A Mechanistic Redefinition of Adverse Effects – a Key Step in the Toxicity Testing Paradigm Shift

Kim Boekelheide [1] and Melvin E. Andersen [2]

[1] Department of Pathology and Laboratory Medicine, Brown University, Providence, RI, USA; [2] Program in Chemical Safety Sciences, The Hamner Institutes for Health Sciences, Research Triangle Park, NC, USA

Summary

The efforts of the committee that produced the report on Toxicity Testing in the 21st Century reflected the need to look at the issue of assessing risks to humans from exposure to various chemicals through a lens of 21st century biology. The problem – determining if there is a risk of specific exposures – is as old as humanity; every generation brings its own perspective and tools for examining the problem and coming to answers and solutions. Bringing this generation's tools to bear requires us to see the problem of chemical risk assessment in a different light, both in terms of testing of toxicity pathways in vitro and in the interpretation of the tests for estimating whether exposures will be safe. One key issue will be to assess when pathway perturbations are believed to be excessive, i.e., when they are deemed adverse. Redefinition of adversity based on in vitro testing will require a new perception of dose response functions as probabilities of failures, with multiple underlying processes acting sequentially and in parallel leading to failure at a cellular and an organism level. These dose response relationships for adversity will also require a computational systems biology approach for examining toxicity pathway dynamics and stress pathway overload. While the overall approach of defining adversity for in vitro endpoints and using this definition of adversity for risk assessment can be painted in broad brush strokes, as we have done here for DNA-reactive compounds, it will take implementation with a series of prototypes to show the process in practice.

Keywords: toxicity testing, toxicity pathway, adverse response, carcinogenicity testing

1 Introduction: defining adversity

In 2007, the National Research Council released a report entitled "Toxicity Testing in the 21st Century: A Vision and a Strategy" (NRC, 2007, 2007) that laid out a roadmap for revamping toxicity testing. At the center of this vision is a re-orientation of testing to focus on molecular mechanisms of toxicant effects. Using high throughput assays and human cells *in vitro*, this new testing system would replace the current testing approach that relies on phenotypic responses in animals. The NRC committee focused on the concept of "toxicity pathways," proposing a systems approach to integrate large amounts of molecular- and cellular-level data to delineate toxicant-induced modes of action. A distinct advantage of the *in vitro* approach is the ability to test across large concentration ranges to develop extensive dose response relationships for various toxicity pathways.

A key element of the reorientation of toxicity testing is the way in which adverse effects are understood, the topic of this paper. Currently, adverse effects are defined in the context of the whole organism. For example, Lewis et al. (2002) defined an adverse effect as "A biochemical, morphological or physiological change (in response to a stimulus) that either singly or in combination adversely affects the performance of the whole organism or reduces the organism's ability to respond to an additional environmental challenge." A similar definition is readily available online: "Abnormal, harmful, or undesirable effect on an organism that causes anatomical or functional damage, irreversible physical changes, or increases the susceptibility to other biological, chemical, or environmental stresses." (BusinessDictionary.com, 2010)

Defining adversity at the level of the whole organism is clearly incompatible with a systems biology based toxicity testing approach that relies upon interpreting toxicant-induced modes of action at the molecular and cellular level. The incompatibility of the current definition of adversity with the toxicity testing of the future leads to the following questions:

1) What are the barriers to redefining adverse effects from the organism to the molecular level?

   a) Are existing regulatory and legal processes so constrained by the current definition of adversity that a radical rewriting of legislation is required to make this change?

b) Is the state of science sufficiently advanced to allow such a redefinition of adversity?

2) How will adverse effects be identified and defined at the molecular level?

Question 1a will be addressed in this section; the example of carcinogenesis as an apical endpoint will be used to explore the state of the science with respect to question 1b. The latter sections of the paper discuss a possible implementation framework for applying a molecular definition of adverse effects for compounds that are potentially DNA-reactive with likely high dose mutagenicity and carcinogenicity in animals.

The legal perspective

In a report prepared by the Duke Center for Environmental Solutions for the Dose Response Specialty Group of the Society for Risk Analysis entitled "'Adverse Effects' and Similar Terms in U.S. Law," Professor Jonathan Wiener and two students (Kelsey Stansell and Mark Marvelli) reviewed the use of adversity-related terms in federal statutes, federal agency regulations, and federal judicial opinions issued since 1970 (Stansell et al., 2005). The conclusion of this analysis was that, while terms such as "adverse effect" are ubiquitous in federal statutes, little guidance is provided as to their meaning. Indeed, the report notes that "the particular scientific methods used to assess "adverse effects'" – such as hazard identification, dose-response assessments (including linear, threshold, or hormetic low-dose extrapolations) – appear to be largely within the discretion of agency scientific staff and their supervisors." As its major conclusion, the report suggests that "the absence of clear definitions of 'adverse effects' and related terms, and the typical deference shown by the judiciary to agency expertise…..may appropriately leave the interpretation and identification of adverse effects in the hands of the most expert institution in the government: agency staff (especially agency scientists)…."

Professor Wiener expounded on the policy and legal implications of including new scientific approaches in US EPA decision-making at a recent meeting (National Research Council Symposium of the Standing Committee on Risk Analysis Issues and Reviews entitled "Toxicity Pathway-based Risk Assessment: Preparing for Paradigm Change," Washington, DC, May, 2009). He noted that several statutes now indicate that agencies should use the best available science, and that the new toxicity testing approaches could be argued to constitute the best science. The overall conclusion: both the use of the new toxicity paradigm and the redefining of adversity at the molecular level rest upon the scientific validity of the new approach and its ability to deliver on the promise of better science.

2 Redefining adversity – the example of carcinogenesis

Cancer is a particularly complex apical endpoint that only manifests all of its well described behaviors – initiation, promotion, progression, metastasis – in the context of the whole organism. In addition, cancer is an all-or-none phenomenon, meaning that within specific categories of cancer, an increasing dose is associated with an increasing incidence of disease but not an altered character of disease. There has been great progress in molecular understanding of neoplasia in the last half century, and these advances should have allowed design of new mechanistic test systems for mutagenic and carcinogenic potential. Unfortunately, the growing understanding of the molecular and cellular underpinnings of cancer has been accompanied by only incremental changes in current carcinogenicity testing protocols.

