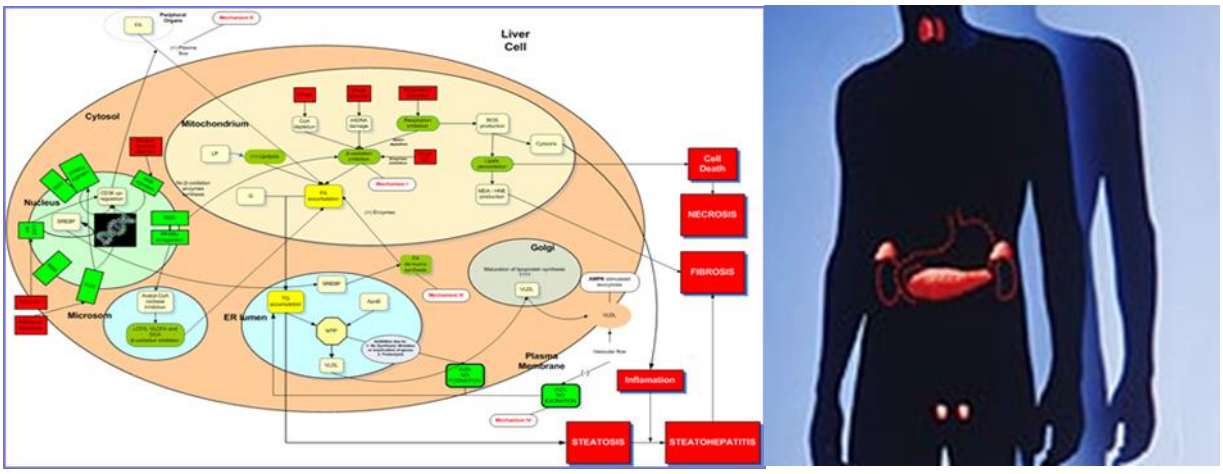
European Commission

JRC SCIENTIFIC AND POLICY REPORTS

Toxicity Pathways – from concepts to application in chemical safety assessment

Maurice Whelan
Melvin Andersen

2013



Report EUR 26389 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Health and Consumer Protection

Contact information
Maurice Whelan
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 202, 21027 Ispra (VA), Italy

http://ihcp.jrc.ec.europa.eu/

This publication is a Scientific and Policy Report by the Joint Research Centre of the European Commission.

JRC 86467

EUR 26389 EN

ISBN 978-92-79-34910-2 (pdf)
ISBN 978-92-79-34911-9 (print)

ISSN 1018-5593 (print)
ISSN 1831-9424 (online)

doi: 10.2788/49626

Luxembourg: Publications Office of the European Union, 2013

Printed in Italy

Front page images: Schematic of toxicity pathways related to steatosis taken from Landesmann et. al., JRC Scientific and Policy Report, Nr. JRC75689 (2012). doi:10.2788/71112 [Public domain] and Human endocrine system, via Wikimedia Commons.

Toxicity Pathways – from concepts to application in chemical safety assessment

Maurice Whelan and Melvin Andersen

Acknowledgements

The invited experts that participated in the "Toxicity Pathways Workshop" (September 2011) provided the basic content on which this report is based. They were as follows; Leonidas Alexopoulos (National Technical University of Athens), Gordana Apic (Cambridge Cellnetworks Ltd.), Kim Boekelheide (Brown University), Edward Carney (The Dow Chemical Company), Mark Cronin (Liverpool John Moore's University), David Gerhold (NIH Chemical Genomics Center), Thomas Hartung (Johns Hopkins University), Miriam Jacobs (EFSA), Richard Judson (US EPA), Robert Landsiedel (BASF), Avi Maayan (Mount Sinai School of Medicine), Bette Meek (University of Ottawa), Pierre Moulin (Novartis), Kathleen Plotzke (Dow Corning Corporation), Julio Saez-Rodriguez (EMBL-EBI), Michael Schwarz (University of Tübingen), Russell Thomas (THI), and José-Manuel Zaldívar (JRC).

The deliberations of workshop participants were recorded by colleagues from the JRC, namely, Elisabet Berggren, Sharon Munn and Andrew Worth.

The Workshop organisers gratefully acknowledge the financial support provided by the American Chemistry Council's Long-Range Research Initiative (ACC LRI) to participants travelling from North American, in addition to the support provided by the JRC to EU participants and to host the event.

CONTENTS

Introduction

Few would deny that the NRC report (NRC, 2007), "Toxicity Testing in the 21st Century: A Vision and Strategy", represented a re-orientation of thinking surrounding the risk assessment of environmental chemicals. The key take-home message was that by understanding Toxicity Pathways (TP) we could profile the potential hazard and assess risks to humans and the environment using intelligent combinations of computational and in vitro methods. In theory at least, shifting to this new paradigm promises more efficient, comprehensive and cost effective testing strategies for every chemical in commerce while minimising the use of animals. For those of us who embrace the vision and the strategy proposed to achieve it, attention has increasingly focused on how we can actually practice what we preach. For a start, 21st century concepts described in the report have to be carefully interpreted and then translated into processes that essentially define and operationalize a TP framework for chemical risk assessment.

In September 2011 the European Commission's Joint Research Centre (JRC) and the Hamner Institutes for Health Sciences co-organised a "Toxicity Pathways" workshop. It was hosted by the JRC and took place in Ispra, Italy. There were 23 invited participants with more or less equal representation from Europe and North America. The purpose of the meeting was to address three key questions surrounding a TP based approach to chemical risk assessment, namely – *What constitutes a TP? How can we use TPs to develop in vitro assays and testing strategies?* And, *How can the results from TP testing be used in human health risk assessments?* The meeting ran over two days and comprised a series of thought-starter presentations, breakout sessions and plenty of group discussions. The outcome was captured by rapporteurs and compiled as a workshop report which is available for download (without charge) from the JRC website. Here we expand on selected deliberations of the workshop to illustrate how TP thinking is still evolving and to indicate what pieces of the puzzle still need to fall into place before TP based risk assessment can become a reality.

Defining a Toxicity Pathway

The NRC panel originally coined the term "Toxicity Pathway" and defined it as a cellular response pathway that would result in an adverse health effect when sufficiently perturbed). (NRC, 2007). This definition basically infers that a TP possesses three key attributes. First,

the underlying biology of a TP resides in the molecular and cellular domains. Second, a TP is essentially a cellular mode-of-failure i.e. a description of the particular conditions and sequential steps involved when an external force causes a cellular sub-system to malfunction and fail (Boekelheide and Andersen, 2010; Boekelheide and Campion, 2010). And, third, the mode-of-failure captured by a TP which is manifest as a cellular phenotype will result in adverse health effects at the organism level under certain specified conditions. Fulfilling these attributes is obviously not trivial and it is not a surprise therefore that what is often presented as a TP is not actually a TP. For example, identifying just one step along the pathway such as receptor-binding or apoptosis, or omitting an explanation of the concentration and time dependant dynamic relationship between two sequential steps, or not being specific about the conditions under which a TP will actually lead to adverse effects at the organ and organism level simply falls short of the bar.

Although one could reasonably argue that the NRC report gave a clear definition of a TP, there was however a relatively diverse landscape of opinion between workshop participants. This fact was highlighted in an exercise where everyone was asked to write their own short definition, all of which are included as the annex A. Although the definitions collected were quite different, an attempt was made to agree on one common working definition to help maintain a common thread during the course of the meeting. The working definition proposed was, *"A Toxicity Pathway is a sequence of intracellular events which regulate normal biology which, when sufficiently perturbed by a xenobiotic, leads to an adverse outcome at the level of the cell, and possibly the whole organism"*. On the face of it this definition seemed very much in line with the NRC report but as the workshop progressed the devil in the detail became apparent and made it evident that even experts in the field can interpret the same definition and terms in a very different way.

