

The FDA Critical Path Initiative is a new toolkit developed by government and industry to improve drug development. It recognizes the urgent need for better biomarkers to improve predictability and efficiency. Computation toxicology can identify gaps where toxicology data are limited or lacking, and applications include lead pharmaceutical screening, evaluating contaminants and degradants, decision support information for toxicology issues related to drug products, food contact substances, environmental and industrial chemical toxicity screening, hypothesis generation, and research prioritization. The FDA has proprietary clinical and toxicology data for review and non-proprietary data for approval and post-approval cycles, and applies the data for guidances, decision support, R&D, and to develop computational toxicology models.

Precise nomenclature is needed to unique identify chemical substances in electronic databases and ICSAS decided to use the ".mol" file, the digital chemical structure-based substance inventory, which contains an enormous amount of information. The "chemoinformatic" knowledge base built around this allows comparisons to and searches of clinical and toxicology databases, clinical study and pharmaceutical toxicology study summaries, adverse event reporting systems, and other databases. The knowledge can then be applied for regulatory decision support, to search for similarities among compounds, for example, and provides a powerful tool for regulatory review.

Computational toxicology, then, is the application of computer (*in silico*) technology and information processing to analyze, model, and predict toxicological activity, or "e-tox." The area of predictive toxicology is an emerging one, filling a need to transform data to numerical form, add chemical structure, plus structure-activity relationship software to yield toxicity response and dose predictions.

ICSAS has cooperative agreements to work with several companies to develop predictive toxicology. They have tried to select companies that have products that are unique or algorithms with promise, and also have an agreement with the National Institute on Drug Abuse to screen drugs for carcinogenicity and teratology.

Converting data into a weighted numerical potency scale is an essential step in QSAR modeling. Rodent carcinogenicity was the first target—to convert tumor findings based on properties of a tumor (multiple or single site, weak or equivocal single site), stratify the data, and rate it on an activity scale.

Different programs are based on different principles. MCAST/MC4PC uses the molecular fragment theory, reducing a molecule to all possible 2–10 atom fragments, comparing the fragments to active and inactive fragments in the control database associated and not associated with toxicity, to reach an endpoint that estimates toxic potential. To relate chemical structure to biological activity, MDL software uses unique e-state descriptors, which do not disclose proprietary structures.

The use of computational toxicology has shown promise in work with rodent carcinogenicity. Carcinogenicity testing is required for all new drugs and is costly and time-consuming, taking 3–4 years that could alter approval of a drug. It is possible that computational methods could eventually replace studies for compounds that are highly represented in the carcinogenicity database, freeing resources for testing compounds that are truly new molecular entities poorly represented in the carcinogenicity database.

CDER is working on human modeling, which will be discussed in greater detail later in the workshop. Endpoints include adverse effects in the liver, cardiovascular system, kidney and bladder, and immune system; and dose-related estimates of safe starting dose, maximum recommended therapeutic dose, and no-effect level.

Current software limitations include the inability to deal with inorganics, high molecular weight polymers, very small molecules with few atoms, and compounds that are poorly covered in the training database. Challenges for the regulatory acceptance of *in silico* testing include the development of accurate, validated software; standardization; experience and training; and database and data sharing with adequate protection of proprietary information. Open-minded regulatory scientists and managers willing to consider and use new approaches are needed, along with objective appraisal of the limitations of current testing methods.

Pharma 2005: An Industrial Revolution in R&D (PricewaterhouseCoopers, 1998) looked at the future in R&D, which is now based on primary studies in labs and patients, with secondary science using computers, and concluded that in the future this process could be reversed, with electronic science predictive, and laboratory and clinical science confirmatory.

DISCUSSION OF OVERVIEW TALKS

The following questions were asked and points made:

- In response to a question about the state of development of other model systems, Dr. Tomaszewski said there are obvious limitations to an *in vitro* system. Scientists are working with living liver and lung tissues *in vitro* looking for answers not found in a cellular system, but they need systems that are more physiologically based than current models.
- The FDA has recommended online therapeutic databases, but there have been administrative and technical problems, including proprietary uses.
- Regarding predictability in humans from other species, Dr. Tomaszewski said the size of the mouse precludes some measures; that dogs vomit easily but are no more sensitive than other species and more predictive; that more evaluations can be done in a large animal, but some animals might not have the resistance of others. In general, the dog tends to be a better predictor than the mouse or rat.
- In response to a question about which *in vitro* systems are better, Dr. Tucker said the challenge is to transport relative to the enzyme, and researchers do reasonably well in predicting exposure, but the big challenges are understanding and modeling predictive toxicology.
- Regarding cytotoxicity of oncology drugs, Dr. Tucker said that pharmacokinetics and pharmacodynamics are being used to identify this in newer drugs. In response to another question, he said that many promising compounds have not yet been related to pharmacokinetics. Dr. Robertson added that there are more tools that can be used to predict toxicity, and only about 50 percent of toxicities in man are being predicted, so there is still a long way to go.
- A participant commented that American professional societies are concerned about the few academics entering these fields, and this is in part because of increasing demand from industry for these scientists. Recognizing this, industry is sponsoring more fellowships. Dr. Tucker said that industry in the United Kingdom is training its own people, but an academic presence in the field is desirable.
- In response to a question about dealing with drug interactions in the database, Dr. Contreras said that that is not yet being done, and would be another dimension for another generation of development.
- In response to a question about toxicology tools replacing animal testing, Dr. Contreras said that is not likely in foreseeable testing, but considerable unnecessary and duplicative testing goes on now and the *in silico* approach offers possibilities for streamlining and maximizing efficiency. Dr. Robertson and Dr. Tucker agreed that it probably won't happen soon but it is something to strive for.

IN VITRO AND IN VIVO DRUG TOXICITY MODELS SESSION

SCREENING THE RECEPTOROME TO DISCOVER MOLECULAR TARGETS RESPONSIBLE FOR DRUG TOXICITIES

Bryan Roth, M.D., Ph.D.
Case Western Reserve University Medical School

This work was pioneered by exploring how the psychoactive drug screening program of the National Institute of Mental Health (NIMH) could be used to discover molecular targets for drug toxicity. Two basic methods can be used to avoid toxicities in humans: to find better humans, i.e., tailored approaches that identify single nucleotide polymorphisms (SNPs), which are validated using mouse models; and to find better drugs through comprehensive profiling.

In a recently published study about adverse events to selective serotonin reuptake inhibitors (SSRIs), suicidal ideation associated with paroxetine was found to be induced by a single polymorphism, a truly dramatic genotypic effect that lays the groundwork for this discussion of the receptorome and how it can be screened for molecular targets responsible for drug toxicities. While sequencing the genome has uncovered a large amount of information, little of this has yet translated to useful toxicology, with the function of a large number of molecules still to be elucidated.

The receptorome is the entire complement of receptors in the genome, and the goal is to screen the receptorome to predict drug toxicities in a nonbiased and discovery-based approach. However, this could turn out to be a molecular fishing expedition with no way of knowing what a large-scale screen will turn up. G-protein coupled receptors (GPCRs), for example, are a large family of receptors that have not been focused on as targets for drug toxicity but are likely to represent many known targets for drug therapeutic actions. The approach is straightforward—get clones for human GPCRs, ion channels, transporters and enzymes; devise assays (radioligand binding and function assays); and then profile compounds.

An example is the identification of the molecular target for heart valvulopathy caused by fenfluramine/phentermine, the appetite suppressant known as fen/phen. Fen/phen, the most effective appetite suppressant found for humans, was prescribed for this purpose beginning in 1992, with clinics springing up around this combination of drugs. Initially, few adverse side effects were noted, but in 1997, 24 cases of valvular heart disease related to fen/phen were reported. Similar reports followed and eventually it was withdrawn from the market and led to the single largest pharmaceutical product liability suit ever—$16 billion and counting.

Working back from this known effect, scientists screened for actions at a large number of cloned GPCRs by fenfluramine and other valvulopathic drugs, as well as weight-loss agents that do not cause valve damage. Other serotonergic drugs were also screened. The profile of the receptorome shows peaks and valleys and it was clear that one receptor, 5-HT_{2B}, fulfilled all the receptor requirements. Researchers were surprised to find that the parent compound fenfluramine showed weak activity but norfenfluramine, a principle metabolite, was very potent. It is probably also the metabolite responsible for anorectic properties of the drug. The 5-HT_{2B} receptors were found to be enriched in human cardiac valves.

Currently all approved drugs that hit the 5-HT_{2B} receptors are being screened for valvulopathic properties. Pergolide (Permax), used to treat Parkinson's disease; and dihydroergotamine

(DHE-45), a migraine treatment, may have this property, and Lilly has issued a warning on Permax. It appears that restrictive valvular heart disease may be more common than suspected in Parkinson's patient on Permax—it was not previously detected because no one was looking for it. About 1,000 drugs that bind with 5-HT_{2B} have now been screened and most are not anorectic; others were antihypertensives, nasal decongestants, and anti-arrythmics. In a study of agonist activity at 5-HT_{2B} receptors published in 2000, Dr. Roth and colleagues recommended screening for 5-HT_{2B} and that clinicians consider suspending use of all medications with significant activity at this receptor.

Another adverse effect studied was weight gain associated with antipsychotic drugs, which can be a substantial problem with metabolic side effects (i.e., diabetes, hyperlipidemia). Weight gain has been observed with use of both older and newer antipsychotic drugs, and comparative study and statistical analysis of the complex receptorome profiles has allowed predictions of which drugs have the propensity to cause weight gain. Affinity with three receptors—H1-histamine, $5HT_{2C,}$ and α_{1A}-adreergic—best predicted weight gain.

Receptor affinity also allows toxicity predictions for other drugs. For example, agranulocytosis, potentially fatal decreases in white blood cell counts seen in patients who take clozapine (used to treat intractable schizophrenia), is associated with H4 receptor affinity. An Ephedra action due to selective effects as transporter substrate for norepinephrine transporters (NET) predicts that NET substrates will have unfavorable cardiovascular actions. And an NCI screen of about 2,000 compounds for a role in HERG blockade, which can lead to heart arrhythmia and is known to be caused by a variety of medications in susceptible individuals, found about ten percent "hits." However, structure-based predictions for HERG blockade are not likely to be successful because of the structural diversity of hits.