Current method

Today, the "gold standard" for predicting human carcinogenicity is the 2-year rodent bioassay. While evolving incrementally over time, this test has become standardized by the National Toxicology Program (NTP), which has conducted some 600 2-year rodent bioassays over the past 30 years. In its standard protocol, the NTP conducts 2-year bioassays in two species (rats and mice) beginning at 5-6 weeks of age, in both sexes, and with three doses and a control. The 2-year bioassay is supported by preliminary short-term animal studies providing absorption, distribution, metabolism, and excretion data important to dose setting. The highest dose tested is often the Maximum Tolerated Dose (MTD) with additional doses often selected as the MTD/2 and MTD/4. The MTD is estimated from the short-term studies and is operationally defined as the highest dose that does not significantly increase the death rate or induce more than 10% weight differential compared to controls in the 2-year bioassay.

The 2-year rodent bioassay is recognized as having significant shortcomings as a predictive tool for human carcinogens: some claim that it is too sensitive and non-specific (Ennever and Lave, 2003); others claim that it is not sensitive enough (Huff et al., 2008). The use of the MTD as the top dose is particularly problematic, since most chemicals would be identified as carcinogens in rodents at the MTD with sufficiently large sample sizes (Gaylor, 2005). The MTD is justified because high doses increase the likelihood of seeing a response with group sizes of near 100 rodents. However, the MTD induces biological responses (such as cytotoxicity) that are irrelevant for humans exposed at environmental levels. The induction of irrelevant responses in rodents exposed at high doses creates significant problems of interpretation, i.e., separating those effects that are due only to high dose from those that would persist at lower levels of exposure. This interpretive problem is made even more difficult by the lack of human data for almost all of the identified rodent carcinogens. Indeed, the 2-year rodent bioassay has never been validated, and the small amount of data available suggests that it produces both many false positives and many false negatives (Ennever and Lave, 2003). Despite the fact that the 2-year rodent bioassay itself has never been validated as a predictor of human carcinogenicity, it has become the gold standard for judging the validity of other tests. With a poorly validated model on the pedestal, there is a real and continuing problem in gaining acceptance for a new toxicity pathway-based approach for carcinogen risk assessment.

Cancer modes-of-action

A fundamental mechanistic assumption is that carcinogenicity results from a progressive accumulation of mutations that se-

lect for a growth/survival advantage in the altered clone. Thus, the ultimate major goal in defining *in vitro* tests for detecting carcinogens has been to identify and quantify alterations in the molecular pathways associated with the generation and fixation of mutations. The two major categories of chemical carcinogens are DNA-reactive carcinogens (genotoxic carcinogens) and non-genotoxic carcinogens that alter DNA indirectly through an effect on the cell (Williams, 2001). Genotoxic carcinogens are directly DNA-reactive, are mostly electrophiles, and generally act through the formation of DNA adducts. Thus, genotoxic carcinogens can directly initiate the neoplastic transformation of a normal cell. Non-genotoxic carcinogens, on the other hand, act through the alteration of cellular pathways associated with the modulation of reactive oxygen species, proliferation, apoptosis, endocrine controls, or immune surveillance, etc. This distinction between genotoxic and non-genotoxic carcinogens has been operationally important because risk assessment procedures have treated genotoxic carcinogens as having no threshold, while non-genotoxic carcinogens may be considered to have a threshold for risk assessment purposes, leading to regulation through the use of safety factors. In many ways, it is disappointing that the major contributions of the large body of mechanistic work conducted to study carcinogenic responses (Fig. 1) have only provided risk assessment input for the manner of extrapolating to lower doses (linear versus threshold) or for determining that certain rodent cancers should not be regarded as relevant for humans (such as hydrocarbon nephropathy in male rats or thyroid tumors in rats associated with enhanced thyroid hormone clearance). The challenge faced with *in vivo* mode-of-action studies or with the new *in vitro* mode-of-action toxicity pathway assays is to bring a fundamentally new perspective to dose response extrapolation, not simply toggling between linear low dose and threshold. The assays should optimally provide information to assess a point of departure for risk assessment and to support new approaches for dose response, including dose-dependent transitions and hormesis.

Mutagenicity assays

Dr. Bruce Ames recognized the importance of mutagenicity to carcinogenesis and developed his eponymous bacterial test for the purpose of detecting and classifying mutagens (Ames, 1973). The association of mutagenicity with carcinogenicity is fundamentally important, giving rise to the field of genetic toxicology. Over time, genotoxicity testing has expanded to include a set of *in vivo* and *in vitro* assays, including bacterial and mammalian cells, metabolizing systems, and the investigation of specific pathways of DNA disruption, such as mutation and clastogenicity. Much like other kinds of testing, with increasing knowledge of the underlying mechanisms of toxicant action, genotoxicity tests have been expanded and refined, resulting in the development of a standard genotoxicity battery, including *in vitro* tests, the Ames test, the mouse lymphoma assay, and an *in vitro* micronucleus or chromosome aberration assay. Over the years, these tests have been significantly modified to improve their predictive power, primarily through a lowering of the dose range and a concomitant reduction in non-specific cytotoxicity.