Toxicity Pathways in relation to Adverse Outcome Pathways and Mode-of-Action

Toxicity Pathways, Adverse Outcome Pathways (AOP) and Mode-of-Action (MoA) all reflect concepts that represent a shift towards a knowledge-driven approach to toxicological hazard and risk assessment (Sonich-Mullin *et al.*, 2001; Boobis *et al.*, 2006; Boobis *et al.*, 2008; Julien *et al.*, 2009; Ankley *et al.*, 2010). They capture and formally describe mechanistic understanding of biological systems and the various ways that toxicants can interfere with them to cause adverse effects. Although they originated in different

communities and contexts, a TP, an AOP and a MoA all describe a causal chain of biochemical and biological events, starting typically from an initiating molecule event that leads to an adverse effect. An MOA and an AOP are more expansive in that they include events at higher levels of biological organisation, at least to the level of whole organism for MOA, and possibly to the population level for AOP. In the context of toxicity assessment, decisions supported by a TP based paradigm would rely on effect information at molecular and cellular levels, whereas MOA or AOP based approaches generally include consideration of effects at the level of the tissue, organ and organism. In practice, these various concepts are interrelated. TP knowledge comes from multiple sources, including experience with testing in intact animals and population-level responses. In turn, the understanding of causal relationships within MOAs and AOPs are important considerations in defining the TPs that need examination in any test battery.

Challenging a knowledge-based paradigm however, it was argued that a mechanistic description of the underlying biological system and the manner in which pathway perturbation leads to adversity might not be actual prerequisites to carry out reliable risk assessment using, for example, in vitro data. The NRC report itself stopped short of explicitly stating the need for a mechanistic explanation of the perturbation and events that lead to the adverse cellular effect. The array of 'omics tools now available for measuring for example gene expression, protein interactions and metabolite flux provide experimental platforms for surveying cellular response with exceptional resolution and coverage (Rusyn and Daston, 2010). The quest for many is to search these massive datasets for signatures comprising collective variations in tens, hundreds and possibly thousands of related biomarkers that associate with significant phenotypic changes within a cell. Thus, 'classifiers' – algorithms or prediction models constructed for example using statistical or machine learning approaches - correlate characteristic patterns in signalling or expression 'space' with specific adverse effects .

Pattern-based prediction can also extend to identifying the trajectory or rate/direction of change in a biomarker pattern rather than using an individual pattern related to a particular set of experimental conditions. Thus information on the toxicity of a chemical lies more in how the cell actually responds to an insult (i.e. which direction it takes and how fast it sets off) rather than what state it ends up in. In any case, if mechanistic understanding takes a backseat and becomes a 'nice to have' rather than a 'need to have', and massively high-content

measurements become feasible and affordable, then one might want to embrace another definition of a TP, namely, *"Toxicity Pathways are alterations of signalling motifs, such as altered gene expression, that are predictive of toxicity".* This is in tune with more of a purely data-driven paradigm rather than a knowledge-driven one, but nonetheless it can prove very effective in identifying and classifying a potential toxicant without relying on any real mechanistic understanding up front to design the experiment.

Toxicokinetics and Toxicodynamics

Activation of a TP is dose dependent. Below some critical exposure levels, an agent may modulate a normal signalling or metabolic pathway but due to the ability of the system to adapt and compensate, minor modulation will not result in an adverse effect. However, adverse consequences will occur with perturbations beyond certain limits (Krewski *et al.*, 2011). Methods to estimate or calculate these dose-related limits or thresholds are necessary for using data derived from pathway-based tests for risk assessment. Purely qualitative descriptions of TPs will serve more for hazard identification, grouping and read-across. Besides the dose level, other contextual factors contribute to assessing the relevance and activation conditions of a TP. These factors include species, gender, life-stage, critical times of exposure, genetic predisposition. The description of a TP needs to refer to the biological region of applicability. In other words, in vitro pathway assays need to be carefully designed based on knowledge of the biology and the distribution and relevance of specific pathways in particular tissues/cell-types.

A TP describes a biological pathway and the toxicodynamic processes that may lead to an adverse effect upon alteration of the activity of the TP. The toxicity of an agent arises through activation or inhibition of specific pathways. The toxicokinetic behaviour of the agent is also important because the absorption, distribution, metabolism and excretion dictate what particular pathways contribute for specific exposure conditions. We should consider that toxicokinetics is agent-specific, whereas a TP has biological specificity. In principle, the description of a TP need not include direct reference to specific agents that affect this pathway. On the other hand, when endeavouring to characterise the toxicological hazard of a particular agent both the toxicokinetic and toxicodynamic aspects specific to that agent and the pathway become relevant.

Capturing knowledge on Toxicity Pathways

Surveying the literature shows that diverse methods and approaches are available to describe MOA and TPs. Although the existing knowledge is extensive, the heterogeneity in its organization and curation makes it difficult to systematically and effectively identify, categorise and map TPs. To make progress in building a comprehensive ontology and knowledge base of TPs, contributors will need to define and adopt suitable pathway-specific terminology and nomenclature, leading eventually to development of a consistent methodology in describing and presenting pathways. Recently, OECD has issued guidance for describing and evaluating AOPs (OECD, 2013), together with a glossary of frequently used terms and definitions associated with pathway concepts With further maturation, this type of guidance could provide the basis for a harmonised approach to pathway categorisation and mapping across a broad community of contributors. Emergence of purpose-built IT systems such as Effectopedia [see http://effectopedia.org/] and "wiki" platforms for capturing and exchanging pathway descriptions could facilitate a harmonised collective approach. Ultimately, better organisation of TP information will allow the exploitation of the vast amount of existing knowledge in the extant literature, facilitate the identification of knowledge gaps and stimulate the toxicology and biomedical research communities to target their research efforts to ensure progress in furthering the definition of TPs and organization of this information for risk assessment.

There was discussion of a ‘grand plan’ that would consist of a systematic mining of literature to assemble and map mechanistic MOA knowledge onto a well-structured framework, assessing normal biological pathways and their susceptibility to external perturbation (Hartung and McBride, 2011). This effort would provide candidate TPs. Once we reach a sufficient coverage of the “pathway landscape”, the process of inferring pathway networks and identifying nodes and key events (nodes) can begin in earnest in an attempt to create a rational, purpose-driven design of integrated testing and assessment systems. Any initiative of this scale will require the engagement of national and international scientific communities that go far beyond the capacity of any individual project or consortium. The solution may lie in the establishment of 'reference' goals, objectives, standards and guidance, subsequently communicated to, and accepted by, a large section of the scientific community stakeholders, united in a common cause to further contemporary risk assessment. Alternatively, or as a complimentary measure, one could exploit the somewhat competitive nature of the scientific

community. For example, a tool used in the systems biology community to capture the attention of state-of-the-art practitioners is to invite a competition open to all players. The stage is first set with a clear definition of the problem, for instance, using the in vitro assay results for risk assessment. Participants receive datasets, metadata, and a description of what constitutes success. Evaluations of proposed solutions use specific success criteria, and results made public. Such a competitive approach not only taps into the latest thinking and techniques, but also is highly educational for all concerned. One could conceive of adopting such a model not only to engage a broad section of the scientific community in a focused manner to identify and map TPs, but also to apply pathway concepts and knowledge to solve safety assessment problems more quickly.

Toxicity Pathway based assays and testing strategies

The discussions in the second session highlighted a diverse range of opinions about the detail required to make TPs useful for toxicity testing and safety assessment. There was emphasis on approaches for mining current knowledge in order to collect, organise, interpret and remap available information. This exercise would include broad screening of mechanism-based data both on MOA-based toxicity prediction methods and currently used in vivo apical toxicity test methods. With respect to archival results from apical studies, TP identification should include adverse outcomes associated with both activation and inhibition.

While molecular initiating events (MIEs) provide a starting point for looking at the range of TPs, there is a considerable biological distance between an MIE and an adverse effect. It is clear that challenge of defining adversity in relation to TPs remains a significant obstacle both in identifying TPs and in determining assays for measuring TP perturbations. For instance, is it necessary to have a linkage to adversity in the definition of a TP? Looking ahead, organizing compounds around common key event should produce a list of likely TPs that could form the basis for assay development. Characterization of downstream biological responses (key events) for a group of chemicals according to these shared MIEs might provide a qualitative grouping needed to start the categorization process for TPs and develop more detail about the function of the pathways.