Dr. Roth recommended that a roadmap for predictive toxicology include a comprehensive receptorome screen on all available drugs, with a focus on drug failures due to unpredicted toxicities. This requires a comprehensive database of drug affinities, efficacies, and side effects and involves no technological barriers, just FDA toxicity data, which is proprietary.

For more detailed information, Dr. Roth directed participants to the NIMH Psychoactive Drug Screening Program K_i database: http://kidb.cwru.edu.

IDENTIFICATION OF SEROTONIN 5-HT$_{2B}$ RECEPTOR CARDIOVASCULAR FUNCTIONS, OR HOW COMBINED MOUSE MOLECULAR GENETICS AND PHARMACOLOGY REINTERPRET THE TOXICOLOGY OF FENFLURAMINE (AMPHETAMINE DERIVED) ANOREXIGENS

Luc Maroteaux, Ph.D.
Strasbourg University

Chronic exposure to the anorexic drug fenfluramine can increase the risk for developing pulmonary hypertension, although this toxicity is somewhat rare; about 50 per 10,000 develop the disease.

The pathogenesis of pulmonary hypertension begins with an overproliferation of endothelial and smooth muscle cells, leading to an irreversible occlusion. Risk factors include female gender, viral infections, use of drugs such as anorexigens and amphetamines, environmental factors, sickle cell anemia, and cardiopathies. Genetic susceptibilities include the BMPRII dominant mutation and platelet storage disease. The serotonin 5-HT$_{2B}$ receptor may represent another susceptibility.

Dr. Maroteaux and colleagues have examined the role of 5-HT$_{2B}$ receptors in chronic fenfluramine exposure and pulmonary hypertension, using hypoxia-induced pulmonary hypertension both in wild-type mice and mice null for 5-HT$_{2B}$. Adult mice lacking the 5-HT$_{2B}$ receptor exhibit left ventricular hypoplasia and dilated cardiomyopathy, with normal blood pressure, heartbeat, aortic and mitral flow, and pulmonary artery flow; and they are resistant to hypoxia-induced pulmonary hypertension and chronic fenfluramine exposure. In both wild-type and null mice, more than 99 percent of fenfluramine is converted to nor-dexfenfluramine, an agonist for serotonin receptors, and it is this metabolite that mediates the chronic effects of fenfluramine. In addition, the 5-HT$_{2B}$ receptor is required for hypoxia-induced vascular proliferation, and 5-HT$_{2B}$ expression is increased in pulmonary hypertension. A similar increase has been observed in a pilot study of human patients with hypertension.

Hypoxia-dependent increases in plasma serotonin levels depend on 5-HT$_{2B}$ receptor activity. Plasma serotonin concentration correlates with right ventricular systolic pressure (RVSP) values and 5-HT$_{2B}$ receptor sites, but not with whole blood serotonin, suggesting that plasma serotonin levels are regulated independently of whole-blood serotonin. In addition, hypoxia-induced modifications in lung serotonin-uptake activity require 5-HT$_{2B}$ receptors. Yet acute-injection experiments with fenfluramine or BW723C86, a high-affinity agonist of the 5-HT$_{2B}$ receptor, with or without treatment with paroxetine, suggest that fenfluramine regulation of plasma serotonin levels is independent of 5-HT$_{2B}$ receptor activity and partially paroxetine sensitive.

On the basis of these studies, Maroteaux and colleagues conclude that 5-HT$_{2B}$ receptors are required for chronic, hypoxia-dependent fenfluramine effects; that the chronic effects of fenfluramine are mediated by nor-dexfenfluramine; and that nor-dexfenfluramine uses 5-HT$_{2B}$ receptors to regulate serotonin levels. More work is needed to further explore 5-HT$_{2B}$ function in chronic fenfluramine exposure and pulmonary hypertension, including bioavailability in acute and chronic situations, pharmacology and coupling of all metabolites, and established animal models of suspected diseases. In addition, multifactor diseases must be considered.

A GENETIC ROADMAP FOR PREDICTIVE TOXICOLOGY
Guochun Liao, Ph.D.
Roche Palo Alto

Genetics has an impact on all stages of research and development, particularly the identification of better targets, improved efficiency in pre-clinical and clinical development, and lead optimization and improved product usage. Dr. Liao and colleagues have used a genotoxicity assay to identify alterations in gene expression that occur in response to treatment with various compounds.

The genotoxicity assay relies on the differentiation arrest caused by genotoxic stress. In a model published by Puri and colleagues in *Nature Medicine*, C2C12 cells undergo differentiation arrest when starved and resume differentiation when shifted to the appropriate medium. Pretreating cells with a genotoxic compound before shifting them to differentiation medium prevents the cells from resuming differentiation upon a shift to favorable conditions. Dr. Liao and colleagues have pretreated C2C12 cells with six different compounds, five of which cause DNA damage, and used microarrays to examine alterations in gene expression. They have identified 86 genes that were differentially expressed; 60 of these genes have annotated functions. Of those, 26 are regulated by p53 or c-myc, and 32 are involved in the DNA damage response. Genotoxic agents can therefore be detected by studying a small subset of genes.

Computational tools also can accelerate the pace of genetic discovery. Traditional methods, which involve the examination of inbred mice, are costly and time-intensive. Roche Palo Alto has an NIH-funded mouse SNP database, http://mouseSNP.roche.com, which comprises more than 110,000 SNPs, 21 strains, and 1,700 genes. Dr. Liao and colleagues used computation mapping of haplotypic patterns among strains to identify genes involved in responses to toxicity, and they used inbred mice to verify their results. This method has been used to identify genes involved in aryl hydrocarbon response and to identify the functional elements within the H2-E□ gene.

The method has two limitations thus far: the gene must be in the SNP database for investigators to examine phenotypic differences, and the database has been purposefully constructed to contain genes known to be involved in drug metabolism. Even so, computational genetics provides a powerful platform for studying drug metabolism. Genomics thus promise to improve genetically based therapeutic target identification, accelerate the pace of drug discovery, and further elucidate the mechanism of drug metabolism.

GENETIC AND PHENOTYPIC ENGINEERING OF MOUSE MODELS OF COMMON HUMAN DISEASE

Joseph Nadeau, Ph.D.
Case Western Reserve University School of Medicine

A number of mouse strains have been developed in the past few years that have tremendous potential for productive use in predictive toxicology. Complex traits such as obesity, hypercholesterolemia, and a host of others are influenced by multiple genes and environmental factors. Genetic heterogeneity leads to phenotypic heterogeneity, and strategies and tools are needed to dissect genetic control and engineer phenotypic variants in a defined, controlled, and reproducible manner. Chromosome substitutions strains (CSSs), known as consomic strains, may provide these tools.

In a realistic model, multiple genes lead to multiple phenotypes, and it is easy to lose the connection between gene and phenotype. Inbred strains provide a means to disentangle these relations, but despite phenotypic similarities, mice are not humans, in health or disease. The problems are to detect the genes and dissect the phenotype. Compared to traditional genetics, CSSs offer increased power to detect more genes and phenotypes. CSSs are a simple and powerful way to identify quantitative trait loci (QTL) affecting a variety of processes.

To make the CSSs, offspring that inherit the chromosome of interest are identified and crossed back for 10–15 generations, creating a new inbred strain with the chromosome of interest. Use of the CSSs is relatively simple. Advantages of CSSs, which are commercially available from Jackson Laboratories, are strong statistical power, linkage crosses readily made for fine mapping, congenic panels that can be made with fewer generations, functional studies have been done, and they are a permanent, replaceable, living resource.

These strains have been used to detect a number of genes associated with metabolic syndrome, and will provide a detailed understanding of the genetic and functional nature of metabolic syndrome. The syndrome metabolism studies pinpointed obesity resistance to chromosome 6^{A} and smaller adipocytes and increased liver triglycerides. Further phenotype dissection work is being done on feeding behavior, exercise physiology, metabolic tracer studies, imaging studies, calorimetry studies, blood pressure analysis, metabolic pathway analysis, activity studies, telemetry studies, metabolism mapping studies, energetics, and gene expression analysis.

Dr. Nadeau emphasized that CSS represents an important new strategy and offers tools to find genes and engineer mice with defined reproducible phenotypes. The CSSs also could be useful for detecting mechanisms of drug action and drug toxicity.

DISCUSSION

Dr. Roth asked if anyone has done a strain survey for hepatotoxicity, one of most serious toxicities. Dr. Nadeau said he was not aware of such work, but almost everyone who studies a phenotype finds something. Dr. Roth added that his suspicion is that no one has looked.

Dr. Chand (NIDA) asked if anyone else is using this model in industry or NIH work. Dr. Nadeau said that the information has been shared with everyone who asked; 30 or 40 different labs asked for and got the mice. To the best of his knowledge, no one from industry or NIH has

asked for them, but they are available. Dr. Donald Robertson (Pfizer) noted that as a member of industry, he is very interested in these mice. Dr. Lois Lehman-McKeeman (Bristol-Meyers Squibb) said that to a toxicologist this is a very powerful tool with potential use in the pulmonary adenomal sensitivity locus. Dr. Nadeau added that the strategy allows implementing genetics and biology at same time.

METABONOMICS SESSION

INTRODUCTION

Lois Lehman-McKeeman, Ph.D.
Bristol-Meyers Squibb

Dr. Lehman-McKeeman presented a brief overview on metabonomics and systems biology in mechanistic and predictive toxicology. Metabonomics affords the ability to measure relative concentrations of endogenous metabolites in biological fluids (urine, plasma, cerebrospinal fluid, tears, synovial fluid, tears, saliva, bile) and tissues. The principle tool is nuclear magnetic resonance (NMR), but mass spectrometry and any other techniques that detect metabolites are also used. NMR readily and rapidly detects all molecules that contain protons. It is an unbiased measure of all constituents with low but usually adequate sensitivity, and generates complex datasets.

NMR detects intermediary metabolites—these include biofluids (citrate, succinate, hippurate, creatinine, glucose, taurine, amino acids) and tissues (choline/phosphocholine, glycogen, glucose, fatty acids).

Metabonomics offers a number of advantages: the potential to measure most diagnostic changes early; nondestructive sampling, making it easy to design experiments to gather crucial time-course data to follow a course from dosing to insult to regeneration; diagnostic metabolites are often a part of natural physiology; and no new technology is needed. In trying to relate an NMR spectrum to toxicity, the attempt is to use applications in a way to predict a toxic outcome or to better define and understand the underlying cause of a toxic outcome.