However, an analysis of the performance of this standard battery of *in vitro* genotoxicity tests compared to the 2-year rodent carcinogenicity test showed that this battery is highly sensitive but has very low specificity (Kirkland et al., 2005, 2006). The reasons for this low specificity were explored in a Report of the European Centre for the Validation of Alternative Methods (EC-VAM) Workshop entitled "How to reduce false positive results when undertaking *in vitro* genotoxicity testing and thus avoid unnecessary follow-up animal tests" (Kirkland et al., 2007). This report focused on the details of the test systems and pro-
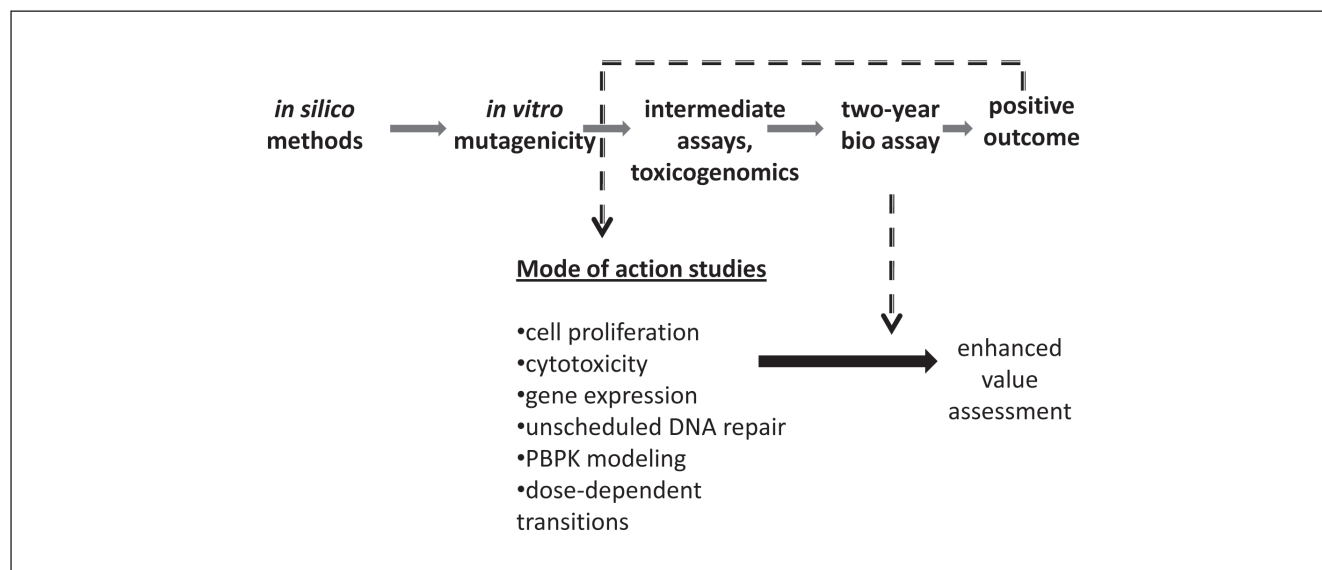
Fig. 1: Mode of action studies for cancer.
In the past this research has usually followed identification of a chemical as a high dose animal carcinogen. Subsequent studies, regardless of detail, usually are used to move from linear low dose to threshold models or to argue for lack of human relevance for a particular endpoint.

posed a research agenda to address their shortcomings, including: (1) changing the culture medium composition and limiting toxicant oxidation, (2) improving the cell lines, (3) designing systems with more normal metabolism, (4) reducing cytotoxicity as a confounder, and (5) designing new cell systems, such as three-dimensional culture models. Ultimately, the report called for elucidating the underlying molecular mechanisms giving rise to a positive test response so that this could be correctly interpreted in the context of our knowledge of carcinogenesis (Kirkland et al., 2007).

Assessing DNA damage and mutagenicity
Remarkable advances in analytical capability now make it possible to identify the formation of DNA adducts and DNA mutations with sensitivities in the order of 1 per 10^8 bases. Swenberg et al. (2008) presented a Framework Analysis approach to examine whether the default assumption of a linear extrapolation of cancer risk was appropriate for genotoxic chemicals. For many chemicals, adduct formation is a key genotoxic event, and it is critical to understand the dose-response relationship between DNA adduct formation and mutation. In the Framework Analysis, four key events were identified: (1) genotoxicity, (2) DNA adduct formation, (3) mutations in reporter genes, and (4) mutations in cancer genes. While the dose-response for adduct formation (a biomarker of exposure) may be linear and extrapolate to zero, the dose-response for mutations (a biomarker of effect) merges with the background or spontaneous mutation rate at some non-zero level of exposure (Fig. 2). The practical implication of this analysis is that genotoxic chemicals are expected to have a threshold equivalent to the background mutation rate and that an understanding of the machinery involved in the removal and fixation of DNA adducts is necessary to ferret out the mechanistic underpinnings of the carcinogenic process and accurately describe the dose-response behavior.

This discussion of the low dose behavior of genotoxic carcinogens (linear extrapolation *versus* threshold) is not simply to revisit arguments about thresholds (Keshava et al., 2008; Lutz and Gaylor, 2008; Starr and Swenberg, 2008; Swenberg and Starr, 2008), but to point out that the question of thresholds for mutagenic compounds is now technically within our grasp. Bringing sophisticated new technical approaches to bear on dose response of precursor steps and mutagenicity will resolve the debate. The lesson for the future of toxicity testing is to design an improved panel of "genotoxicity" tests to measure the effects of chemicals on the molecular pathways involved in the key events associated with induction, recognition, and repair of both endogenously and exogenously derived DNA adducts and their fixation as mutations.

Alternative modes-of-action for carcinogens
Non-genotoxic carcinogens may act through a variety of normal cellular pathways that are inappropriately modulated to give rise to excessive mutations. For example, highly reactive free radicals, such as reactive oxygen species, nitric oxide, and lipid peroxidation by-products, have the potential to interact with, and alter, DNA (Williams and Jeffrey, 2000); cell proliferation can be stimulated as a reparative process, resulting in mutations that are fixed within the genome due to a failure to repair in the setting of more rapid replication (Butterworth et al., 1992). Because non-genotoxic carcinogens act through the modulation of normal cellular pathways, a level of exposure is required that is sufficient to perturb normal homeostatic mechanisms, resulting in a threshold for the induction of excess mutations. Since normal cellular pathways are altered by non-genotoxic carcinogens, assays that measure dose response in a diverse suite of toxicity pathways should indicate relative sensitivity of particular pathways – non-genotoxic carcinogens would activate alternative pathways at doses lower than those pathways affecting DNA-reactivity or mutagenicity.
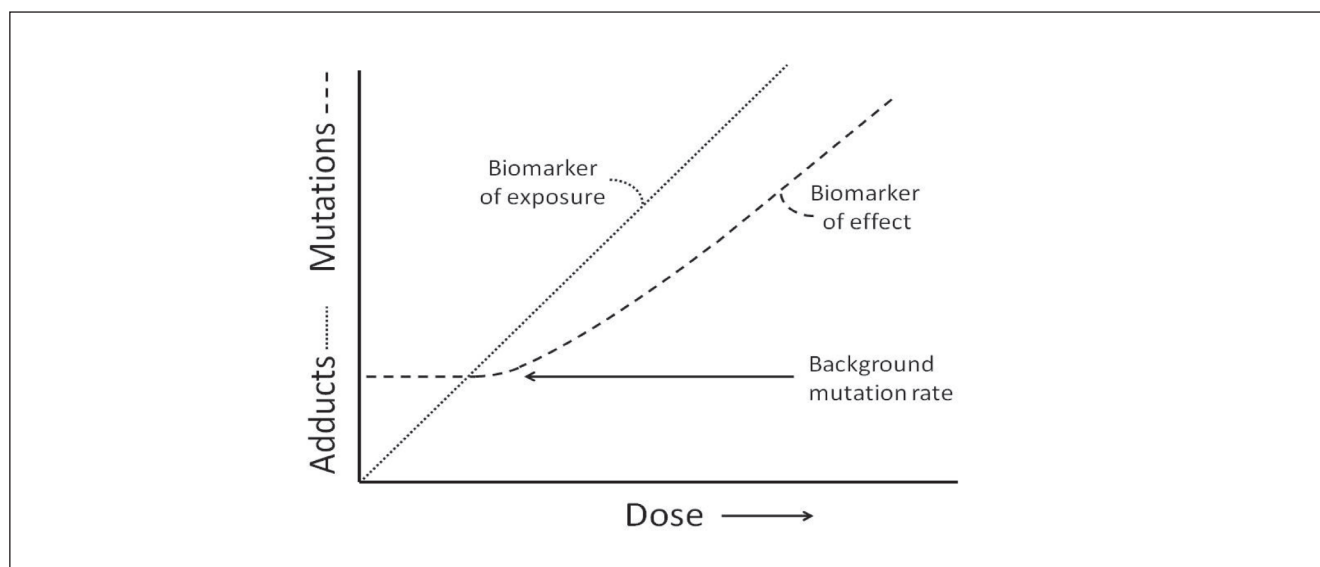
Fig. 2: Biomarkers of Carcinogenicity.
While adducts (biomarkers of exposure) linearly extrapolate to zero, mutations (biomarkers of effect) have a threshold equivalent to the spontaneous background mutation rate.