Compiling and combining biomarker information from animal experiments for categorization would provide a parallel opportunity to identify key/relevant TPs and the biological events that comprise the pathways. The goal would be to distil knowledge from diverse animal

studies to achieve wide coverage of possible TPs leading to definitions of adversity with a goal of testing the TPs with the fewest assays/tests. Thus, a key issue for consideration was is to find a combination of individual TP assays covering common pathways for well-defined MOAs and suites of assays capable of identifying TPs for new compounds. Another question that resurfaced was identifying which cell types to use in testing when there are so many available. This question, however, is likely to be assay dependent, i.e., the optimal cell for use in a TP assay will likely depend on the TP itself.

Broadly, the identification of TPs spans quite a range of biology – MIEs, initial cellular interactions, a set of embedded steps/nodes, recruitment of networks , leading, finally, to adverse cellular/tissue level consequences. In applying the TP concept, it may turn out that there are a more limited set of pre-defined TP "modules", e.g. signalling motifs (Alon, 2007; Zhang *et al.*, 2010). These "common structural pieces" of pathways", i.e. mini-pathways, recur repeatedly across different types of adverse effects and/or species. Apical adverse effects probably arise from perturbation of a network rather than that of a single pathway. Other than pharmaceutical molecules, most compounds of concern for regulation do not appear to have specific targets, sometimes referred to as selective toxicity (Thomas *et al.*, 2013). The deciding factor for adversity probably is the concurrent perturbation of related pathways that culminates in a breakdown of homeostasis and biological networks. Placing TPs in the context of these networks should aid in clarifying the linkages between MIEs, TPs and adversity.

Using results from Toxicity Pathway based testing for risk assessment

Following identification of functional TPs, test procedures will generate results that serve as a basis for risk assessments. Eventually, standardized evaluation of TP assays, in the context of an overarching TP/MoA/AOP framework, will be necessary to convince stakeholders of the practicality of in vitro, mechanism-based assay approaches for risk assessment. Such an evaluation approach would assess test systems against their ability to capture TPs, both qualitatively and quantitatively, and rely on positive and negative controls for assessing assay reliability and performance. Careful development of assays and standardization will help immensely in transitioning from current practice. However, these activities will come later after getting some TP-assays accepted as preferred options for use in testing and risk assessment.

Another topic of considerable interest was the likelihood that tiered approaches will evolve for using TP results in risk assessment (Jaworska and Hoffmann, 2010). Specifics of a tiered strategy may include computational toxicology-QSAR methods and high-throughput screening to infer MIEs. Inferences about target TPs from these test modalities would then trigger specific TP testing. TP assays would then evaluate specific MoAs. A variety of quantitative-high throughput screening (q-HTS) tests used in the US EPA ToxCast program provide insight into specific modes of action (Judson *et al.*, 2010; Kavlock and Dix, 2010; Reif *et al.*, 2010) To date, development of many of the current quantitative-high throughput screening (q-HTS) tests has not followed a purpose-driven design process, in which the design manner drew heavily on knowledge of mechanism and MoA. Because of these deficiencies with many current high-throughput assays, the extant knowledge on in vitro TP assays probably needs 'reprocessing' to achieve a better understanding of key toxicological processes and to provide a blueprint for the design of integrated testing and assessment systems from TPs. The general conclusion was that information on MIE/TPs has several possible uses. They could serve as the basis for the first step in a tiered testing strategy; they could assist in identifying chemicals that are relatively non-specific; they could find use in establishing regions of safety based on no-response regions of treatment; or, they could serve as a basis for investigation of a specific MoA.

Overall, then, a first tier might look at cellular perturbations in a broader set of assays to assess thresholds, no-effect-concentration, or margins-of-exposure; a second tier could look more closely at specific TPs, suggested from the earlier tier. The decision-point for moving to a second tier with these in vitro tests should consider both TP (MoA) and take into account for pharmacokinetic properties of the compounds in order to model expected dosimetry at target sites in vivo (Thomas *et al.*, 2013). The ability to predict in vivo dosimetry at target tissues need to be included in any form of integrated testing (Rotroff *et al.*, 2010; Wetmore *et al.*, 2011). Another factor is that tiered-testing strategies that include TP oriented assay testing must be applicable to chemicals irrespective of MOA/MIEs, i.e., they must be applicable both to those with highly specific targets and to those with very non-specific MoAs.

Endpoints useful for regulation

The use of in vitro points of departure to establish regions of safety implies the use of assays that do not relate to specific adverse outcome at the organism level. For a TP-based risk assessment process to succeed, then, the regulatory emphasis has to move away from the idea

that complete information on adverse outcome at the organism-level is necessary for a risk assessment. Although it may be difficult to predict specific adverse effects in relation to alterations in gene expression or TP based test outputs, it may be possible with these assay results to predict regions of safety. This region of safety approach would identify concentrations that do not lead to toxicologically significant perturbations of signalling pathways in vitro and translate these concentrations to predict the exposure required to produce similar concentrations in vivo (a process called reverse dosimetry). In the context of the idea that network alterations drive toxic responses in the intact animal, it may be possible to find regions where minimal alterations of specific TPs are not sufficiently large to alter overall network behavior and toxicity (Andersen and Krewski, 2010).

Some early efforts to evaluate no-effect-responses are promising. When comparing doses giving minimal alterations in vivo and those concentrations affecting ToxCast assays, the in vitro assays gave values for effective concentration that were within about a factor of 10 of those causing genomic changes in vivo . (The in vitro values were, in general, lower than the in vitro values.) Neither of these assays predicts in life adversity, but both types provide estimates of regions with minimal responses. An in vivo approach of this kind or a similar strategy with in vitro assays could define "regions of safety" where adverse responses are unlikely. The conversion from regions of safety of this kind to an exposure standard is not at all straightforward. Neither is it clear which in vitro measure, e.g., an AC50, BMC10, is best for risk assessment calculations. It may be that some aggregate measure across assays will be the most appropriate as the point of departure.

For instance, assays currently in use by ToxCast are overlapping with respect to MIEs and TPs. Will there be greater value looking at results across multiple assays to define TPs and develop results for risk assessments or will is work better to have specific integrated in vitro assays on which to base a formal risk/safety assessment? Grouping in vitro (ToxCast) assays based on the biological activity appears to be a more accurate and robust way of predicting apical toxicity than would be possible by taking each assay in a battery individually (Kleinstreuer *et al.*, 2011; Martin *et al.*, 2011; Sipes *et al.*, 2011). Approaches that account for a broader range of in vitro responses in developing predictive models do not fare as well (Thomas *et al.*, 2012). One consideration here though is whether the goal is prediction of likely qualitative responses (prioritization) or of applications of the data for formal risk or

safety assessments. For the latter, assay design criteria and standardization take on a higher level of importance.

For applications beyond hazard identification/prioritization, there will be greater need to move to multiple read-outs for TP perturbations, to include systems approaches for conducting dose/response assessments, and to apply pharmacokinetic tools for reverse dosimetry and in vivo-in vitro extrapolation. Computational modelling of MIEs at the cellular level will require an on-going process to create the ontology of perturbations of TPs, with expected consequences of pathway perturbations in vivo, and to understand the range of cellular response patterns and signaling motifs that need inclusion in order to create biologically-structured computational models.

Implications for risk assessment

The last session evaluated questions about the possible use of results from in vitro tests for TP perturbations in more formal risk assessment approaches. The discussions focused on how lessons from other initiatives could guide a TP approach. To the extent that compounds target a specific TP with a well-defined MIE, TP assays enable a mechanistic approach to risk assessment. A transition to an in vitro TP test platform is a natural extension of the efforts over the past 20 years to create MoA frameworks for human relevance. MoA approaches almost exclusively focus on well-studied compounds with large amounts of animal test results and mechanistic studies both from in-life and in vitro studies. The pattern has developed of conducting in life studies and then moving to mechanistic studies to explain animal results. Re-orientation to a TP-approach to toxicity testing would likely reverse this order (Andersen and Krewski, 2010). In vitro studies would provide the basic results for a risk assessment and in-life studies (at least in the US) might follow for high volume chemicals with widespread exposures or for specific endpoints. The TP, as depicted in the overall AOP, is a sub-component of the MOA, focusing on cellular and perhaps intercellular levels of response. The Risk21 project at ILSI-HESI in the US (see http://www.hesiglobal.org/i4a/pages/index.cfm?pageid=3492) looks to extend the MOA from discussions of human relevance to include dose-response. The experience in the risk assessment community with MoA should provide confidence in moving more aggressively in a TP direction.