METABONOMICS AND DRUG SAFETY EVALUATION

Donald Robertson, Ph.D.
Pfizer Global Research and Development

Dr. Robertson, a traditionally trained toxicologist, and therefore, he said, a trained cynic, noted that metabonomics has always challenged him to rethink what he thought he knew. He relearned some principles of toxicology while conducting metabonomic studies:

- Systems biology: Rats are more than collections of isolated target organs.
- Biologic variation: No two rats are the same.
- Temporal variation: The same rat is a different rat the next day. This can be measured with new technologies and no longer ignored.
- Dynamic toxicity: Rats don't get the concept of weekends.
- Effects don't restrict themselves to drug actions.
- For every toxic action there is a not necessarily equal and not necessarily opposite reaction that is not in any biochemistry book.
- Traditional safety studies frequently ignore most if not all of the above.

Metabonomics demands that the toxicologist think beyond results in terms of the empirically contrived dose selection and sample limitations imposed by pragmatic concerns. The bottom line is economics—the "omics" that matters most in the pharmaceutical industry. Metabonomics offers a list of advantages for pharmaceutical R&D: it is targeted, it has the potential to identify novel biomarkers, it is non-invasive, no *a priori* decisions about samples are needed, little sample processing is needed, it minimizes compound requirements, it offers relatively fast analysis, it identifies outliers pretest, it offers the potential for metabolic information, and complete time course data can readily be obtained.

From pharmaceutical applications, reasons for variations in response are not always that clear-cut. Trying to comprehend a dynamic response using fixed time points is a difficult proposition, and what is often attributed to inter-animal variability may actually be due to temporal variations. A relatively small difference in age can make a big difference. Protocol details that scientists traditionally have paid little attention to can have a big impact. Everything that is going on with the animals must be attended to, and contaminations can occur and clear quickly, escaping detection.

In a real-world application, metabonomics helped explain the case of two MEK inhibitors, one that was efficacious but toxic, the other not efficacious and nontoxic. A clear metabolic difference between the two similar compounds was evident, and time-dependent changes in an unknown drug metabolite were seen in the toxic compound. The ability to separate the active from inactive compound allows toxicity screening on a multivariate basis, a new approach for toxicologists.

The work indicates that mice are as suitable as rats for metabonomic studies, meaning minimal drug requirements. An extra benefit is that unknown and critical metabolic information was obtained as byproduct of the study. Despite the presence of confounding metabolites, pattern separations were still evident as early as 24 hours and were biphasic, suggesting at least two biochemical events.

In another example, metabonomics were added to early discovery. Compound X was in an early phase of development, when unexpected and unexplained deaths occurred in a pharmacology study with many compounds from same class as Compound X. All mid- and high-dose animals were found dead or dying by Day 3 and a single low dose female died on Day 6, with no clinical pathology or histopathology and no apparent cause of death. Metabonomic results showed a profound increase in urine glucose excretion, β-hydroxybutyrate, and creatine. The low-dose animals also showed abnormal analyte levels, but returned to normal by Day 7, despite continued dosing. Metabonomics were able to identify a mechanism of toxicity not evident with standard toxicity endpoints. The effects found, with subsequent regression, have significant mechanistic ramifications regarding the mechanism of action of the compound and may offer a relatively simple way to monitor potential toxicity in further preclinical and clinical development.

These experiments demonstrate a dramatic increase in the breadth and depth of metabonomic applications. Biomarker identification, safety screening, and mechanistic applications are all being explored vigorously, and the focus is shifting from observation to understanding. Metabonomics will serve as a powerful adjunct to other "omic" technologies, but it requires a comprehensive awareness and understanding of all aspects of a toxicity study.

MASS SPECTROMETRY-BASED METABONOMICS; DATA ANALYSIS AND PATHWAY MAPPING
Susan Sumner, Ph.D.
Paradigm Genetics, Inc.

Dr. Sumner moved from a background in nuclear magnetic resonance (NMR) spectrometry to toxicology because she became intrigued by the many signals that were found that were unexpected. Paradigm Genetics has been a pioneer in biochemical profiling using mass spectrometry and in the development of machine vision tools for quantitative tissue phenotyping. Its current focus is on biomarker and biomarker-enabled drug discovery.

This process is based on two core concepts: that biomarkers are needed to monitor normal and pathogenic process and to predict patient response, and that understanding disease and drug action requires the unbiased analysis and interpretation of multiple data streams. Linking gene biomarkers can help scientists learn pathways and drug actions on disease, leading to further understandings of diagnostics and drug targets. Combining data streams provides a systems biology context for discovery, new information that otherwise would have been hidden, and improved knowledge about mechanisms to understand toxicities.

Biomarkers are better defined in metabonomics than in genomics or proteomics, with peaks assignable to specific compounds within defined biochemical pathways and linkage to potential causal mechanisms. Using liquid and gas chromatography/mass spectrometry, a broad spectrum of metabolites can be detected and viewed in a number of ways. Signals from metabolites can be assigned to specific compounds. With pathway mapping, other data streams can be brought in, for example, gene expression analysis. Once proteins and metabolites have been mapped to the same pathways, biochemical hot spots may become apparent.

Automated histopathology uses machine vision to differentiate normal from abnormal tissue and export quantitative data for subsequent analysis. Quantitative histomorphometry removes subjectivity in determination of phenotype, which remains a critical element in defining disease and damage. Objective tissue data provide a phenotypic anchor for other data streams.

Metabonomics, then, provides analytic data and a data management system, both targeted, as well as global profiling methods for analysis, data interpretation and analysis, and data coherence tools to create algorithms for mapping data to pathways and integrating data streams. The biomarker discovery process centers on the metabonome dictionary, which has information about all known metabolites, their genes, proteins, and pathways. This allows for the development of mechanism based-markers.

An example is seen in a liver toxicity study. The liver is a model system because it plays a central role in diverse metabolic processes and has high medical/pharmaceutical relevance because of drug induced hepatotoxicity, a key regulatory issue. An acetaminophen study sought to develop markers for *in vitro* systems such as hepatocytes and *in vivo* systems such as rats, and then human, using human volunteers who received high clinical acetaminophen doses; and then integrate data from other liver studies. More changes were seen in individual metabolites in higher dose groups.

A tool called Nodeworker analyzed what metabolites have in common and allowed visualization of metabolic pathways, helping researchers converge on important metabolites, and paving the way for work that relates gene changes back to specific mechanisms. Combining multiple data

streams can uncover previously unknown data. Scientists are also trying to develop means of bridging biomarkers, to find what biomarkers are available in serum and urine that are reflective of disease state in tissue. Accessible mechanism-based biomarkers can be developed to provide convenient indications of drug safety and efficacy.

DISCUSSION

In response to a question about the role of these tools, Dr. Sumner said they can be used in *in vitro*, preclinical, and clinical phases. Pharmaceutical companies are primarily exploring preclinical uses.

Another participant asked about making correlates between metabolites in urine and tissue, and Dr. Sumner said that urine, tissue, and serum are sampled and used for comparisons. A study design with dose response to develop biomarkers to assess trend of data is needed, she said. Dr. Lehman-McKeeman added that there are more tissue samples available than time to sample them, but there is clear interest in following that up, and the work is likely to be forthcoming.

SYSTEMS TOXICOLOGY AND THE CHEMICAL EFFECTS IN BIOLOGICAL SYSTEMS KNOWLEDGEBASE
Michael Waters, Ph.D.
National Institute of Environmental Health Sciences (NIEHS)

NIEHS's National Center for Toxicogenomics (NCT) has been involved in toxicology work for some time, and the Chemical Effects in Biological Systems (CEBS) knowledgebase is being developed as a public resource. Toxicogenomics is defined as the study of the response of a genome to environmental stressors and toxicants, and it combines genetics, transcriptomics (genomic-scale mRNA expression), proteomics, metabonomics and bioinformatics with conventional toxicology in an effort to understand the role of gene-environment interactions in disease.

The knowledge base uses data and information to carry out tasks that create new information and new understandings, and CEBS aims to be a dynamic system for integrating large volumes of disparate information in a framework that serves as a continually changing heuristic engine. The goal is for CEBS to evolve in content and capabilities to become a system of predictive toxicology.

Just as systems biology provides an overarching description and understanding of how components of a biological system work together and measure changes globally, systems toxicology is a complete description of the toxicological interactions within a system, perturbing a system and measuring changes to develop a better model of the system. The objective of toxicogenomics is to compare toxicogenomic effects of chemicals or stressors across species, yielding signatures of altered gene and protein expression and providing a phenotypical anchor of these changes with conventional toxicology data. Global changes are then delineated as adaptive, pharmacologic, or toxic outcomes. Data integration is the critical challenge.

The vision for CEBS is to assemble high quality internally coherent molecular expression, toxicology, and pathology datasets for environmental chemicals and drugs; use international database standards to fully annotate these datasets; develop a knowledge base by actively assimilating and refining information from multiple public databases and from current scientific literature on an active knowledge template; and to be able to query the knowledge base from sequence alone, to discover new information about genes, gene groups, pathways, toxicity, and disease. Two hallmarks of CEBS are sequence anchoring, in which probes are anchored in genomic sequence; and phenotypic anchoring, using controlled vocabulary. If the genotype is understood, researchers can then think in terms of susceptibility.

Building a knowledge base presents bioinformatics and interpretive challenges and opportunities. Starting with a chemical treatment, and proceeding through effects on multiple genes and proteins, functional characterization, sequential events, and network and systems behavior allows adverse effects to be related back to the chemical treatment. Focusing on the literature, gene/protein, toxicology, and pathology ontologies can be compared to experimental observations to set up a loop to verify phenotypes and sequences and refine gene group membership.

The immediate objectives are to establish a database to capture, store, and analyze gene expression data produced from toxicogenomic experiments in different laboratories; interrogate gene expression data using queries from genomic, experimental and toxicological domains; and gain knowledge of relationships between gene expression changes and toxicological endpoints.

The main challenge is to provide internally consistent data and allow comparability among many toxicogenomic experiments. Standards are essential.

NCT has worked on this project with the National Toxicology Program (NTP) and international partners including the European Bioinformatics Institute (EBI) and the International Life Sciences Institute Health and Environmental Sciences Institute (ILSI-HESI). The partnership grew out of efforts to support the Microarray Gene Expression Database (MGED) project and the MAGE object model. The database is not public yet, but came online to collect data in late 2003.