Toxicogenomics approaches have been proposed as tools to improve the specificity of carcinogenicity assays (Aubrecht and Caba, 2005; Ellinger-Ziegelbauer et al., 2009). The advent of these data-dense and powerful toxicogenomics assays has markedly increased the ability to identify alterations in stress-response pathways after toxicant exposure. With prototypal genotoxicants, such as ionizing radiation, these microarray-based approaches yield sensitive dose-dependent insight into the activation of numerous pathways, including the p53 pathway and mitogen-activated protein kinase pathways (Snyder and Morgan, 2004). Toxicogenomics should allow the investigation of the underlying mechanisms responsible for genotoxicity. Support for this concept comes from 13-week *in vivo* carcinogenicity studies examining the lung response to compounds identified as carcinogens through the rodent 2-year bioassay (Thomas et al., 2009). These short-term assays relied on toxicogenomics signals and had an estimated predictive accuracy of about 80%.

Toxicity pathway approaches and DNA-reactive compounds
Before outlining the framework to define adversity based on a toxicity pathway-based approach to testing DNA-reactive compounds, it is necessary to have a clear focus on the expected application of test results for human health risk assessment. We would emphasize that the goal of toxicity testing is not to predict high dose outcomes in rodents but to accurately predict the likely risks of low dose exposures in human populations. With respect to mechanistic dose response behaviors, it is also necessary to assess the conditions under which perturbations of DNA-structure are likely to become fixed in the genome and provide altered cells the potential for autonomous growth. Identifying this transition from interactions that are adequately controlled by DNA-repair processes to such extensive alteration that repair is incomplete is the key step in assessing "adversity" at the cellular level for DNA-reactive compounds in the new testing paradigm.

3 A systems framework is needed to define adversity at the cellular level

DNA-reactive, mutagenic compounds and new testing paradigms
The approach for using *in vivo* studies for cancer risk assessment (Fig. 1) can be set in contrast to the likely use of test results with implementation of an *in vitro* mode-of-action based toxicity evaluation scheme (Fig. 3). These new *in vitro* cellular methods will provide increasingly more useful *in silico* tools as high throughput databases become available for QSAR modeling of toxicity pathway assay results. After *in silico* assessment of the potential of activating various toxicity pathways, a suite of pathway assays (here designated from assay 1 through assay n) would provide a catalog of likely modes of action and of the relative ability of compounds to activate particular assays based on potency (by providing an effective concentration causing 50 or 10% maximum responses, i.e., a EC50 or EC10). Some subset of the pathway assays would test for DNA-damage, DNA-damage associated genomic responses in cellular systems, and mutagenicity (tests a, b, c, and d in Fig. 3). Compounds that had lower EC50 in these DNA-reactivity oriented assays than in the remainder of the assays would be considered likely to have mutagenic and carcinogenic potential at higher doses. For cases where mutagenicity occurred at higher EC50s than toxicity or stress pathway activation, the conclusion would be that effects at higher doses could have a mutagenic mode of action secondary to primary modes-of-action related to the more sensitive pathways. However, this enumeration of activities across a suite of tests and the doses at which they produced pathway perturbations would not of itself define "adversity" at a cellular or molecular level. This aggregation of EC50s would provide a more qualitative assessment of likely apical outcomes and a relative measure of *in vitro* potency. Conclusions about adversity, however, would require a more quantitative integration of

in silico methods	*in vitro* assays		DNA damage mutagenicity Assays (a-d)	Qualitative Conclusion
Estimating	**1**	**a**	low EC50	Compound likely to
likelihood of	-	**b**	-	have mutagenic
toxicity,	-	**c**	-	mode of action and
metabolism,	-	**d**	-	likely to be a
cancer, etc.	-	-	-	carcinogen at high
	-	-	-	doses – i.e., with
	n	n	-	"high degree of
			highest EC50	perturbation"

Fig. 3: Using results from multiple assays to define adversity.
One possibility for using the *in vitro* tests in risk assessment is to focus on a panel of pathway-relevant assays and the concentrations at which the various pathways are activated. The adverse level is defined more as an algebraic combination of pathway responses than as aspects of severity of perturbations. The panel would provide output at different levels of perturbation and likely different levels of biological organization in the *in vitro* assays.

results of pathway assays examining different aspects of DNA-damage and repair.

Pathway assay design & degree and severity of perturbations
Toxicologists are familiar with *in vitro* test batteries meant to predict specific target organ toxicity. The goal of mode-of-action, human biology-based testing is not to generate batteries of tests to provide a prediction of animal toxicity test results for various endpoints. Instead, these methods are intended to determine regions of exposure that will not cause any adverse responses in exposed human populations. These panels of assays, evaluating specific pathway targets – such as DNA damage and repair – would be designed to capture increasing degrees of perturbations. In this example, we propose several assays for increasing severity; however, in practice, the choice of specific assays needs to have significantly more input than from the two current authors – neither of whom is an expert in mutational assays. The *in vitro* assays for DNA-damage repair might be micronuclei formation, comet assays in treated cells, p53/H2AX binding to DNA, dose-dependent genomic changes in target cells, and mutation in various test systems. Each test in the panel would be designed to provide dose response over broad ranges of treatment and to assess some aspect of pathway perturbation for DNA-damage, DNA-repair, and mutation. The organization of these perturbations across levels of biological

response could be utilized in a failure model structure, as in Fig. 4 (Boekelheide and Campion, 2010) or for use in a more quantitative process similar in structure to genetic progression and waiting time models (Beerenwinkel et al., 2007). In either of these approaches, the degree of perturbation and adversity would be provided by a composite probability associated with contributions from each of the stages with increasing degrees of perturbation as defined by dose response characterization of the sequential pathway assays. The overall goal in such an aggregate analysis is to estimate, with some degree of confidence, regions of exposure that will not increase mutation frequency, i.e., regions of exposure in which DNA-repair will be adequate to deal with any DNA-damage.