We already have many instances of success in replacing in life studies with in vitro methods. These in vivo methods have moved through a formal process for validation by governmental organizations, such as ECVAM. Reliable and cost-effective in silico and vitro assays identify TP-based MIEs that are relevant in a MoA analysis. While MoA did not encompass pharmacokinetics, physiologically based biokinetic (PBBK) modelling has contributed to understanding plasma and tissue dosimetry (Blaauboer, 2010). This technology will now contribute to in vitro in vivo extrapolation (IVIVE) – one of the critical pieces in predicting exposures that will produce tissue concentrations similar to the in vitro concentrations that affect specific TPs. Advances in computational systems biology pathway modelling will support extrapolation across concentrations and provide concentrations for IVIVE (Bhattacharya *et al.*, 2011).

Using case studies to drive progress

An opportunity in the transition to a TP-based test platform could take advantage of case studies using compounds that were subject of MoA framework procedures (Andersen *et al.*, 2011). These case studies – such as those with CAR, PPAR or AhR activation as the MIE – could compare and contrast TP methods with the MoA linked to conventional extrapolations based on in life results. The cases studies could begin by looking at how existing in vitro data could guide a TP based risk assessment and grow to focus on the necessary laboratory work to design assays and the modelling tools for a detailed TP-test based risk assessment. Gaining experience with case studies is extremely important. Such examples will develop prototype safety assessment frameworks based squarely on TP concepts using data from TP-assays. Experience with case studies, as was done with specific applications of the MoA framework (Boobis *et al.*, 2006; Boobis *et al.*, 2008; Julien *et al.*, 2009), will become a key opportunity for communication, thereby engaging developers/contributors, end-users, risk assessors and stakeholders. Some case studies could span from AOP (OECD QSAR project) through MoA (WHO/IPCS) and on to applications of TP-assay results for risk assessment.

Support of different risk assessment approaches

Not every application using TP-assays and extrapolation tools for IVIVE and computational dose response modelling would proceed to a formal risk assessment. TP-methods also support tiered-approaches that have sequential decision points and reflect tonnage, margins-of-exposures evaluated by reverse toxicokinetic modelling and the specific TP targeted by the

compound, leading to formal in-life toxicology studies in margins of exposure are below some cut off level (Thomas *et al.*, 2013). . Case studies focus on use of established TP assays for compounds with relatively specific targets. Other assays will examine cellular responses to evaluate likely pathway targets more broadly. The specific targeted pathways will be critical for case studies and gaining confidence in the new TP-related tools. Of course, longer-term success in moving in these new directions will include education, both of existing professionals in doing toxicity testing and risk assessment and in programs training the next generation of practitioners. Training should begin as soon as teaching materials and informative case study materials become available.

Conclusions

Although it remains a challenge to reach consensus regarding the definition of a TP and even what constitutes a TP, a general agreement on the terms is essential for moving the central concept of a TP from an amorphous generalization to applications in contemporary risk assessment. Even without a clear definition, it is possible to specify attributes or criteria that define a TP that capture the intentions of the NRC report. The biology of a TP should cover the molecular and cellular domains. A TP should comprise a normal biological process. Activation of the TP in the intact animal leads to an adverse health outcome at sufficiently great levels of exposure. TP assays needs to provide numerical thresholds for perturbation for quantitative prediction. A TP should only include toxicodynamics of the pathways, and not toxicokinetic considerations specific to agents. Moreover, a TP description should comply with a standardised vocabulary of terms, nomenclature and description template. It is clear too that even possessing these attributes a one-size TP definition will not fit all, and thus the particular description of a pathway may differ depending on its intended use. A pathway description that serves as the basis for building a computational prediction model for risk assessment will likely differ from a description of the same pathway used to select a battery of in vitro assays for hazard identification.

Why did we find such a diversity of definitions and expectations of assay designs and applications from our relatively small number of participants? The workshop participants represented a diverse array of skills and backgrounds. In addition, the balance between European and North American participants may also have led to differences in familiarity with the idea of TPs. Alternatively, it may indicate that the very idea of 'Toxicity Pathway'

needs refinement to be useful in the on-going dialog around defining in vitro toxicity testing tools for chemical risk assessment. However, in our opinion the issue is not that we need a better or clearer definition, but in fact we need to invest more time and energy in making sure that the concepts and consequences of TP thinking are clearly communicated and understood in the wider community. Only then can we expect more consistent use of the term, more comprehensive TP descriptions that address the three key attributes that a TP should possess, and more conviction in the development and application of truly TP-based test systems for application in chemical risk assessment.

Annex I Toxicity Pathways Workshop Report

"Toxicity Pathways Workshop"
28-29 September 2011, La Quassa, Ispra (VA), Italy.

Co-organised by,
The European Commission Joint Research Centre and The Hamner Institutes for Health Sciences.
Hosted by the JRC's Institute for Health and Consumer Protection.

Meeting co-chairs: M. P. Whelan (JRC) and M. E. Andersen (THI).

Introduction

In 2007, the US National Research Council report, *Toxicity Testing in the 21st Century: A Vision and A Strategy,* envisioned a not-too-distant future where virtually all toxicity testing would be conducted assessing perturbations of toxicity pathways *in vitro*. Toxicity pathways were described as normal signalling pathways whose function could be altered by chemicals. Pathway perturbations of significant magnitude and duration would be expected to lead to adverse health consequences. In the intervening four years, there have been many discussions around various aspects of the 2007 report; however, little progress has been made in elaborating the 'toxicity pathways' concept and exploring how it can be exploited for human health risk assessments. Therefore, the European Commission's Joint Research Centre (JRC) together with The Hamner Institutes for Health Sciences (THI) organised a workshop to bring together a relatively small group of scientists (23 participants) from various backgrounds to bring some clarity (1) to what constitutes a 'toxicity pathway', (2) to identify the key pathways linked to cellular dysfunction, and (2) to propose fit-for-purpose in vitro assays and testing strategies. In addition, consideration was given to the analysis tools needed to apply results from pathway assays for human health risk assessments.

Participants provided short statements during the meeting reflecting her/his ideas about 'toxicity pathways' and how the pathways might be assayed, supported by specific examples. Background materials, such as key literature references were shared with participants beforehand. The meeting began with an overview of toxicity pathway thinking and related

concepts such as mode-of-action and adverse-outcome-pathways. Plenary sessions were complimented with breakout group discussions to examine three key questions, namely i) how do we define a toxicity pathway, ii) how can toxicity pathway thinking influence assay and test system design, and iii) how results pertaining to toxicity pathways could be used for safety assessments with respect to human health.

The deliberations of both the plenary sessions and the breakout groups were captured by rapporteurs as sets of bullet-points, as reported here. The organisers felt that this approach worked best to faithfully capture the substance of the discussions in an objective manner, providing 'raw' content to the participants and others to summarise, analyse, interpret and communicate in a manner which they see fit.

Participants

The invited experts that participated and contributed to the discussion were, in alphabetic order: Leonidas Alexopoulos (National Technical University of Athens), Gordana Apic (Cambridge Cellnetworks Ltd.), Kim Boekelheide (Brown University), Edward Carney (The Dow Chemical Company), Mark Cronin (Liverpool John Moores University), David Gerhold (NIH Chemical Genomics Center), Thomas Hartung (Johns Hopkins University), Miriam Jacobs (EFSA), Richard Judson (US EPA), Robert Landsiedel (BASF), Avi Maayan (Mount Sinai School of Medicine), Bette Meek (University of Ottawa), Pierre Moulin (Novartis), Kathleen Plotzke (Dow Corning Corporation), Julio Saez-Rodriguez (EMBL-EBI), Michael Schwarz (University of Tübingen), Russell Thomas (THI), and José-Manuel Zaldívar (JRC). The discussions were further assisted by the JRC note takers: Elisabet Berggren, Sharon Munn and Andrew Worth (all JRC).