To uphold standards, scientists are looking at chip technology to represent mouse and rats in terms of toxicogenomic studies, with a strategy of sequence alignment and results indexed according to their probe. The standard is the Cancer Bioinformatics Infrastructure Objects (caBIO). The CEBS System Biology Object Model is publicly available at http://cebs.niehs.nih.gov, and a major goal is to have a toxicology arm this year. These are steps toward public data exchange.

CEBS is now gaining content from five academic members of the Toxicogenomics Research Consortium, and other intramural and extramural partnerships. Guidelines for content include an array with well-characterized probe sequences; proteomics spot data and peptide sequence; description of experimental design; materials and methods information; measure of sample integrity; relevant measure(s) of treatment outcome; and overall data documentation. NCT collaboration with the Environmental Protection Agency is now beginning.

In the future, scientists aim to use computational toxicology for integration of molecular databases, interfacing with a large number of resources. Toxicogenomics can change the way toxicology is performed; it will contribute new methods, data, and interpretation to environmental toxicology; and CEBS will be a key component in toxicological interpretation, linking transcriptomics, proteomics, metabonomics, and toxicology to generate new knowledge.

Dr. Okita adjourned the session at 5:41 p.m.

National Institutes of Health (NIH)
Natcher Auditorium

WEDNESDAY, JUNE 16, 2004

Dr. Okita opened the session at 8:15 a.m.

Wednesday presentations:

- **Absorption, Distribution, Metabolism, Excretion (ADME) —Presentations**
  - *Use of X-Ray Crystal Structures to Predict Mammalian Cytochrome P450 Function:* James Halpert, Ph.D.
  - *Physiological Models Of Drug-Drug Interactions:* Stephen D. Hall, Ph.D.
  - *Structure-Toxicity Relationships as a Tool to Investigate Mechanisms of Drug Toxicity:* Sidney Nelson, Ph.D.
  - *Pharmacogenetics of Phase II Drug Metabolizing Enzymes:* Rebecca Blanchard, Ph.D.
  - *Renal Transporters in Predictive ADME:* Kathleen Giacomini, Ph.D.
  - *Impact of Active Transport on Tissue Distribution of Drugs:* Noa Zerangue, Ph.D.

- **Toxicogenomics - Presentation**
  - *Toxicogenomics: Application of Gene Expression in Toxicology Screening, Mechanistic Elucidation, and Biomarker Identification:* Donna Mendrick, Ph.D.

- **Databases/Data Analysis — Presentations**
  - *Data Analysis Using Structured Biological Knowledge: Experience with a Database of Protein Functional Interactions:* Alan Ruttenberg
  - *Construction of a Human Adverse Effects Database for Modeling Quantitative Structure-Activity Relationships (QSAR):* Naomi Kruhlak, Ph.D.

ABSORPTION, DISTRIBUTION, METABOLISM, EXCRETION (ADME) SESSION

Steven Wrighton, Ph.D., Eli Lilly and Co., moderated this panel and introduced the speakers.

USE OF X-RAY CRYSTAL STRUCTURES TO PREDICT MAMMALIAN CYTOCHROME P450 FUNCTION

James Halpert, Ph.D.
University of Texas—Medical Branch

The first X-ray crystal structure of P450 was recorded in the mid–1980s, and there is currently a great deal of interest in this field and how it can be used to predict pharmacokinetics of new chemical enterprises, looking at the mammalian structures themselves. Four mammalian cytochromes were found in the past year.

Researchers would like to use this technology to predict the pharmacokinetics of new chemical enterprises (NCE). This includes, at an overarching level, pharmacokinetics of the NCE, drug interactions, functional consequences of single nucleotide polymorphisms (SNPs) in individuals and populations, and how to engineer around a metabolic problem with an NCE. At a more detailed level, they want to be able to predict if a substance will be a substrate, a non-substrate, or an inhibitor; sites of metabolism; and time of turnover.

In general, cytochrome P450 catalysis allows substrate entry into P450 and binding, presumably near the site of the heme ligand. X-ray structures of the same P450 with different substrates suggest a very flexible active site topography, and some substrates can bind in more than one orientation. The P450 must assume an open conformation for the substrate to move into the active site, and there has been some controversy about how much free rotation substrates have. It remains an open question.

Looking at rates of binding, four approaches for predicting P450 function are suggested: X-ray crystal structures; pharmacophores (3-dimensional representations); homology models; or combined protein and pharmacophore models, for enhanced accuracy. Work using X-ray crystal structure of bacterial P450 is not all completely transferable to mammalian structures; 5 of 11 predicted substrates yielded organic products, but the accuracy of predictions could be increased by increasing the distance allowed between atom of substrate and atom of enzyme. Homology and pharmacophore and combined models are a fruitful approach that industry can take.

Issues that remain outstanding include substrate access, conformational flexibility, protein-protein interactions, protein-membrane interactions, and atypical kinetics, an issue of great interest to the pharmaceutical business. Dr. Halpert and colleagues tried a basic structural approach to crystallography of P450 2B enzymes and got crystals up to 1 mm in length. The 2B4 structure is very similar to 2C5-DMZ with one exception, an opening that would allow substrates to enter the enzyme.

Researchers looked at active site residues and redox partner binding to determine if the open conformation can turn over substrate and concluded that conformational shift may accommodate sequential substrate binding and catalysis. This has great ramifications for the homology model and may address the longstanding question about how a compound gets into a site and how the enzyme responds once a compound gets in. Most likely, an open substrate

allows substrate binding, then clefts close for selective metabolism; another way of looking at it is that the enzyme is opening and closing all the time. Different enzymes have different mechanisms. It is also clear that protein-protein interactions allow for conformational changes.

CYP3A4 presents a number of complications that make predictions of toxicity difficult, even with existing understandings of P450s. These include sigmoidal steady-state kinetics, stimulation by some substrates of oxidation of others, and partial or no inhibition in the presence of two substrates. Several working models have been proposed to account for atypical kinetics of CYP3A4 reactions, the spatial relationship between the effector and substrate oxidation sites. Two possible scenarios are suggested: 1) a given compound binds at a distinct preferred location within the binding pocket of CYP3A4, in which case interactions between two ligands can be predicted based on knowledge of where they bind; and 2) the binding pocket is plastic, and the precise way in which a compound binds is influenced by a second compound, in which case interactions between two ligands would be very difficult to predict.

Many gaps in understanding remain, and NIH funding should support X-ray studies of P450s in a variety of environments, computational methods that incorporate protein folding algorithms with docking to predict induced fit, and solution structure methodologies.

DISCUSSION

In response to a question about computational resources available for this work, Dr. Halpert said that more power is needed than is currently available, and it is not clear that commercial software packages are amenable to the process. Another participant asked about global approaches, and if the study of perturbation of a system by a toxin would be helpful. Dr. Halpert said that is an interesting point, and there is certainly evidence that one P450 that binds a compound influences another.

PHYSIOLOGICAL MODELS OF DRUG-DRUG INTERACTIONS
Stephen D. Hall, Ph.D.
Indiana University School of Medicine

Dr. Hall discussed whether metabolism based drug-drug interactions are a problem in drug development, why quantitative prediction of the extent of drug interactions is poor, and how predictive models of drug interactions can be improved.

From 1995–1999, there was an increase in new molecular entities (NMEs) withdrawn from the market for safety reasons. While the actual numbers are small, this puts a huge burden on the health care system. Recent drugs withdrawn—terfenadine, astemizole, cisapride, mibefradil—all rely on CYP3A substrates/inhibitors and demonstrate drug/drug interactions. Other drugs are refused approval and never get to the market because of drug/drug interactions.

Adverse drug reactions are one of leading causes of death in the United States. Type A adverse drug reactions, which are responsible for 80 percent of the total, are dose-dependent, often predictable, often due to excessive expression of known pharmacologic effects, and often caused by drug-drug interactions. Most drug-drug interactions are metabolism based and most metabolism is CYP dependent. An extensive list of drugs used to treat a variety of conditions involve metabolism by P450 enzymes.

CYP3A4 inhibitors are ranked as potent, moderate, or weak; it is likely that a potent ranking would require a black box label. Drugs that demonstrate a substantial group of drug/drug interactions have lost considerable market share when a competitor with fewer interactions was introduced. Cimetidine is one example.

To predict interactions, scientists are looking at enzyme induction, when elevated expression and decreased clearance and bioavailability of drug is seen. The amount of drug is a balance of the rate of synthesis and rate of degradation, but there is virtually no data about predicting *in vivo* from *in vitro* data, despite devastating consequences of some of these drugs. Complicating the issue is variability, illustrated by the effect of rifampicin on hepatic and intestinal wall CYP3A activity. This is a challenge that will require attention in the future, because many new drugs have the ability to induce or inhibit 3A.

In vitro competitive inhibitor potency is determined routinely for drugs under development, and there have been some successes using this data to predict actions *in vivo*—for example the clearance of warfarin. But others exhibit large discrepancies, such as the interaction of simvastatin and diltiazen, where *in vitro* data would predict no interaction, but interaction is seen clinically. A third example of a false negative from *in vitro* data is midazolam following clarithromycin.

Reasons for poor predictions from *in vitro* include contributions of drug transporters, extra-hepatic metabolism, a possible role of metabolites as potent inhibitors, incorrect inhibition models, and failure to account for multiple pathways of elimination. These all can be characterized for specific drugs, but quality information is rarely available in early stages of drug development, and unknown even now for some drugs in wide use. A simple model using time averaged concentrations of inhibitor in systemic blood cannot account for concentration and time-dependent changes in substrate, inhibitor, and enzyme concentration. The solution is to use physiologically based pharmacokinetic models. Another barrier is inappropriate estimation

of inhibitor concentration; possible solutions include cell-based estimation of effective Ki, estimates of effective *in vivo* Ki values from animal models and humans.

In a physiological model of midazolam disposition, hepatic metabolism can be characterized by measuring what goes in the liver and what comes out. In comparisons of the effect of ketoconazole on midazolam in rat and human studies, the comparison is very good, and scaling up from a rat to human model is a valid approach.

In predicting drug interactions, population variation is the key to the worst case. The success of drug/drug interaction prediction is currently evaluated at an average level. However, variation of average area under the drug concentration curve ratio (AUCR) before and after interaction in the population of interest is as important as the mean, because it represents range of interaction and worst-case scenario. Pharmacokinetic parameters are uncertain, and this needs to be considered when predicting the AUCR.