Pathway Assay Validation
What does it mean to have a validated assay? One commonly stressed point is that the assay provides consistent output for compounds with known modes of action. For instance, we should see mutagenic responses in a mutagenicity pathway assay for compounds known to cause mutations. Once again, to a large extent this represents a form of qualitative validation. To fully implement mode-of-action based approaches in a quantitative manner, validation of the pathway assays will increasingly demand that we know the circuit structure of the toxicity pathway (Alon, 2006, 2007a, b) and have the ability
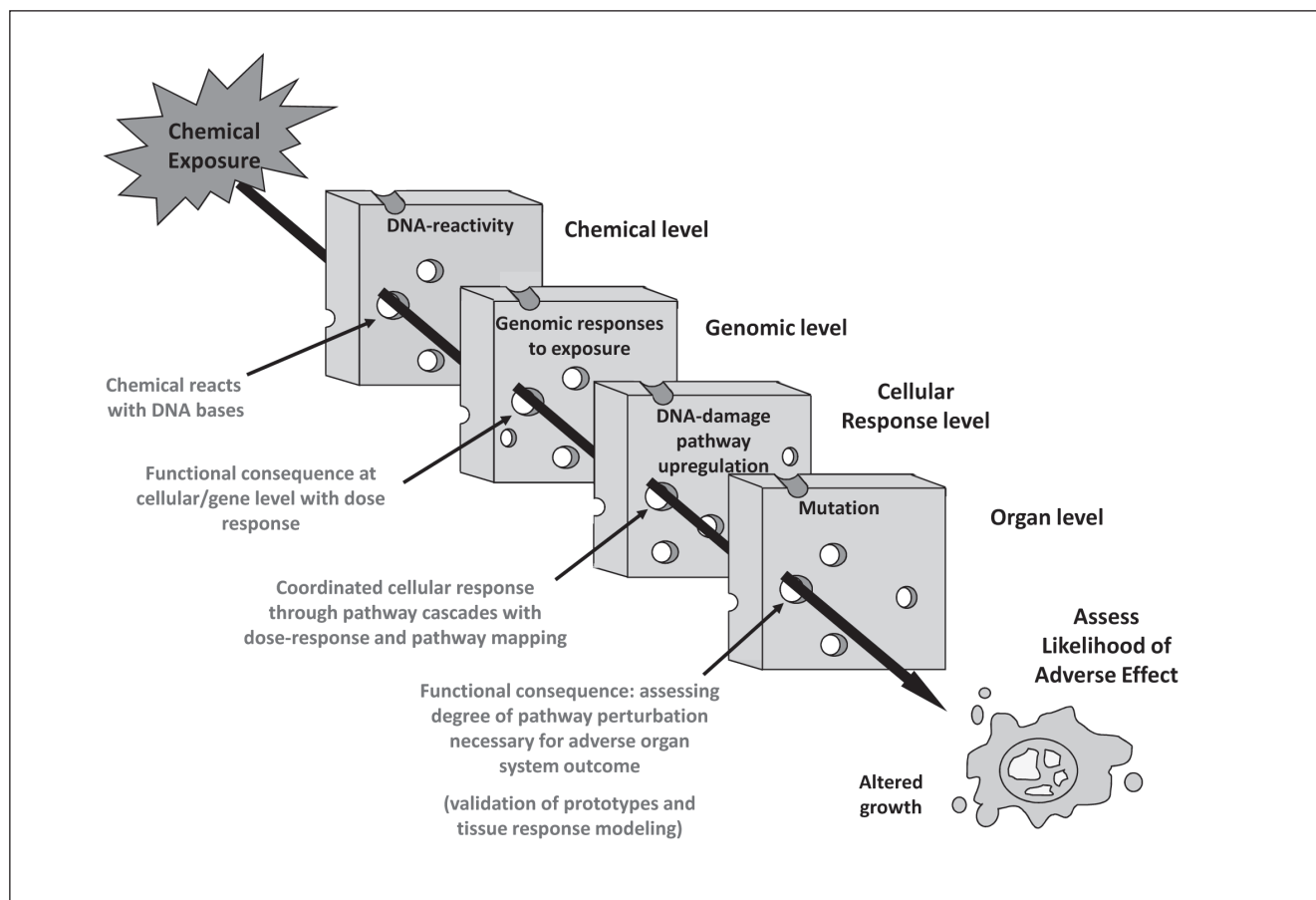

Fig. 4: A Swiss cheese model for adversity for mutagenicity/carcinogenicity.

to model the pathway dynamics (Aldridge et al., 2006) in response to varying degrees of perturbation by treatment with a chemical stressor. These computational systems biology models of the toxicity pathway assays themselves would then describe the dose range of stressors that are expected to be without any consequence on the biological system, the ranges in which damage is sensed with some degree of adaptation, and the ranges where the system has been excessively perturbed. These model structures could give a quantitative, mechanistic evaluation of the conditions leading to excessive pathway perturbation. Such systems biology models would also allow better assessment of susceptibility factors that might need to be considered within a population by identifying key signaling nodes in pathways and how these nodes vary across a population. In addition, these computational systems biology models are key in aiding interpretation of *in vivo* dose response models that also could be improved beyond the process of using mode-of-action studies to simply decide if the process would be linear low dose or based on a threshold calculation (Fig. 1).