Abbreviations

TP Toxicity Pathway
MoA Mode-of-Action
AOP Adverse Outcome Pathway
MIE Molecular Initiating Event

1. Defining Toxicity Pathways

1.1. Following Eric Berlow[1], it was noted that a complex problem is not necessarily complicated (to describe and model) and that "simplicity lies on the other side of complexity". For instance, with toxic responses, chemically-induced biological/toxicological effect can be more predictable if viewed in the broader context of a TP/MoA/AOP. Multiple biological processes take part in causing pathway perturbations - including absorption, action at target tissues, molecular events on a cellular level, inhibition or activation of toxicity pathways, adaptation to stressors and finally, at high degrees of perturbation, adverse consequences . Nonetheless, we expect that there are key nodes (events) in these pathways that are sufficient for understanding and predicting an adverse outcome and that the other parts of the pathways are secondary with lesser importance for the adverse outcome.

1.2. There are diverse methods used to describe MoA and TPs in the literature today; however, they are not necessarily very helpful when trying to categorise and map TPs. Therefore, it is important to find a consistent terminology and methodology to describe pathways leading to adverse health effects, and to set up a framework to better understand and organize available information. A consistent organization of TP information would assist in identifying knowledge gaps and stimulate the scientific community to target toxicity testing research efforts where actually needed.

1.3. How does the TP concept relate to the more well-established concept of MoA? After considerable discussion, the two concepts appear to be very compatible: the TP is lies within a more expansive yet less detailed MoA. To reinforce this complementary aspect, it would be worth considering refining/revising/expanding the definition of a MoA while attempting to define a TP. It might very well be the case that you can't have a MOA without a TP.

1.4. In the context of an AOP, much like the MoA, the TP was regarded as the core element of the AOP. In that respect every TP has essentially a MIE (or perhaps

[1] Eric Berlow: How complexity leads to simplicity. YouTube video, http://www.youtube.com/watch?v=UB2iYzKeej8

multiple MIEs) which may eventually result in an expression of adversity at higher biological levels. It was stressed that a TP is nothing more than a normal (canonical) signalling pathway that can be altered or perturbed (e.g. by xenobiotics) so as to trigger a cascade of events that ultimately lead to an adverse effect. The perturbation leading to toxicity would typically be related to dose, but could also depend heavily on site and time of action (e.g. sensitive time-windows of exposure, in certain anatomical regions, particularly during development).

1.5. In the absence of clear and objective criteria, the determination of adversity can be very subjective. Adversity can be defined at multiple levels of biological complexity. While it probably does not make sense to define it at the molecular level, it can in principle be identified with responses at the cellular level, if these perturbations can be related to adversity at higher levels that cause concern.

1.6. An exercise was conducted to collect anonymous suggestions on how to define TP among the participants (see Annex II). Based on the common elements of the collected definitions, the following consensus definition seems to emerge:

"A toxicity pathway is a sequence of intracellular events which regulate normal biology which, when sufficiently perturbed by a xenobiotic, leads to an adverse outcome at the level of the cell, and possibly the whole organism."

1.7. It was agreed that a set of criteria (essential attributes) should be established to actually identify a TP, although there were different views on what exactly these criteria should be. For example, while it was agreed that a pathway should show the link between a pathway perturbation and some measure of adversity, there were differing views as to whether a TP should include specific events such as an identifiable MIE, and also if quantitative criteria for pathway sensitivity would need to be defined (i.e. a threshold for the toxicologically-relevant perturbation).

1.8. If the definition of a TP is considered as context (use-case) dependent, then more than one definition may be required, or possible. However, if a TP framework is to be credible in a regulatory context, and the predictive tools that are derived from it

are to be trusted, then a (formal) definition is a clear prerequisite and thus overly subjective definitions will likely confuse the issue.

1.9. There might be a bit of a paradox emerging when endeavouring to better define the TP concept. Following the original thinking (NRC 2007) the biological 'domain' of a TP doesn't extend past the cellular level, and thus 'adversity' associated with a TP is described in terms of cellular pathology (responses). However, most discussants agreed that a (normal) signalling pathway can't be considered as a potential TP unless adversity occurs at a higher level of biological organisation (organ, organism). As a workaround, do we need to first sufficiently describe the MoA in which the TP is embedded, where adversity is described in a sufficiently apical manner, and then work backwards to give the cellular events an apical-adversity context? To illustrate this point, consider apoptosis – this is a cellular event that is necessary in a healthy organism, but it could also be an event implicated in a MoA associated with an adverse health outcome (e.g. cancer). Likewise, activation of a nuclear receptor doesn't constitute a TP either.

1.10. An interesting variation of the definition and possibly equally relevant could be: *"Toxicity pathways are alterations of signalling motifs, such as altered gene expression, that are predictive of toxicity".* This definition implies that, for example, the MIE does not necessarily need to be identified, but only the main nodes in relevant pathways that the chemical agent perturbs. Due to the effect and potency of perturbations of the node(s), the motif will vary and might lead to a sufficient basis for understanding the toxicity profile of the agent and the predicted adverse outcome. 'Classifiers' (prediction models) would relate likely adverse outcomes to alterations in characteristic/phenotypic motifs in signalling or expression 'space'. In addition, prediction may not only rely on detecting changes in motif state per se, but also on identifying the trajectory of change. Often the sign of adversity is captured in how a biological system attempts to change/adapt to the perturbation.

1.11. One could extend this 'motif' basis of a TP framework to any biomarker/descriptor space, including biomarkers associated with cellular read-outs from in vitro assays. Thus a change in a motif in biomarker space (the motif being a set of biomarkers

linked in some way) that can be linked to an adverse outcome would be considered as a TP. The features/properties of such TPs would be different than for TP definitions based on a MoA concept. In this way, the criteria that one would use to define and identify TPs would also differ. Thus, a single definition will not be possible if there remains multiple conceptualisation of what constitutes a TP.

1.12. The dose (magnitude, kinetics) of a certain chemical agent will dictate whether a TP is sufficiently 'activated' to lead to an adverse effect. For instance, a pathway that is 'modulated' may not result in an adverse outcome due to the ability of the system to adapt/compensate. It was argued that modulation wouldn't constitute in fact a TP – a TP would only exist if there is toxicity, otherwise it's just a pathway. The 'perturbing' dose is then the dose that would feed into a decision framework for safety assessment. Thus, a purely qualitative understanding of the pathway would be a component but not a driver of a risk assessment. For a quantitative application the dose dependent characteristics of a TP have to be well described/modelled.

1.13. A biologically-based mathematical model for a pathway should be developed to be the basis for the risk assessment. A next step might be how to describe interacting pathways and introduce this information into the pathway model. The aim would be first to identify and validate a pathway or network of pathways, and then identify the main nodes in that pathway/network that could be used as a basis for prediction of adversity and ultimately a safety assessment. If it were possible to identify the key nodes in the pathway or interacting pathways, then a much less complex testing system would be a sufficient basis for a safety assessment. This approach provides linkage between the more exact identification of a TP and the motif approach. It is possible to identify the essential nodes, either through the detailed knowledge of the TP or by observing patterns that correlate with adverse outcomes.

1.14. As an example it might be possible to look at hepatotoxic molecules for which we have mechanistic studies and already know something about the cellular events. In vivo transcriptomic data could be a basis for understanding pathways. The nodes of the pathways could be then associated with specific processes of proteins, such as tubulin, proteasome, mitochondrial electron chain, etc.

1.15. Other considerations were whether all biological pathways can become TPs, and whether there are cases of TPs that are not based on normal biological perturbed beyond (local) homeostasis. For instance, can novel biological processes/pathways be 'created' by a chemical agent? The participants at the workshop concluded that are no TPs that could be defined that do not stem from perturbations of existing (signalling, metabolism) pathways within the cell. In addition, it was recognised that probably all normal biological pathways could lead to 'toxicity', but only a limited number of those would have a real impact at the tissue or organism level. It was further recognised that often only a combination of certain patterns of up-regulated or down-regulated pathways would lead to an adverse outcome. This behaviour has for example been demonstrated for carcinogenicity, where certain up-regulated or down-regulated events by themselves would not be regarded as TPs, but support the possibility for an adverse effect at the cellular or intercellular level.

1.16. The AOP is distinctly different from a TP in that it is described across multiple biological levels of complexity/organisation. This expansion adds to the complexity for setting up prediction models for the adversity based on upstream events. The ability to predict downstream events tends to decrease as the extent of extrapolation from the upstream event (e.g. the MIE) increases, due in part to the stochastic nature of biological systems.