A strategy for building a physiologically based pharmacokinetic (PBPK) model is to: 1) establish bounds for PBPK parameters from animal, *in vitro*, and *in vivo* data, and use them as prior knowledge; 2) use published individual drug concentration data to update PK parameters and their population variations; and 3) verify and select the best drug interaction model if the clinical drug interaction data is available.

Using Bayesian methods, incorporate all variation sources, update PK parameter estimates for a complicated PBPK model from simple models, predict false negative rate by simulation, and perform model selection. With these parameters determined, how bad interactions can be in a population can be predicted. The biggest stumbling block remaining is determining tissue inhibitor concentrations.

DISCUSSION

Dr. Tucker observed that it seems that research in the United Kingdom is a bit ahead of the U.S. on this subject.

STRUCTURE-TOXICITY RELATIONSHIPS AS A TOOL TO INVESTIGATE MECHANISMS OF DRUG TOXICITY

Sidney Nelson, Ph.D.
University of Washington School of Pharmacy

The structure of drugs mediates therapeutic activity as well as toxicity. Important determinants of toxicity are chemical structure, dose, rates of toxic metabolite formation versus detoxification, immune recognition of macromolecular adducts of a drug or its metabolites, and efficiency of macromolecular repair mechanisms.

Investigation points in drug induced liver injury—*in vitro* and in humans and animals—include susceptibility factors (initiating and progression mechanisms) and adaptive and protective factors. These are variable; the example of acetaminophen illustrates that at certain high doses, everyone who takes it will end up with some liver damage, and some will experience toxicity at lower doses.

In general, metabolite-mediated drug toxicities, like oxidation, can lead to intrinsically toxic reactions, immune reactions, or idiosyncratic reactions, or the macromolecular adducts can be excreted and/or repaired. Often drugs that are withdrawn or given black box warnings cause liver problems, and reactive metabolites are the culprit.

Drugs that contain or can be metabolized to the following structures may cause toxic effects via formation of reactive metabolites:

- ❑ Arylacetic or aryl propionic acids
- ❑ Thiophene, furan, or pyrrole
- ❑ Anilines or anilides
- ❑ Quinone and quinone imines
- ❑ Medium chain fatty acids
- ❑ Halogenated hydrocarbons and some halogenated aromatics
- ❑ Nitroaromatics
- ❑ Groups that can be oxidized to acroleins or form reactive allylic alcohols
- ❑ Thiols, thiono compounds, thiazolidinediones

But structure analysis can be misleading, and more than one part of a structure may be involved. In some cases (i.e., atenolol) switching the position of the isomer removes the potential for toxicity.

In the case of acetaminophen, other related structures have been studied but each has a slightly different synthesis. With *N*-acetyl-*m*-aminophenol (AMAP), for example, most of the binding is to microsomal proteins, while acetaminophen binds more to mitochondrial proteins and is a mitochondrial toxic.

Work is also being done with genomic responses to acetaminophen in stress-responsive changes, cell cycling and growth inhibition, adhesion and structural components, inflammatory changes, cell signaling and death, and cell metabolism. Some mouse strains (i.e., B6) are more susceptible than others (i.e., SJL), though little difference is seen in the metabolism of acetaminophen. The differences are in heat shock proteins, which are expressed much less in the B6 strain.

Proteomics analysis takes a global approach to examine differences in proteins in two strains. Several proteins of interest are more abundant in SJL mice; some are mitochondrial, some are not.

Looking at nitroaromatics, a common group, two mechanisms may form metabolites and oxidated stress leads to damage. In the case of androgen receptor antagonists, flutamide was recently removed from the market because of hepatotoxicity. Nilutamide was also removed, but bicalutamide remains available, with few adverse drug events (ADEs) observed. Troglitazone, one of the thiazolidinediones, was removed from the market because of liver toxicity, while a much lower incidence of ADEs was associated with rosiglitazone and the profile for pioglitazone is not yet clear. In the case of quinolone antibiotics, temafloxacin was removed from the market in 1991 because of ADEs including hemolytic anemia, central nervous system problems, and hepatotoxicity; trovofloxacin reduced to restricted use because of hepatotoxicity; and enoxacin has demonstrated phototoxicity and ophthalmologic problems.

Further studies of many of these compounds are needed to determine mechanisms, and dose differences and potency need to be explored. Structural differences can be used as tools to probe toxicity. Needs for researchers include:

- Development of accurate database of structure-toxicity relationships, with input from FDA, pharmaceutical companies, and working groups to evaluate case reports.
- Development of better in vivo and in vitro models of idiosyncratic drug-induced liver disease (DILD) to probe with toxicants versus analogues that are less toxic or intrinsically toxic.
- Use of different kinds of toxicants to identify biomarkers of different kinds of toxicity.
- Use of toxicokinetic and toxicogenomic data to retrospectively evaluate the potential for toxicity of marketed and prospective drugs.

PHARMACOGENETICS OF PHASE II DRUG METABOLIZING ENZYMES

Rebecca Blanchard, Ph.D.
Fox Chase Cancer Center

The classic method for therapeutic drug monitoring looks at plasma drug concentration over time to determine ADME and endpoints that are toxic, therapeutic, or ineffective. Much is now known about the genetics that go into ADME processes, and this presentation will look at genetic variations. Patients with genetic variations may require different treatments or doses or toxicity prevention. In addition to plasma levels, researchers also are increasing knowledge about metabolic processes in target tissue for a drug to be effective, and how targets interact. The metabolic processes in target tissues may be the same or different as in the liver.

Integrating pharmacogenetics into therapeutics has received a great deal of recent attention, including a statement last year from the FDA commissioner, emphasizing the need to identify patients who need different doses or are prone to certain toxic effects, since these efforts can help maximize drug benefits and minimize toxicities for all.

Work in a cancer center requires understandings not only of genetic variations of the host genome, but also genetics of the disease state, the tumor. Drugs can be selected by the presence of targets, as in the cases of herceptin and Gleevec, and as more rational drugs are developed, this method will become increasingly important. Virtually all phase II enzymes are genetically variant. Clinical studies are needed for these compounds.

The classic pharmacogenetic approach for finding single gene versus multi-gene phenotypes goes from phenotype to protein to cDNA/gene to variation, and the variation is then correlated back to the phenotype. As more knowledge is generated about genetic variations, this process is better understood. Thiopurine methyl transferase (TPMT) deficiency, in which the TPMT alleles cause complete TPMT deficiency, is a good example of the paradigm of how genetic variation can affect enzyme activity and genetic information can determine an individualized dose that is most effective and least toxic.

Now the field needs to move from SNPs to clinical predictors, to develop dosing levels for other drugs in which genetic variation plays a role. An example of this is work that has been done with tamoxifen metabolism, which has been linked to a polymorphism that is associated with a low level of sulfotransferase 1A1 (SULT1A1) activity. The hypothesis is that those patients with rapid activation/slow inactivation will have higher plasma concentration and higher toxicity, those with slow activation/high inactivation will have less response.

Looking at the pharmacogenetics of another anti-cancer drug, the major metabolic pathway of irinotecan is SN-38, and that is what is responsible for the compound's antitumor activity. Variations in UGT1A1 effect level of expression of SN-38-G, and studies in Dr. Blanchard's laboratory found an association between low activity UGT1A7 alleles and favorable tumor response, as well as low toxicity, contradicting previously published reports. Similar data were found for UGT1A9.
The toxicity findings seem counterintuitive, highlighting the need for sophisticated modeling to incorporate genetic modeling into toxicity. Further considerations for implementing pharmacogenetics in drug therapy are the need for pharmacogenetic research centers, the knowledge of the clinical team about using genetic information to individualize treatments, availability of genetic tests, patient privacy, and liability, which may have a large role in driving this research.

DISCUSSION

A participant asked if dietary components might have a role in metabolizing endogenous proteins, and Dr. Blanchard agreed that was likely.

RENAL TRANSPORTERS IN PREDICTIVE ADME
Kathleen Giacomini, Ph.D.
University of California–San Francisco School of Pharmacology

While it has been determined that 39 percent of drugs that enter clinical trials fail for pharmacokinetic reasons, other categories—lack of efficacy (30 percent) and adverse drug reaction (ten percent)—may also be pharmacokinetic in nature. The determinants of pharmacokinetics are ADME, and predictive ADME, as illustrated in previous presentations, must consider population variations.

In preclinical drug development, the focus on drug metabolizing enzymes is in the liver. However, membrane transporters are also of interest, working in concert with drug metabolizing enzymes in the intestine, liver, and kidney to aid absorption and elimination. Approximately 30–40 percent of currently marketed drugs are eliminated primarily by the kidney, and the kidney may be the sole route for many drug metabolites. This presentation focused on elimination and drug accumulation in the kidney.

In considering predictive ADME, one question is what to predict. Dr. Giacomini proposed renal clearance, drug accumulation in the kidney, and variation in renal clearance.

The functional unit of the kidney is the nephron, and the process involved in renal handling of drugs includes filtration by the glomerulus, secretion and reabsorption in the proximal tubule, and passive reabsorption in the distal tubule. Renal clearance is the sum of filtration and secretion clearance less reabsorption, a formula that works for filtration drugs. Active secretion in proximal tubule epithelium requires polarized distribution of transporters. Little is known about the role of transporters in reabsorption, but this is a rapidly evolving field with a number of transporters identified and associated with specific drug actions.

Organic cations and organic anions are transporters in the kidney in the SLC22 transporter family. With a candidate drug, several things must be determined. Does it interact with these transporters? Is it a substrate? Is it an inhibitor? What are its kinetics? And what are potential drug-drug interactions?

To develop predictive models of renal elimination, researchers need to know if the transporters in the kidney are localized, what their mechanism and direction of transport is, the expression level of transporters, substrate and inhibitor specificity, kinetics, and regulation. Tools needed to develop predictive models include basic transporter studies, particularly of transporters in pathways; stably transfected cell lines from animals and human transporters; isoform specific inhibitors; knockout mice to learn specific roles; and predictive models to develop and test.

There is much less variation in renal clearance than hepatic clearance, although this differs from drug to drug. Factors that can contribute to variation in renal clearance include level and activity of transporters, which may be influenced by age, gender, and concomitant drugs, but also by

genetic variation in transporters. Researchers are looking for polymorphisms in the coding region of renal drug transporters in the SLC22 family.

The aims in examining the pharmacogenetics of the membrane transporters project are to move from genotype to phenotype—from genetic variation to cellular phenotype to clinical drug response. Genetic variants of the transporters are identified in ethnically diverse populations and are listed at the Pharmacogenomics and Pharmacogenetics Knowledge Base, www.pharmgkb.org.