DNA-repair pathways are under intense scrutiny to produce models of damage and repair. The dynamics of mdm2-p53 signaling is an area of keen interest due to the importance of this pathway for genomic integrity and cancer (Batchelor et al., 2009). An mdm2 dimer negatively regulates p53. Activation of various proteins is controlled by p53 as well as upregulation of mdm2 to limit the period of activity of p53 responses to damage (Wade et al., 2010). The dynamics display pulsatile, single cell behavior to regulate multiple pulses in the face of ongoing DNA-damage (Lahav et al., 2004; Ma et al., 2005). Recent studies have described that the basal level of activity of the DNA-stress pathway is controlled by post-translational modulation of p53 (Loewer et al., 2010). The partnering molecules and dynamics of the kinase cascade transducers are known, but their quantitative behaviors are only now being resolved. These stress responsive pathways, despite the variety of partners in the control of specific stressors, have similar architecture (Simmons et al., 2009) and appear to be designed to function as negative feedback controllers with high-loop gain to insure robustness in the anti-stress responses (Zhang and Andersen, 2007). With increasing resolution regarding pathway function, "adversity" becomes associated with those regions of exposure that lead to excessive activation of pathways, including an association of pathway activation with an adverse outcome in the case of DNA-reactive compounds with mutation. The output of the analysis across the panel of DNA-damage, genomic responses, and mutagenicity could be, for example, an *in vitro* EC10 that represents a level of perturbation likely to lead to adverse responses in an intact organism exposed to these concentrations for prolonged periods of time.
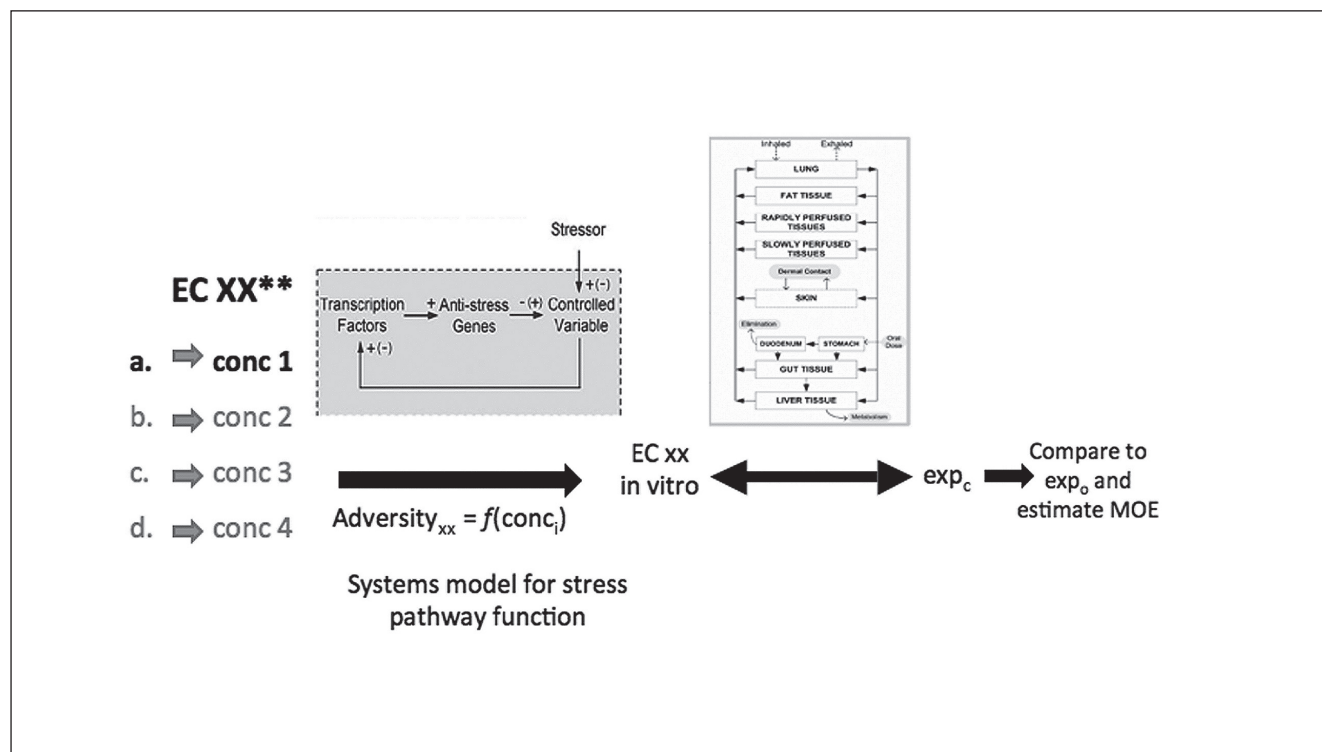
Fig. 5: *In vitro-in vivo* extrapolations require estimate of "adverse" concentrations from the *in vitro* assays, computational systems models of pathway function, and the use of reverse dosimetry to estimate the expected human exposures required to produce tissue concentrations equal to those causing adverse responses in the *in vitro* assay systems.

Here, EC XX= effective concentration for a xx % response; exp_c = predicted human exposure (mg/kg/day) giving rise to $conc_i$ *in vivo*; exp_o = observed exposure in human populations; MOE = margin of exposure from results of testing with toxicity pathway "y". The figure to the left shows a generic negative feedback controller for oxidative stress pathway function (Zhang and Andersen, 2007).

## 4 Applying a redefined concept of adversity to risk assessment

The *in vitro* panels and computational systems modeling of pathway structure and function provide an estimate of a concentration that would be considered to lead to excessive perturbation of the toxicity pathway. The aggregation of effective concentrations in pathway assays with adjustments required accounting for polymorphisms in the pathway and other pharmacodynamic aspects of pathway function would comprise a point-of-departure for the risk assessment. The low dose extrapolation would proceed based on the pathway circuitry and the dose response for the activation of the key circuit involved in pathway activation.

The process of setting acceptable exposure levels requires the ability to estimate human exposures that would give rise to tissue or blood (or plasma) concentrations equal to those concentrations found to be "adverse" *in vitro* in the tests with human cells (Fig. 5). These calculations require contributions from disciplines of pharmacokinetic modeling and, more specifically, physiologically based pharmacokinetic modeling. The process of reverse dosimetry (Tan et al., 2007; Clewell et al., 2008) provides tools to estimate ranges of exposure expected to give rise to specific plasma or tissue concentrations. The overall process for moving from perturbations of a suite of pathways of toxicity to a health risk assessment (Fig. 6) utilizes these various components: targeted assays, computational systems biology dose-response models for the pathways, and

in vitro-in vivo extrapolations to convert active concentrations *in vitro* to human exposures expected to give similar concentrations in exposed individuals.