1.17. Another aspect of TPs involves a focus on response and not dose. Since in vivo biokinetics is such a fundamental part of hazard prediction, should we be speaking not only about toxicodynamic pathways, about also about toxicokinetic pathways?

1.18. There is a need to develop a suitable nomenclature specific to toxicity pathways that reflects the fact that a pathway comprises a chain of events starting from a MIE and leading to an observable/measurable biomarker of pathology/cellular response. Currently a TP is often labelled or referred to by a typical/reference chemical that induces it, by its MIE, an intermediate event, or by a pathological biomarker. The lack of a working nomenclature leads to imprecise shorthand descriptions of TPs and may impede communication between scientists.

1.19. The set of assays to define a pathway or nodes in a pathway could primarily be a tool to group chemicals according to their MoA, rather than to be directly predictive of an adverse effect. Predictions of adversity for data-poor chemicals could then be obtained by read-across from “similar” data-rich chemicals in the same group. This is an extension of the (chemical) category approach to toxicity prediction. Furthermore, predictions can be expressed as probability statements (likelihood that a chemical in a given category has an adverse effect), which in itself would change the traditional basis of a safety assessment.

1.20. In response to the question “what is a TP” there seemed to be some consensus that these were normal signalling pathways that could be perturbed by the interaction of a toxicant (xenobiotic). Moreover, these pathways were part normal homeostasis processes that can be altered by dose, and which might have different levels of sensitivity depending on context (life stage, sex, (epi)genetic profile, etc). Threshold dose/response transitions were considered to be key in determining when a normal pathway becomes a TP.

1.21. It was agreed that in order to judge whether a pathway was "adversely" activated or just activated/modulated, a linkage to some measure of phenotypic change would be required. The term “chemical epidemiology” was proposed to describe the evidence-based determination of the threshold value.

1.22. It was considered that not all pathways were as important as each other with respect to potential for perturbation by toxicants. Glycolysis was given as an example of a pathway which is very important but for which there are few examples of perturbations by small xenobiotic molecules. There is a need to identify those TPs that are important and relevant, and provide clear examples of what constitutes a TP and what doesn't.

1.23. The Ankley schema [Ankley et. al (2010)] depicting the hierarchy of a TP as being part of a network which is part of a MoA, which is part of an AOP was seen to be very useful. Each has different levels of granularity, the pathway being most detailed and touching on the lowest biological level of molecular information. All are valid and need to be looked at in a complementary manner.

1.24. It was suggested that one could either start by probing a specific pathway or start with a cellular response and work backwards to identify the pathways leading to that response. It was not clear which approach was likely to lead to the greatest progress.

1.25. Although for simplicity it may be preferred to think in terms of (linear) pathways it was agreed that signalling pathways do not act in isolation and in reality there is certainly more of a network of activity, the sum of which leads to a cellular response: context, magnitude and time/timing being important factors in determining a phenotypic outcome. Thus, focusing on understanding the connectivity of signalling pathways and how such interconnectivity/interdependence can influence adverse effects might actually be the decisive objective of establishing the utility of the TP concept, rather than concentrating on a single TP in isolation.

1.26. Evidence from using animal models that indicate no effect on phenotype when certain genes are knocked out seems to indicate redundancy in the function of normal biological pathways. Considering the biological processes of evolution would also support this view.

1.27. The spatial, temporal and cellular contexts of a TP were stressed as being crucial to the relevance of a TP to an adverse outcome. They will also affect transition of perturbations from one level of biological organisation to the next. For example, it is more important to define interaction of xenobiotics on TNF alpha pathway/target in kidney proximal tubule cells than in a macrophage or hepatocyte? The impact of a xenobiotic interaction with estrogen receptors (ERs) in the breast is likely to have a different impact than interaction with ERs in other tissues. Consequently, it is important to overlie what we know about biology in considerations of perturbations of pathways and their significance for toxicity.

1.28. There may be impacts of exogenous substances on up- or down-regulating gene expression of cells, but this does not necessarily lead to overt toxicity (e.g. isoflavones). The lack of an adverse outcome means that the pathway would not be considered a TP in this context; however, in another case an exogenous substance may trigger the same pathway which may be linked to an adverse phenotypic

response (e.g. DEHP [ref]). In this case, the pathway becomes sufficiently perturbed that it would be considered a TP for the latter substance.

1.29. In considering the level of biological organisation, it was most relevant to focus on the point of first interaction of a xenobiotic with the cell machinery i.e. the MIE, and then follow as far as possible the pathway leading from first interaction to a phenotypic change associated with an adverse health outcome.

1.30. It was considered that TPs are at the moment incomplete in their description, with little overlap and integration between activities of proponents in the field. It would be highly desirable to harness efforts in a collaborative fashion. This may be facilitated for example through a "wiki" forum designed along a pathways theme/concept. Postulated pathways should be developed from, and challenged with, real data and information on biology and chemicals, to test their robustness and validity within a broad scientific community.

1.31. One could conceive of a 'grand plan' – a systematic mining of literature to assemble and map mechanistic/MoA knowledge onto a well-structured framework, assessing pathways and their susceptibility to external perturbation and thereby considering their candidacy for nomination as a TP. Once the TP landscape is described with a sufficient level of coverage and granularity, the process of assembling a TP network and identifying key/choke events/nodes can begin, which will ultimately inform a rational, purpose-driven design of integrated testing and assessment systems. The scale of such an endeavour is indeed 'grand', and will require the engagement of scientific communities that go far beyond the capacity of any individual project or consortium. The solution may lie in the establishment of 'reference' goals, objectives, approaches, and standards, subsequently communicated to, and accepted by, a critical mass of the scientific community, thereby ultimately engaging a diverse and defuse set of groups/entities in a common cause.

1.32. A tool commonly used in the systems biology community to capture the attention of state-of-the-art practitioners is to pose a problem looking for a solution – a competition open to all players ("Dream", "Improver" initiatives). The stage is set with a clear definition of the challenge, and the provision of the necessary datasets,

metadata, and a description of what constitutes success. Solutions proposed are assessed by experts against success criteria, and made public. Such an approach not only taps into the latest thinking and techniques, but is highly educational for all concerned. Could we conceive of adopting such a model to engage a broad section of the scientific community in a focused manner, to develop the TP concept and apply TP methodology to solve safety assessment problems?

2. Using Toxicity Pathways

2.1. There were differences of opinion regarding the level of detail describing a TP that would be required before it could be actually used to guide the design of a related test system. However, a general principle that was identified was to "explore lots and use little". In other words, it is often necessary to carry out extensive research to on the mechanisms of toxic action before the most informative and discriminating key events can be determined and short-listed for use in risk assessment.

2.2. In applying the TP concept, it could be useful to refer to pre-defined TP "modules". These are "commonly reused chunks of pathways", i.e. non-redundant mini-pathways that are found to recur across different types of adverse effects and/or species. A term with similar meaning, "the protein interaction sub-network" was also proposed. There are conceptual parallels here to be drawn with synthetic biology philosophy.

2.3. It would be useful to measure the extent and time course of a perturbation, i.e. determine how many genes are disturbed and whether they stay up-regulated over time or resume their initial state.

2.4. We're asking questions of in vitro systems that we never asked of in vivo models.

2.5. Further it must be estimated whether the effect is connected only to the disturbed pathway, or does it have a major effect on other pathways that may be relevant for an adverse outcome, i.e. a cascade of events involving many pathways of which several might contribute to the adverse outcome. Considering the non-target-specific

nature of most molecules (non-pharma) then perhaps the deciding factor is more about the concurrent perturbation of related pathways that culminates in a breakdown of homeostasis and a progression towards dysfunctional and adversity.

2.6. The standardised evaluation of tests, in the context of an overarching TP/MoA/AOP framework, would be an important added value. There should be a standardised approach to test-method evaluation, which would make comparisons more straight forward and more reliable. Such an evaluation approach would assess test systems against their ability to capture TP(s) both qualitatively and quantitatively.