One study, SOPHIE (Studies of Pharmacogenetics In Ethnically Diverse Populations), has enrolled 600 individuals. In looking at cellular phenotypes, site-directed mutagenesis is used to construct protein-altering variants of transporters. Functional diversity of transporters, which clearly could contribute to clinical differences, has been observed. Each ethnic group studied had reduced function variants, and the variants were population-specific. Variants that retain function are more common than reduced function variants.

Looking, then, at clinical drug responses, one of the problems with pharmacokinetic studies is the difficulty in determining if drug response runs in families, since the genetic component to the variation is unknown. In a study in twin pairs of metformin, an antidiabetic medication that is eliminated exclusively in the kidney, highly concordant renal clearance was found in healthy identical twins. Preliminary heritability was estimated at 95 percent, although this would decrease in less healthy pairs.

In an on-going clinical genotype-to-phenotype study of Adefovir, an antiviral drug used to treat hepatitis B that is cleared through the kidney, the OAT1 transporter appears to play a role and this is instructive about human biology.

IMPACT OF ACTIVE TRANSPORT ON TISSUE DISTRIBUTION OF DRUGS
Noa Zerangue, Ph.D.
XenoPort

Active transport plays an important role in tissue distribution of drugs, and it is clear that there are xenobiotic transporters that are relevant in both the liver and kidney. Active transport can influence drug absorption, distribution and elimination. While the small intestine gets a lot of attention, absorption in the colon is less understood, but it is important for extended release drugs.

Drugs have very different levels of permeability, illustrated at one end by Gleevec, with high permeability and mostly passive diffusion; and on the other by gemcitabine, with negative passive permeability, its ability to get across membranes dependent on transporters.

There are more than 250 active transport proteins in the human genome that could potentially interact with drugs. The best characterized xenobiotic transporters are ABC efflux pumps and OAT/OATP/OCT, but there are also several families with many orphans of interest.

The transport of lipophilic amines is not well understood; many are recognized through screening, but the physiologic part of the puzzle is missing. Broadly expressed transporters recognize a diverse range of substrates, and also interact with xenobiotics. The active transport of drugs often occurs at barrier tissues; the important barrier tissues are the blood brain barrier

(BBB), intestinal mucosa, and others that are less studied but important to consider in terms of full body transport of drugs. These include the choroid plexus, renal and hepatic tissues, blood-testis barrier, placenta, and mammary epithelia.

Active transport affects oral drug bioavailability. Efflux transporters may limit oral drug absorption, but usually saturate at typical drug concentration in the intestinal lumen. A variety of antibiotics, amino acid analogs, and nucleoside drugs are actively absorbed. Valiciclovir, an anti-HIV drug, is an example of a compound with active uptake in the intestine; it conjugates to PEPT1, a peptide that improves bioavailability.

Gabapentin prodrug is an example of a compound with improved bioavailability through colon uptake. However, there is little published literature on transporter segmental expression along the length of the human gut, or estimates of individual variability. In current work, researchers are working with more than 100 intestinal mucosal biopsy samples from routine endoscopic procedures in different parts of the gut. With quantitative polymerase chain reaction (qPCR), they measured 277 transporters and 63 metabolic enzymes in each sample, to derive a dataset that provides a clear picture of which transporters are expressed, segmental patterns, and variability from individual to individual. This will be published in the near future.

Penetrating the BBB is an important unsolved problem in pharmaceutical development, and less than three percent of compounds show high brain penetration. Existing commercial central nervous system (CNS) drugs are usually small, with molecular weight below 350. The poor drug penetration translates to high rate of failure for CNS drugs. *In vivo* studies are usually difficult and inconclusive, and better BBB permeability screening tools are needed.

Tight junctions prevent diffusion of compounds into brain, and brain vasculature is much tighter than in other organs. Brain microvasculature cells highly express many transporters. MDR1 is a multidrug transporter that has a profound effect on brain absorption of drugs. This P-glycoprotein (Pgp) has a more definitive effect on brain penetration than oral bioavailability, and the Pgp level is predictive of what drugs get into the brain. Other drugs get into the brain through active transporters.

XenoPort is collaborating with Pfizer on the validation of active transporters in BBB. Scientists are investigating gene expression analysis in purified capillaries from humans, mice, and rats; protein analysis and subcellular localization in rats and humans; screening a panel of permeable drugs for interaction with BBB transporters; and *in vivo* validation using inhibitors. Active transport is also being studied in relation to chemotherapy, looking at drug resistance, antimetabolite drug uptake into tumors, and patterns of adverse toxicities. Understanding transport may aid in judicious targeting.

Active transport of cytotoxics can cause tissue selective toxicity, and interaction with concentrative transporters may improve activity of chemotherapeutics. This is seen in uptake of gemcitabine, which may have implications in treatment of lung and pancreatic cancers.

In conclusion, efflux transporters (Pgp) are well validated; active transport is likely to be important for drugs that mimic amino acids, nucleosides, peptides, and folates; and many innocuous transporters harbor an unrecognized capacity to transport xenobiotics. Active transport may have unexpected effects (positive or negative) on permeation through barrier tissues by drugs with modest or low passive permeability.

Areas recommended for further research areas include upgrading tissue culture models; looking at some important and little studied barrier tissues (testis, retina, placenta) other than BBB; validation outside the liver and kidney to address whether transporter inhibition or accumulation leads to toxicity or variable pharmacokinetics; and improving existing pharmacological tools for *in vivo* validation.

Toxicogenomics Session

TOXICOGENOMICS: APPLICATION OF GENE EXPRESSION IN TOXICOLOGY SCREENING, MECHANISTIC ELUCIDATION, AND BIOMARKER IDENTIFICATION
Donna Mendrick, Ph.D.
Gene Logic, Inc.

Gene Logic is a company that integrates genomics with drug discovery and development. It works with a number of pharmaceutical companies in database development. Two systems are being tested to augment biomarker discovery: ToxExpress® is a large reference gene expression database, and BioExpress™ is a large gene expression database of normal and diseased human tissues.

The company uses toxicogenomics to respond to customers' wishes for two test systems: *in vitro* primary hepatocytes and short-term *in vivo* experiments in humans as well as rats. These systems can be used for statistical prediction of proprietary compounds to induce toxicity in humans; mechanistic insight to a compound's toxicity by comparing its gene dysregulation to that of known drugs and chemicals; and biomarker discovery. Key elements of experimental design include range-finding studies, *in vivo* experiments that look at adverse events, and *in vitro* experiments with 10–20 percent reduction in mitochondrial functioning. This phenotype anchoring assures a gene expression response.

The *in vitro* studies focus on primary rat hepatocytes; the *in vivo* work on rat liver, kidney, heart, bone marrow, and blood. This presentation will focus on rat liver.

Predictive toxicogenomic models are not yet *in silico*. Steps are to expose a rat or cell culture to an NCE, then researchers investigate if is likely to be a human toxicant, what mechanism of injury it will induce in humans or rats, and what other compounds it resembles. Rat-specific hepatotoxicants, human specific liver toxins that are not toxic to rats, and multi-species hepatotoxicants were used to build the models, with non-toxic negative controls, including controls that are non-toxic to the liver but exert pharmacologic action.

Success rates were high (98 percent for hepatotoxicants; 92 percent for non-hepatotoxicants), with human liver effects predicted and cross validated correctly for a number of compounds including flutamide (human-specific hepatitis), acetaminophen (necrosis), tamoxifen (cholestasis), tetracycline (microvesicular steatosis), and metformin (pharmacologically active and nontoxic). Gene Logic customers have used the predictive models with more than 150 compounds.

An example of detection of human-specific toxin in rats is seen with felbamate, an anticonvulsant approved in 1993 and restricted the following year because of hepatotoxicity.

Preclinical animal testing did not foresee hepatotoxicity, but predictive modeling revealed it to be a potential human hepatotoxin that might cause hepatitis.

Mechanisms of toxicity outcomes support drug development decisions to rank order by toxicity and/or mechanism and evaluate sister compounds in parent series for toxicity, and to explore gene expression changes to find potential candidate biomarkers. Gene Logic has conducted analysis on individual genes and more than 360 biological pathways to derive a toxicogenomic assessment.

Toxicogenomic studies on clofibrate, a hypolipidemic agent that induces increase in mitotic activity in hepatocytes and liver enlargement with short-term use and liver tumors in rodents with long-term use, predicted that the compound would not induce hepatic injury in humans, a hypothesis supported by mechanistic analysis. These studies raise the comfort level of using this drug and help move studies forward.

Looking at biomarker discovery, possible biomarkers for tissue injury are from RNA derived from tissue biopsy, which tends to be suboptimal for clinical application, or circulating white cells; and protein or analyte in serum or urine, which is easier to do. Biomarker discovery can start with the ToxExpress database and move to BioExpress, the human database, or it can start with a human disease state. An example used was Arginase 1, a known biomarker of hepatotoxicity. The toxicogenomic analysis confirms it as a predictive marker for hepatotoxicity that is only seen in the liver, with the dysregulated gene found in diseased livers.

Dr. Mendrick concluded that the value of large databases is that they identify normal expression ranges of genes; they allow statistical confidence and selection of appropriate toxicity-responsive genes, telling scientists what to ignore and what to focus on; they allow robust toxicogenomic models; they empower mechanistic analysis of NCEs; and they enable biomarker discovery across species.

DISCUSSION

In response to a question about how "dysregulation" is defined, Dr. Mendrick said she and her colleagues look at the amount of mRNA present in gene expression. Another participant asked why an *in silico* approach is not used, and Dr. Mendrick said it is premature for that methodology, because gene expression is not linked to the structure of a compound.

Another participant noted that as part of the NIH team that developed felbamate, it was observed that of the few people who died of liver disease, most of the fatalities were related to aplastic anemia. The participant then asked if this could have been predicted. Dr. Mendrick answered that bone marrow toxicity is being studied as well as hepatotoxicity.

In response to a comment about NIH's large number of existing databases and the possibilities of working partnerships among academia, industry, and government, Dr. Mendrick noted that this is an interesting concept.