Implementing the new approaches to adversity
These new approaches, based on mode-of-action human biology, and a better appreciation of biological responses to stressors and pathway perturbation, promise improved throughput and reduced uncertainty in estimating concentrations causing various degrees of perturbation. Nonetheless, they will naturally be compared, in one way or another, with current cancer bioassay approaches to determine if the new methods provide adequate safety for exposed human populations. The comparison of past practices and these more mode-of-action based approaches will be difficult, but, in our opinion, they can be made by using well-studied prototype compounds whose toxicity has already been examined with *in vivo* and *in vitro* assays. These prototype compounds can support case studies for proof of concept exercises for a mechanistic, pathway of toxicity based approach to human health risk assessment. The choice of the prototypes would also depend on choosing examples that are expected to activate specific stress pathways, e.g., DNA-damage, endoplasmic-reticulum stress, or oxidative stress (Simmons et al., 2009), or that have specific receptor-mediated targets, such as the estrogen receptor (ER), the arylhydrocarbon receptor (AhR), and the peroxisomal proliferator activating receptors (PPARs). These case studies could be more broadly useful in assessing how mode-of-action results should be used to inform risk assessment.
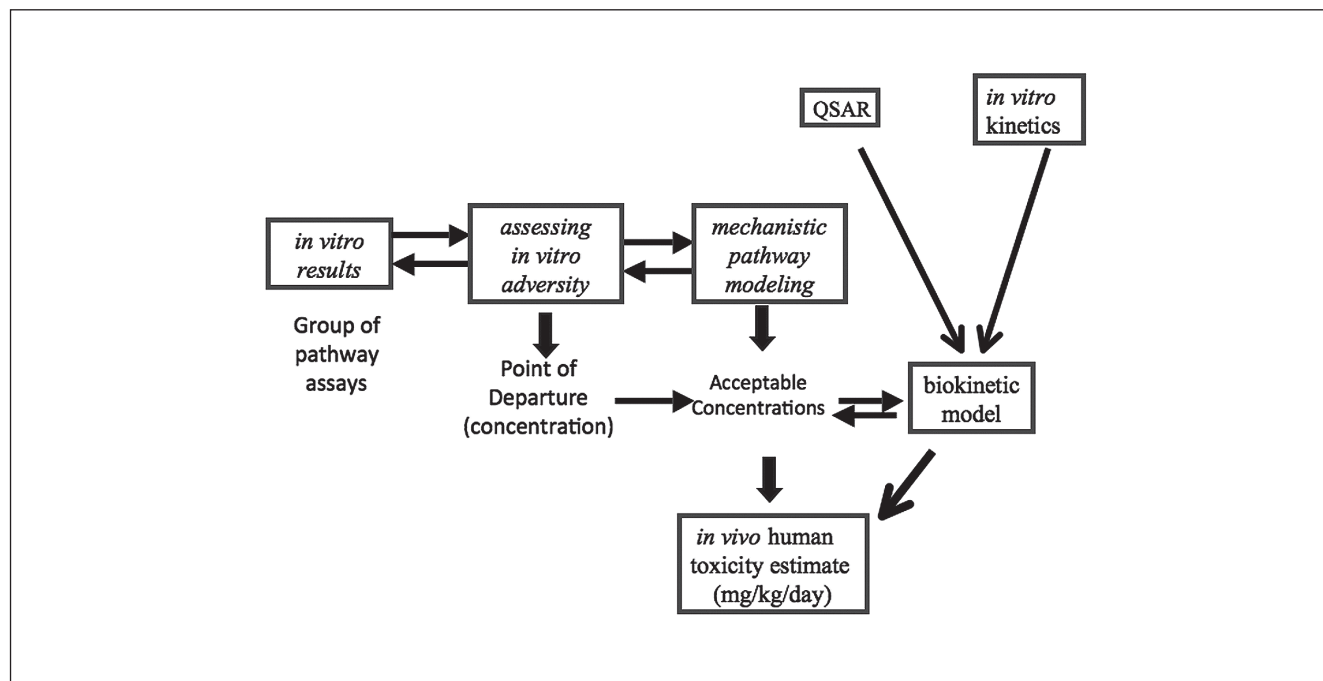
Fig. 6: Organizing information from mechanistic evaluation of pathways of toxicity for human health risk/safety assessment.
The components, similar to those defined in *Toxicity Testing in the 21st Century: A Vision and A Strategy* (NRC, 2007) would need to be integrated to develop recommended exposures for human populations (e.g., mg/kg/day) based on active concentrations in the *in vitro* assays.

Another value of these prototypes/case studies will be the demonstration of the new toxicity testing paradigm in practice – showing how data will be collected, organized, and interpreted to make decisions about regions of exposure that should be without appreciable risk in human populations. Over the past several years, genomic studies of formaldehyde, a natural metabolite found in all cells and tissues, have illuminated dose dependent transitions in mode of action of this compound with exposures in the intact rat (Thomas et al., 2007; Andersen et al., 2008; Andersen, 2010). These exposures range from those causing no net change in tissue formaldehyde compared to background, to levels affecting mucosal formaldehyde and leading to irritancy, to overtly toxic levels leading to large scale changes in cell cycle control and DNA-damage pathways. Formaldehyde would be a useful prototype for the application of these new *in vitro* methods for defining adversity and comparing *in vitro* and *in vivo* mode of action studies.

References

Aldridge, B. B., Burke, J. M., Lauffenburger, D. A. and Sorger, P. K. (2006). Physicochemical modelling of cell signalling pathways. *Nature Cell Biology 8*, 1195-1203.

Alon, U. (2007a). Network motifs: theory and experimental approaches. *Nature Reviews Genetics 8*, 450-461.

Alon, U. (2007b). Simplicity in biology. *Nature 446*, 497-497.

Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Boca Raton, FL: Chapman & Hall/CRC.

Ames, B. N. (1973). Carcinogens are mutagens: their detection and classification. *Environ. Health Perspect. 6*, 115-118.

Andersen, M. E., Clewell, H. J., Bermudez, E. et al. (2010). Formaldehyde: integrating dosimetry, cytotoxicity and genomics to understand dose-dependent transitions with an important endogenous compound. *Toxicological Sciences*, in press.

Andersen, M. E., Clewell, H. J., Bermudez, E. et al. (2008). Genomic signatures and dose-dependent transitions in nasal epithelial responses to inhaled formaldehyde in the rat. *Toxicological Sciences 105*, 368-383.

Aubrecht, J. and Caba, E. (2005). Gene expression profile analysis: an emerging approach to investigate mechanisms of genotoxicity. *Pharmacogenomics 6*, 419-428.