2.7. It was suggested to use a tiered approach in utilising test systems for hazard identification and characterisation. The first tier would look at cellular perturbation at certain thresholds, leading to a second tier where a further understanding of the MoA would be the focus. Also, ADME considerations (exposure kinetics at target sites) could be included in a tiered approach.

2.8. MoA studies could be used to move away from the idea that complete information on adverse outcome is necessary for a risk assessment. A careful selection of case studies should be able to show that this is possible. The case studies would be needed to convince actors in the regulatory field that the information traditionally required would not be needed to make an accurate safety assessment.

2.9. Test assays must be applicable to all types of chemicals. An early tier that is giving much more confidence is requested, this can only be obtained based on a series of collaborations to get a global view of pieces currently contributing to such testing strategies. The current in vitro tests were not necessarily developed in a purpose-driven manner drawing on knowledge of mechanism and MoA. This deficiency is why current knowledge must be re-examined and 'reprocessed' to achieve a better understanding of key toxicological processes and to provide a blueprint for the design of integrated testing and assessment systems.

2.10. Compiling and combining biomarker information from suitably conducted animal experiments would provide an opportunity to identify key/relevant TPs and the biological events that essentially define these pathways. Such distilled information

would then allow a wide coverage of possible pathways leading to definitions of adversity with the fewest assays/tests. Thus, a key issue is to find a combination of individual assays covering common pathways. Related to this is the necessity to identify the cell types that capture these pathways; such efforts at identification of suitable test systems would inevitably lead to a reduction in the number of cell-model candidates. The reduced number of assays could then be tested with a number of well-known reference chemicals of known activity to 'calibrate' the system.

2.11. We need to first define the TP framework as far as possible by mining current knowledge, and then to use this framework as a basis to design fit-for-purpose testing system. The first task is therefore to collect, organise, interpret and remap already available information.

2.12. To extend the knowledge base on TPs, it will be necessary to undertake a broad screening of literature both on MOA-based toxicity prediction methods as well as currently used in vivo apical toxicity test methods. There are already on-going projects to collect currently available information and understand MoA. In parallel it would be helpful to collect TP information related to different animals and organs, and to identify both positive (pathway perturbing) and negative (non-perturbing) reference chemicals. An alternative way forward would also be to use non-reactive chemicals to help to understand and define the pathways for the toxic ones.

2.13. In attempting to map the TP domain, identifying all possible MIEs might seem like a good start. However, it was also stated that the MIE or TP has to be linked to an adverse effect, and how to get the data to establish the relationship between MIE and adverse outcome remains an issue. So far correlation analysis between MIEs and adverse outcomes were not good; an MIE was considered to be far from an adverse apical endpoint in most cases. It was pointed out that pathways also need to be multidimensional to include kinetics.

2.14. Identification of chemicals sharing the same MIE/TP/signalling motif may be useful information which may remove the necessity to completely define the pathway to adversity in every case by applying a grouping approach to assessment based on shared MIEs. Measuring and predicting downstream biological activities (key

events) for a group of chemicals according to shared MIEs will provide a qualitative grouping needed to start the process. A next step may be application of kinetics to introduce quantitative aspects.

2.15. Basing prediction purely on observed correlation between adverse health effects and changes in biomarker/signalling motif type TPs will always have limitations, both in relation to predictive accuracy, transparency and credibility. Systems based on a TP framework that has a phenomenological/mechanistic basis will be more difficult to establish with respect to correlative/statistical based systems, but will ultimately perform better. Perhaps it will be a case of developing both types of system in parallel, transitioning from one to the other as our ability to capture and model TPs grows.

2.16. In terms of computational modelling of MIEs, robust definitions and ontology of effects/perturbations would be required in order to be able to model them. Computational models could also play a role in modelling key events downstream of an MIE provided data could be obtained for these individual processes. The role of chemical reactivity should not be underestimated and QSARs have been found useful in grouping according to chemical reactivity.

2.17. A TP could be either qualitative or quantitative. Qualitative information on the chemicals that trigger and that do not trigger pathways is needed. For application beyond hazard identification there is a need to move to quantification and dose/response assessment.

2.18. In considering the need to relate potency in vitro to potency in vivo, it was considered that much of the apparent difference was related to ADME and examining potency within pathways may not be so relevant. Key events downstream to the MIE/TP were considered to be the likely rate limiting steps (choke points) where quantitative aspects could be best applied.

2.19. One question raised was the level of biological organisation required in any model to be predictive of a specific apical adverse effect. In this regard it was noted that this may be the wrong focus since even at the intact organism level it as a false notion to

believe that we can predict specific effects in humans from those seen in the rat. If there is an adverse effect in one species, there might be a similar or a different adverse effect in another species. Yet, in each they could arise as a downstream consequence of the same MIE.

2.20. Although it may be difficult to predict specific adverse effects as a consequence of alterations in gene expression, it may be possible to predict **regions of safety**. This approach is based on identifying dose ranges that do not lead to toxicologically significant perturbations of signalling pathways and translating these in vitro concentrations to target-site in vivo concentrations related to external exposures (so-called reverse dosimetry). The use of in vitro (as opposed to in vivo) points of departure to establish regions of safety implies the use of threshold values that are not necessarily related to any specific adverse outcome at the organism level.

2.21. Using this approach, comparing in vivo results with the most sensitive ToxCast assays gave values within a factor of 15 lower of the in vivo apical endpoint. Assuming the available TP assays provided good coverage of possible modes of toxicity (i.e. preferably screen entire genome) you could predict what would be safe level although you would not predict the specific hazard. This approach was proposed as a way to define a region of safety where most likely you would not see apical toxicity.

2.22. Grouping in vitro (ToxCast) assays based on the biological activity that hey measure might be a more accurate and robust way of predicting toxicity, as opposed to taking the result from each assay in a battery individually. Groupings should be on a pathway basis, and not just similar types of assays that may be measuring very different things/pathways. Could grouping according to pathways and the knowledge of the biology of the pathway with which they are associated help define the likely adverse outcomes?

2.23. Overall a suggestion was that MIE/TPs can be used as a basis for the first step in a tiered testing strategy, primarily to screen for non-reactive molecules, predicting regions of safety or as a basis for investigation of a specific MoA.

2.24. Case studies were also proposed as a practical next step to demonstrate use of genomic data within a MoA framework. A high priority area such as identification and assessment of endocrine disrupters might be a good example for the first round of case studies.

2.25. An issue was raised of how to name and annotate TPs, and how to share knowledge on them through a data/knowledge-base. Such a knowledge-base would be a valuable resource but how would this be established and governed. These activities will depend on the development of a consistent ontology for TPs and their elements, and the use of this ontology to define the data model of the data/knowledge-base.

2.26. One practical approach suggested in postulating AOPs was to identify biomarkers from the literature such as certain signalling pathways that are disrupted over and over again in the development of an adverse outcome such as cancer. A meeting of experts could decide which subset of signalling pathways is most interesting to use. Computational models could help with this task. From this subset, choke points (i.e. most critical events) could be identified and from this evaluation a set of candidate assays could be developed in suitable cell types. These assays could be challenged with reference substances causing cancer and some that don't, results examined and more assays conducted on the basis of the results in an iterative manner. This approach favours the unsupervised identification of pathways followed by confirmation of whether or not they are relevant. Using crowd sourcing, one could solicit the contribution of scientific peers to this activity.

3. Prospects for Safety Assessment

3.1. Knowledge of TPs should enable a mechanistically-based and hypothesis-driven approach to risk assessment. While this is clearly advantageous in terms of developing the science of toxicology, as well increasing confidence in toxicological assessments, one of the largest challenges lies in obtaining consensus on how to apply the TP concept and to get the risk assessment community to accept the new thinking. Currently, most assessments are made based on observational apical endpoints in vivo without any further knowledge of the background or

understanding of the mechanism leading to toxicity. It would not be necessary to achieve a full mechanistic understanding when assessing a chemical, but a basic idea of key events and consistency with the adverse outcome would be needed. This dose-response analysis based on key events is already being taken into practice by risk assessors within the application of the MoA framework. In order to move forward and embed TP into MoA frameworks it will be necessary to engage the risk assessment communities, including considerable training and outreach to educate stakeholders about the potential of these alternative testing and safety assessment approaches.