DATABASES/DATA ANALYSIS - SESSION

DATA ANALYSIS USING STRUCTURED BIOLOGICAL KNOWLEDGE: EXPERIENCE WITH A DATABASE OF PROTEIN FUNCTIONAL INTERACTIONS
Alan Ruttenberg
Millennium Pharmaceuticals

Mr. Ruttenberg described himself as a tool builder for molecular biologists. A paradox of biological knowledge is that so much is available, but it is so hard to use. Effective work in both discovery and toxicology depends on access to previous scientific results. Researchers are constantly hunting for papers relevant to their problems, a time consuming and error-prone process. Much useful information is buried in research documents, reference books, and regulatory documentation.

Specific to toxicology research, if a compound is found too toxic to be useful at an advanced phase in the discovery process, a great deal of money has been spent and individuals may have been harmed. If toxicity is found too early in discovery, important work may be cut off. Researchers seek to preemptively avoid adverse side effects by searching the literature and regulatory documentation, but that hunt can be prodigious.

Mr. Ruttenberg proposed systematically building a two-pronged approach to exploiting information:

- First, organize and structure knowledge, integrating public and licensed and internal databases, and collaborating with Ingenuity Systems, a company with the world's largest curated database of biological networks.
- Second, develop methods to exploit the ensemble of information as a whole. Investigate network methods, develop tools for working with sets, and make querying easier. Scientists want to be able to compute, to use all the information at once and not just overlay parts, and can best do this in a structured knowledge base that has been standardized for synonyms and with other built-in tools.

This is of interest because specific analysis methods may be able to be applied to toxicogenomics, adding to understandings of drug metabolism. It is a productive approach to gathering, structuring, and use of knowledge, and may have further uses not yet conceived of.

Structured knowledge is the same information represented in the same way, normalized for synonyms of entities, and processes, with taxonomy where appropriate. It is computable, and more intricate and dynamic than any human-drawn diagrams system could be. The tools his company builds, Mr. Ruttenberg said, are not push-button answer type tools but complex time-saving instruments for very experienced biologists.

Case 1, a category server, is one illustration. The goal is to help turn experimentally devised sets of genes into hypotheses about function. The knowledge used includes sets of genes (categories) based on shared properties such as functional class, involvement in similar process, subcellular location, and differential expression in some other experiment. The strategy is to develop scoring functions that enable comparing one set of genes to another, exhaustively

score selected categories, and present the user with a ranked list of results. Millennium's category server includes the Gene Ontology (GO) consortium, gene sets associated with processes extracted from the Ingenuity database, gene signatures extracted from papers by Millennium staff, ligand pathway members, gene expression comparisons, transcription factor targets, genes by subcellular location, and an ever growing list of resources. Category scoring functions look at ranking (e.g., genes sorted by differential expression), overlap, and interaction count (e.g., biomarker set).

A specific category server example is the dynamics of hematopoiesis, beginning with a marrow injury in a baboon. The result of the investigation is a large number of categories related to cell proliferation, indicating that cells were still largely quiescent.

Another example is Case 2, an activity center analysis. The goal is to help biologists understand expression profiling experiments in terms of prior knowledge. The knowledge used includes pairs of interacting proteins, inferred from facts on gene products, across human, mouse, and rat networks where nodes are proteins and edges are interactions. The strategy is to use statistics on average activity of gene and interactors as a less "noisy" signal, and to display using network layout based on the functional "distance" between proteins.

Looking at protein functional relationships as a network, the activity center analysis determines that the full interaction network (functional interactions involving gene products) plus the data defining activity (expression comparisons) yields the active sub-network (hints on the cellular processes, i.e., perturbation by a compound, downstream of a target, involvement in drug resistance). Scoring activity computes an activity score for each gene in the network and uses Monte Carlo simulation to determine distribution of scores to yield a p-value answering how unusual the level of activity.

A specific activity center example is the effects of an IKK2 inhibitor. The primary question is whether IKK2 inhibitor pretreatment reduces response to lipopolysaccharide (LPS) stimulation. Inflammatory pathways were upregulated in the LPS-treated cells, but downregulated when the IKK2 inhibitor was added.

The third example, Case 3, deals with edge-count statistics. The goal is to develop new ways to exploit interaction network structure to mine and analyze data. Knowledge used is a combination of networks and sets, and the strategy is to apply random network theory to category scoring, module discovery, and list expansion. A problem with counting edges is they may not have the same significant variation that needs to be captured. Working with randomized networks, at each step pick two edges and swap end nodes, leaving each node with same number of edges after a swap, but 25 swaps later in the network illustrated, there are four edges between pink and blue sets, compared to one in the initial network. The modified random network is better than the Monte Carlo simulation and can be used to derive analytic formulas. An example application is annotation for DLG3, which controls certain excitatory synapses. There is very little GO annotation on DLG2, but Ingenuity's database includes 21 interactions. This can help biologists navigate gene function.

In conclusion, the genomics technologies demonstrated are powerful tools for understanding. Having structured knowledge opens a fertile field for developing new data analysis techniques, and technology and expertise have advanced to the stage to make building these kinds of databases feasible. This can be applied to toxicology by:

- Exploring these methods for toxicogenomics, e.g., the effect of toxic compounds on cell and tissue systems
- Identifying and prioritizing information sources of value, for example, journals, databases, references, and regulatory documentation
- Sponsoring public and industry efforts to build a knowledge base
- Working with academic publishers to establish standards for submission of structured knowledge, in model of Genbank
- Thinking about other networks of information in individual areas of expertise and opportunities to exploit it

DISCUSSION

A participant asked if when approaching analysis the origin of information used could be documented sequentially in order to tear apart the contributions to the scoring. Mr. Ruttenberg said yes, that is possible; all the information is carefully recorded.

CONSTRUCTION OF A HUMAN ADVERSE EFFECTS DATABASE FOR MODELING QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIPS (QSAR)
Naomi Kruhlak, Ph.D.
U.S. Food and Drug Administration

The Informatics and Computational Safety Analysis Staff (ICSAS) of the FDA's Center for Drug Evaluation and Research was introduced to workshop participants in Dr. Contrera's presentation on Day 1. ICSAS is an applied regulatory research group that develops databases with animal toxicology and human clinical data, develops QSARs, and facilitates this with leveraging agreements to work directly with software developers, benefiting both regulatory agency and industry.

QSAR is a statistical correlation between molecular structural features and toxicological activity at a specific endpoint. It can be fragment based or molecular descriptor based. In both cases, chemical structure is the key field. Construction of QSARs requires toxicity endpoint data, chemical structure, and a suitable modeling platform.

Predictive models are needed when data are limited or lacking, or when findings are equivocal. Specifically for human adverse effects, predictive models are needed when risks are not adequately identified by conventional means. Animal findings do not adequately represent certain types of adverse events in humans, and clinical trials utilize only a small patient population. Current animal effects model systems look at carcinogenicity in rodents, tetratogenicity and reproductive toxicity in mammals, mutagenicity in *Salmonella typhimurium*, other genetic toxicity, maximum tolerated dose, and acute toxicity.

The first thing needed to predict human adverse effects is a good dataset. This can be found in labeling, as reported in the *Physicians' Desk Reference* and FDA/CDER and industry archives; in published literature; and in post-market surveillance adverse effect data, as compiled by MedWatch, the World Health Organization, and other programs. ICSAS decided to use post-marketing surveillance data because it is most accessible to computer use.

ICSAS decided to use the Spontaneous Reporting System (SRS), which contains approximately 1.5 million adverse drug reaction reports from 1969–1997. Problems with adverse effect data are that adverse events are severely under-reported and dependent on severity of effect and length of time a drug is on the market. Also, no chemical structures are provided, data are often listed by trade rather than generic name, and there is no patient exposure denominator.

Raw data were downloaded by drug from an Oracle database and pooled for each drug, with trade names mapped to generic names. Some 8,620 trade names were mapped to 1,861 generics, which were linked to 1,515 organic compounds that were identified as suitable for QSAR modeling. Estimate of patient exposure is needed as the denominator, and shipping unit data was used to derive this information. That was available for 610 of the 1,515 compounds. The denominator is important to yield an incidence rate of adverse effects.

The highest rate and number of adverse reaction reports were in CNS drugs. The reports were grouped under 22 organ system terms, and six organ systems were selected for further investigation: immune system, heart, cardiovascular, liver, kidney, and bladder. Individual and composite models were created. The QSAR model, using MCASE/MC4PC technology, reduces the database to 2–10 atom fragments, identifies those associated with active molecules,

identifies modulators of activity, assigns training set compounds an activity score, and runs a test compound against a model to identify structural alerts. It then prioritizes alerts, which are scored based on the number of chemicals their activities. There are three categories of alerts: decision, possible decision, or nonsignificant. To be positive, the same decision alert must be identified at two or more related endpoints.

A suitable reporting index cutoff is then needed to define active compounds, a value high enough to overcome background noise and low enough to retain sensitivity. The majority of compounds had a reporting index between 1 and 10. At the higher reporting index values, there were fewer active compounds with better clustering. At lower values, there were more compounds, more false positives, and poorer clustering. It was determined that a median cutoff is preferable, and the percentage of decision alerts is maximized at a reporting index cutoff of four. These optimized parameters were used to build four models: liver effects in humans, heart and cardiovascular effects in humans, kidney and bladder effects in humans, and immune system effects in humans.

In conclusion, SRS data are suitable for QSAR modeling but need noise suppression and signal enhancement. The database is flexible and shows potential for use with multiple QSAR software platforms. However, it is on the threshold of having too few chemicals and new data will need to be added.

Future considerations are to prioritize lead compounds based on the likelihood of human adverse effects, and to focus attention in the review process on organ systems where there is a greater chance of adverse side effects.

The meeting then broke into concurrent discussion groups to consider additional areas needed to improve preclinical ADME and toxicology analysis. On Thursday morning, concurrent discussion groups met to consider research needs in two general areas: ADME/metabonomics and in vitro-in vivo models/toxicogenomics. The recommendations of the discussion groups were presented at the closing session of the workshop and on Thursday are summarized below.

Discussion groups

WEDNESDAY, JUNE 16, 2004

PRECLINICAL PREDICTIVE ADME AND TOXICOLOGY ANALYSIS

Recommendations for research not addressed in summit.

GROUP A
Moderator: Curtis Klaassen, Ph.D.