Batchelor, E., Loewer, A. and Lahav, G. (2009). The ups and downs of p53: understanding protein dynamics in single cells. *Nature Reviews Cancer 9*, 371-377.

Beerenwinkel, N., Antal, T., Dingli, D. et al. (2007). Genetic progression and the waiting time to cancer. *Plos Computational Biology 3*, 2239-2246.

Boekelheide, K., and Campion, S. N. (2010). Toxicity Testing in the 21st Century: using the new toxicity testing paradigm to create a taxonomy of adverse effects. *Toxicological Sciences 114*, 20-24.

BusinessDictionary.com (2010). "adverse effect". In *BusinessDictionary.com*.

Butterworth, B. E., Popp, J. A., Conolly, R. B. and Goldsworthy, T. L. (1992). Chemically induced cell proliferation in carcinogenesis. *IARC Sci. Publ.,* 279-305.

Clewell, H. J., Tan, Y. M., Campbell, J. L. and Andersen, M. E. (2008). Quantitative interpretation of human biomonitoring data. *Toxicology and Applied Pharmacology 231*, 122-133.

Ellinger-Ziegelbauer, H., Aubrecht, J., Kleinjans, J. C. and Ahr, H. J. (2009). Application of toxicogenomics to study mechanisms of genotoxicity and carcinogenicity. *Toxicol. Lett. 186*, 36-44.

Ennever, F. K. and Lave, L. B. (2003). Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. *Regul. Toxicol. Pharmacol. 38*, 52-57.

Gaylor, D. W. (2005). Are tumor incidence rates from chronic bioassays telling us what we need to know about carcinogens? *Regul. Toxicol. Pharmacol. 41*, 128-133.

Huff, J., Jacobson, M. F. and Davis, D. L. (2008). The limits of two-year bioassay exposure regimens for identifying chemical carcinogens. *Environ. Health Perspect. 116*, 1439-1442.

Keshava, N., Chao, C., White, P. et al. (2008). Biomarkers in toxicology and risk assessment: informing critical dose-response relationships by Swenberg et al., 2008. *Chem. Res. Toxicol. 22*, 8-10.

Kirkland, D., Pfuhler, S., Tweats, D. et al. (2007). How to reduce false positive results when undertaking in vitro genotoxicity testing and thus avoid unnecessary follow-up animal tests: Report of an ECVAM Workshop. *Mutat. Res. 628*, 31-55.

Kirkland, D., Aardema, M., Henderson, L. and Muller, L. (2005). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and noncarcinogens I. Sensitivity, specificity and relative predictivity. *Mutat. Res. 584*, 1-256.

Kirkland, D., Aardema, M., Muller, L. and Makoto, H. (2006). Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens II. Further analysis of mammalian cell results, relative predictivity and tumour profiles. *Mutat. Res. 608*, 29-42.

Lahav, G., Rosenfeld, N., Sigal, A. et al. (2004). Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet. 36*, 147-150.

Lewis, R. W., Billington, R., Debryune, E. et al. (2002). Recognition of adverse and nonadverse effects in toxicity studies. *Toxicol. Pathol. 30*, 66-74.

Loewer, A., Batchelor, E., Gaglia, G. and Lahav, G. (2010). Basal dynamics of p53 reveal transcriptionally attenuated pulses in cycling cells. *Cell 142*, 89-100.

Lutz, W. K. and Gaylor, D. W. (2008). Dose-response relationships for cancer incidence reflect susceptibility distributions. *Chem. Res. Toxicol. 21*, 971-972.

Ma, L., Wagner, J., Rice, J. J. et al. (2005). A plausible model for the digital response of p53 to DNA damage. *Proc. Natl. Acad. Sci. USA 102*, 14266-14271.

NCR (2007). Toxicity testing in the 21st century: a vision and a strategy. Washington, DC: The National Academies Press.

Simmons, S. O., Fan, C. Y. and Ramabhadran, R. (2009). cellular stress response pathway system as a sentinel ensemble

in toxicological screening. *Toxicological Sciences 111*, 202-225.

Snyder, A. R. and Morgan, W. F. (2004). Gene expression profiling after irradiation: clues to understanding acute and persistent responses? *Cancer Metastasis Rev. 23*, 259-268.

Stansell, K., Marvelli, M. and Wiener, J. (2005). "*Adverse effects" and similar terms in U.S. Law*. Durham, NC: Duke Center for Environmental Studies.

Starr, T. and Swenberg, J. A. (2008). Letter to the Editor. *Chem. Res Toxicol. 21*, 972-973.

Swenberg, J. A., Fryar-Tita, E., Jeong, Y. C. et al. (2008). Biomarkers in toxicology and risk assessment: informing critical dose-response relationships. *Chem. Res. Toxicol. 21*, 253-265.

Swenberg, J. A. and Starr, T. (2008). Response to Keshava et al. (2008). *Chem. Res. Toxicol. 22*, 10-12.

Tan, Y. M., Liao, K. H. and Clewell, H. J. (2007). Reverse dosimetry: interpreting trihalomethanes biomonitoring data using physiologically based pharmacokinetic modeling. *Journal of Exposure Science and Environmental Epidemiology 17*, 591-603.

Thomas, R. S., Bao, W., Chu, T. M. et al. (2009). Use of short-term transcriptional profiles to assess the long-term cancer-related safety of environmental and industrial chemicals. *Toxicol. Sci. 112*, 311-321.

Thomas, R. S., Allen, B. C., Nong, A. et al. (2007). A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. *Toxicol. Sci. 98*, 240-248.

Wade, M., Wang, Y. V. and Wahl, G. M. (2010). The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends in Cell Biology 20*, 299-309.

Williams, G. M. (2001). Mechanisms of chemical carcinogenesis and application to human cancer risk assessment. *Toxicology 166*, 3-10.

Williams, G. M. and Jeffrey, A. M. (2000). Oxidative DNA damage: endogenous and chemically induced. *Regul. Toxicol. Pharmacol. 32*, 283-292.

Zhang, Q. and Andersen, M. E. (2007). Dose response relationship in anti-stress gene regulatory networks. *Plos Computational Biology 3*, 345-363.

Correspondence to

Kim Boekelheide, MD, PhD
Department of Pathology and Laboratory Medicine
Brown University
Box G-E5
Providence, RI 02901, USA
Tel: +1 401 863 1783
Fax: +1 401 863 9008
e-mail: kim_boekelheide@brown.edu