1. Establish sample banks of tissue and DNA from people who experience adverse side effects from drugs. Preliminary work in this area is already being done through the Hepatotoxicity Clinical Research Network of the National Institute of Diabetes & Digestive & Kidney Diseases (NIDDK). Sample banks could work with clinical groups that have samples available that can be used for genomics and other omics testing.
2. Improve methods for quantitative predictions of P450 mediated metabolism. This should consider conformational flexibility of enzymes and steric and electronic properties of substrates; getting new structures; protein dynamics in solution; and new computational methods, including the development of databases about the propensity of metabolism of adjacent sites. Additional theoretical approaches are needed.
3. Develop tissue and cell cultures, co-cultures, and 3D cultures that mimic liver and other organs/tissues. This is a rich arena for development of improved models and systems. Participants urged a program endorsed by Congress to encourage researchers to develop human models that are lacking.
4. Develop methods to standardize "omics" so that scientists can compare data sets (e.g., housing); methods should include tissue and drug levels. For example, how animals are grouped per cage and the effect this has on stress genes is a factor that is often ignored. Housing, feeding, and all other factors related to animals need to be considered. This is also relevant to human studies; understandings of variations in human populations are necessary to reach valid conclusions.
5. Interpretation of "omics" is also needed. Currently there is a lack of annotations in metabonomics, and a need to focus on how to develop the right models to use "omics" data in conjunction with other models such as blood data and cell proliferation.
6. Better and shorter methods to predict carcinogenicity are needed.
7. Phase II drug metabolism studies lag behind the study of P450s and transporters, and more work is this area is needed.
8. More work is also needed with transporters. Specifically, the relative importance of phase II enzymes to transporters and the disposition of drugs need further elucidation.
9. Identify tissue repair and protective mechanisms. While initiating events in toxicology receive considerable attention, much less is knows about protective mechanisms once there is an insult to a tissue. Some preliminary research in the area has begun, but it represents a relatively new way of thinking.

10. Application of imaging technologies such as magnetic resonance is needed for liver and other tissues in humans and animal models, as well as tissues and tissue culture.

11. Research is needed into nonalcoholic steatohepatitis (NASH), an increasingly common liver disorder associated with obesity. It is unknown if individuals with NASH are more susceptible to drug induced toxicity or its effect on pharmacokinetics.

12. Mathematical modeling (*in silico* models) of systems biology is needed. A standard concept of what a system is and how the "omics" are integrated would help further understandings.

13. siRNA models of knockout enzymes and transporters are needed. This would include *in vivo* and *in vitro* delivery of siRNA and further development of conditional knockouts.

14. The importance and interactions of various cell types in liver needs further investigation. This includes what role they play in hepatotoxicity, fibrosis, steatosis, and stellate cells, which store vitamin A and may produce collagen in chronic liver disease.

15. Research is needed into predicting developmental toxicity. The interaction of development and the "omics" should be studied for understandings of using adult drugs in children and in pregnant women. Developmental responses in developing children and adolescents cannot necessarily be predicted from adult genomics.

16. Nanotechnology is an evolving area. This might be useful for diagnostic techniques, traversing the blood-brain barrier, and issues related to delivery systems and toxicity.

GROUP B
Moderator: Kathleen Giacomini, Ph.D.

1. The use of *in silico* methods in predictive toxicology. For example, a small company in Barcelona, Spain, has technology for entering information from conferences into a computer library. This company has a master database of pharmaceuticals, based on information from patents, and programs that can use this information to predict pharmacologic properties.

2. The development of complex, three-dimensional systems that provide an environment for cells to interact and make the extracellular matrix and other tissue components. In the human body, stromal cells drive tissue building, and an environment must be provided where the stroma can interact with parenchymal cells. Simple systems, such as sandwich cultures or seeding cells on scaffolds, will produce nothing more than a two-dimensional monolayer.

3. A clear definition of toxicity. In some cases, potential drugs are discarded because they exhibit toxicities for one indication, but they may have therapeutic effects with other indications.

4. The development of cheaper alternatives to GLP and Ames tests. Replacements for the Ames test are available at the drug-discovery phase. An early-stage reactogenicity test may be useful, but such a test would change the regulatory paradigm and require a great deal of effort. Replacement tests must also be validated.

5. A focus on genetic variation and drug-drug interactions at the drug-modeling and dosing levels. A set of known predictors and the appropriate models should be developed and tested in human populations.
6. Collaboration among government, academia, and industry in drug discovery, making use of available databases, knowledge, and experience. NIH has a database, PubChem, where all NIH Roadmap data is being entered.
7. The development of imaging methods, both *in vivo* and *in vitro*, employing PET or CT. Such modalities could be used to examine both phenotypes and molecular or fluorescent markers.
8. Studies of age-related genetic and phenotypic changes affecting drug toxicity. ADME models are particularly needed for pediatric populations.
9. The exploration of small animal models, such as zebrafish and C. elegans.
10. The development of new, immortalized cell lines to supplement current studies done in organs, and ways to maintain differences in these cell lines.
11. The development of mechanistically based models.
12. The evaluation of combination therapies to improve their efficacy and safety.
13. Improved human dose projections done with allometric scaling, using new paradigms and incorporating animal *in vivo* data.
14. An ethnicity-based stem cell bank that allows investigators to draw from specific populations and examine differences in gene expression.
15. Transgenic and knockout models in animals other than mouse.
16. The application of siRNA, which conditionally dampens gene expression, to ADME studies in animal models.
17. A comparison of pharmacology data between synthetic compounds and herbal compounds and the application of this comparison to ADME.
18. The development and improvement of microanalysis methods to lower the amount of sample required for drug-discovery efforts.
19. Instrument development for mass spectroscopy and other technologies. Efforts are under way in proteomics (the Plasma Proteomics project) and metabonomics (SMRS), but standardization is needed.

DISCUSSION GROUPS

THURSDAY, JUNE 17, 2004

The charge for these groups was to identify and define pressing areas of research and how these areas best could be specified in a Request for Applications (RFA).

SESSION A: ADME AND METABONOMICS
Discussion Leader: Steven Wrighton, Ph.D.

1. Encourage applications for techniques to restore normal phenotype of cells, reflecting liver, lung, kidney, intestine, and heart; and also to improve models of barrier function. These models should also include systems to address population variability.
2. Encourage applications that will address the need for various reagents, specific substrates, and inhibitors of phase II enzymes and transporters, including knockout and transgenic animals.
3. Encourage applications to develop transfected cells for individual transporters and drug metabolizing enzymes and defined multiple enzyme pathways.
4. Develop *in vitro* and *in vivo* models of idiosyncratic and other toxic reactions, which includes banking cells and DNA of patients with known idiosyncratic and toxic reactions. Improve access to this material.
5. Develop databases of ADME/toxicity parameters. Encourage applications that will define, curate, and annotate databases, so information can be acquired efficiently. An adverse events database, as described in one of the presentations, would be very useful.
6. Application of metabonomics to clinical material (human systems toxicology); develop and validate new biomarkers including new analytic techniques to study the metabonome.
7. Studies to understand intracellular trafficking of drugs and toxins to their sites of action.
8. Understand the variability of subjects on "omics."
9. Improved approaches for quantitative prediction of metabolism and transport, including new computational approaches, structure of the proteins, knowledge bases, ligand properties, and physiological parameters, including basic information on variability of proteins in tissues of interest.
10. Quantitative approaches to predict induction of the drug metabolizing enzymes and transporters *in vivo*, from *in vitro* data.
11. Encourage academic-industry collaborations in grants. Industry representatives pointed out that this can be a difficult endeavor, and for success research must be sponsored in a way that protects intellectual property rights.

SESSION B:* IN VIVO *AND* IN VITRO *MODELS OF DRUG TOXICITY AND TOXICOGENOMICS
Discussion Leader: Bryan Roth, M.D., Ph.D.

1. Encourage applications related to idiosyncratic reactions, as related to hepatotoxicities. Idiosyncratic reactions, big killers for drugs on the market and in clinical trials, are difficult to predict, and there is no consensus on how to define them, especially in terms of human-specific reactions. Gene Logics has been able to predict, with surprising accuracy, the idiosyncratic reactions for a number of compounds, and there are some differences in gene expressions between these reactions and other toxicities. The following are needed to further explore idiosyncratic reactions:

   - ❑ Models that mimic the human environment as closely as possible.
   - ❑ Better detection methods.
   - ❑ Humanized, multicellular cultures; work on these systems should continue at the R01 level.
   - ❑ A governmental repository for engineered cells, which would standardize media, processes, and matrices.
   - ❑ Blinded profiling with a large number of compounds, particularly those that have been withdrawn because of the incidence of idiosyncratic reactions; government and academia should collaborate to screen compounds with known toxicities and identify predictive patterns.
   - ❑ Improved access to drugs that have been killed because of idiosyncratic reactions.

2. The use of animal models to examine a narrowly defined goal versus using a global approach. As discussed by Dr. Luc Maroteaux, once a molecular target for a side effect has been identified, one can engineer models and test them in sophisticated and definitive ways. Developing animal models to predict dosing involves a lot of money and time, perhaps too much to pursue such narrow targets, and the number of animals used for these studies raises ethical concerns. Yet animal models are still useful. For example, they may provide a more efficient way to distinguish QT effects from the deadly Torsades des Points, if better assays are developed. At present, many potentially useful drugs are killed because of QT effects, but the relationship between QT and Torsades is tenuous.
3. A database of animal models. Jackson Labs has an online encyclopedia of mutant mice discovered during the past 30 years.
4. The use of smaller organisms such as zebrafish. Many of the genes involved in toxicity may be conserved from small organisms to humans. A group at Massachusetts General Hospital has used zebrafish to test 100 compounds known to cause QT effects and found that they cause brachycardia in zebrafish. Lower organisms offer more convenience and lower costs. Yet these organisms may not be best for screening, because a single amino-acid change in the binding pocket of most compounds changes their pharmacology, and it's likely that some amino acid changes have occurred during evolution. Even if the drug were converted to an active metabolite, zebrafish and other

lower organisms may not be ideal for screening, because the cytochrome P450 substrate specificity may be different, and some metabolizing enzymes present in one species may not be present in another. However, zebrafish and other lower organisms can work well as discovery and validation tools.

5. *In vitro* models and biomarkers. Consensus is needed on how to define, validate, and use predictive toxicity biomarkers. Some current biomarkers, such as tropinin, measure events that have already happened. Concerted efforts are needed on the discovery and basic science end to identify candidate biomarkers, using proteomics and metabonomics, and translate them to the clinic. In addition, past experience can be used to form guidelines for predictive biomarkers. Validation of new candidates should include a determination of their relationship to disease.