3.2. The MoA framework has been applied so far primarily for data rich substances in a very detailed way in order to understand the human relevancy of animal data. It is now necessary to bring the TP information into the MoA framework and then see how it fits into specific outcomes within a MoA. TPs offer a higher degree of resolution to MoA analyses. They may also provide the opportunity for less expensive, predictive assays as a minimum data set to be requested from registrants which would provide some basis to ask for more tailored data where necessary. The overall scope is to come to the point when we can base risk assessment on testing with TPs. We don't yet know how many pathways would need to be covered to be adequate for risk assessment?

3.3. Today there are instances where in vitro methods have replaced in vivo assays. However, in general the animal studies are conducted before MoA studies are contemplated.. It would make much more sense to do it in the opposite way, using case studies on data rich substances to extend to structurally related data poor chemicals, thereby using pathway-based models to demonstrate a shared MoA. In this manner it may be possible to make the case in a regulatory context for acceptance of a TP/MoA-based safety assessment without the use of animal studies.

3.4. Some highly reliable and cost-effective in chemico and vitro assays are available to identify MIEs that are relevant in a MoA analysis. There should not be a real difficulty to introduce those approaches. PBBK modelling has shown to be powerful in predicting internal doses in target organs, which was always identified as a gap of knowledge in in vitro testing. Multiple competences are needed to go through

examples and find the way through to develop a set of assays that could be used to fulfil a safety assessment based on a couple of known PT/MoAs.

3.5. The elaboration of case studies is extremely important in the initial stages to develop and challenge prototype safety assessment frameworks based squarely on TP concepts and data derived from TP based test/assessment systems. They also represent a powerful tool for communication, to engage developers/contributors, end-users and stakeholders. Perhaps some case studies could be developed around recently proposed examples of AOP (OECD QSAR project) and MoA (WHO/IPCS).

3.6. Further, it must be possible to agree on a target where it is acceptable not to test further, every tier of testing is giving more information, but it is necessary to set a limit using a cost benefit approach. For example, is it acceptable to put a chemical on the market that is unlikely with 90% probability to lead to an adverse outcome? It must be realistic to think in probabilities; there are no absolute certainties in current risk assessment procedures.

3.7. An important pre-requisite in gaining regulatory acceptance and widespread application of the TP concept will be to harmonise terminology, and to develop an agreed ontology of intermediate effects (potential key events) and adverse outcomes, as well as reporting formats that describe intermediate effects, TPs and AOPs in a consistent manner.

Annex II Definitions of Toxicity Pathways proposed by workshop participants

At the workshop, the participants were asked to provide their own definition of the term 'toxicity pathway'. Below is a list of the individual definitions proposed by individuals;

1. Set of molecular entities whose perturbation leads to toxic events.

2. A model and a module of interacting cellular components that are variables that together can be used to predict toxicity.

3. A circuit of key molecular and cellular processes which regulate normal biology, but when sufficiently perturbed, lead to adverse outcome.

4. A series of steps leading from an early event to a toxic effect. The early event can be a chemico-biological reaction. Subsequent steps of a TP drive an organism towards an apical/toxic effect. To define a TP, the early event, the apical toxic effects and at least one step in between need to be defined. TP are generally no new events but a sequence of 'normal' biological reactions in the wrong context or at the wrong extent. Here the number of possible TP is finite.

5. A sequence of intracellular events leading to an adverse effect; which could be parts of signalling, metabolic, gene regulatory pathways that are relevant for a toxic outcome.

6. A biological pathway which significantly perturbed by internal/external agents, can result in advance effects in the organism.

7. Molecular level process (including signalling metabolic pathway, collection of molecular interactions, cellular process) that can be perturbed in a measurable way by a chemical; and for which there is a documented association or linkage between perturbations of the process (perhaps through multiple steps) and cellular, tissue or organismal toxicity.

8. A collection of molecular events that comprise a normal signalling pathway and are perturbed by a toxicant. The pathway exists within a series of representations with

increasing granularity starting from the molecular initiating event to the pathway, network, cellular response, and organ/tissue response. The linkage between perturbation of the pathway and the downstream apical response is context dependent and depends on magnitudes, temporality and spatial aspects of the perturbation.

9. A TP has a definable molecular initiating event; involves perturbation of normal biological pathways resulting in a toxic effect; can be defined, although need not to be defined in full; is the core of mode of action and adverse outcome pathway.

10. A TP describes a series of subsequent or simultaneously occurring changes in the function of cellular macromolecules that ultimately cause an adverse effect on the organ or organism level.

11. A normal biological process that when sufficiently perturbed (or stimulated at an inappropriate time) to lead to an apical adverse outcome minimally at the cellular level.

12. Alterations of signalling motifs that are predictive of toxicity.

13. A TP is a normal signalling pathway that can be perturbed by chemical treatment. Perturbations that are sufficiently great or sufficiently prolonged lead to toxic responses entering cell preparations or in vivo.

14. Altered pathway in the biological control, signalling or/and metabolic cellular networks as a consequence of the interaction with a toxicant.

15. TPs are normal multi-step biological pathways that are context dependent and predictive of diversity when sufficiently perturbed (by an exposure).

16. TP: A (biochemical or signal transduction) pathway that has been demonstrated to trigger an adverse event when sufficiently perturbed.

17. TP are molecularly defined chains of not necessarily linear cellular events starting from point of chemical interaction to perturbation of metabolic networks and phenotypic

change. TP are causal – either necessary or aggravating – and will typically have a threshold of adversity.

18. TP are the formal description of toxic modes of actions on the resolution of biochemistry and molecular biology.

19. TP are causal links between a given toxicant and its effect in a Systems Toxicology approach.

20. TP (for a given xenobiotic) sequence of biochemical molecular, cellular, or tissular physiological events, induced by a defined exposure to a given xenobiotic and leading to a collection of functional and/or morphological consequence qualified as adverse.

21. TP (specific definition) sequence of events (biochemical, molecular, cellular or tissular) leading sequentially to irreversible injury when initiated by exposure to a xenobiotic.

22. TP (sensitive definition) sequence of events (biochemical, molecular, cellular or tissular) representing an elementary reaction, common to several xenobiotics, which in isolation or in combination with others, leads to reversible or irreversible injury at the level of the cell, the tissue or in the organism.

Annex III Workshop Agenda

Toxicity Pathway Workshop

28th September, 2011

08:15 *Hotel pick-up*

08:30-09:00 **Welcome and Introduction**
M. Andersen and M. Whelan

Session I: Defining Toxicity Pathways

09:00-10:00 **Plenary – Presentations and discussion**

10:00-10:30 *Coffee Break*

10:30-13:00 **Breakout Groups - Discussion**

13:00-14:00 *Lunch (buffet)*

14:00-15:30 **Plenary – Breakout Group Feedback**

15:30-16:00 *Coffee Break*

Session II: Using Toxicity Pathways

16:00-17:00 **Plenary – Presentations and discussion**

Return to hotel

Dinner on Isola dei Pescatori (boat departs 18:45)

29th September, 2011

08:15 *Hotel pick-up*

Session II: Using Toxicity Pathways (cont.)

08:30-11:00 **Breakout Groups – Discussion**

11:00-11:30 *Coffee Break*

11:30-13:00 **Plenary – Breakout Group Feedback**

13:00-14:00 *Lunch (buffet)*

Session III: Prospects for Safety Assessment

14:00-15:30 **Plenary – Presentations and discussion**

15:30-16:00 **Wrap-up and close.**

Return to hotel, departures to airport etc.

References

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**, 450-461.

Andersen, M. E., Clewell, H. J., Carmichael, P. L., and Boekelheide, K. (2011). Can case study approaches speed implementation of the NRC report: "toxicity testing in the 21st century: a vision and a strategy?". *Altex* **28**, 175-182.

Andersen, M. E., and Krewski, D. (2010). The vision of Toxicity Testing in the 21st Century: moving from discussion to action. *Toxicol.Sci.* **117**, 17-24.

Ankley, G. T., Bennett, R. S., Erickson, R. J., Hoff, D. J., Hornung, M. W., Johnson, R. D., Mount, D. R., Nichols, J. W., Russom, C. L., Schmieder, P. K., Serrrano, J. A., Tietge, J. E., and Villeneuve, D. L. (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ Toxicol Chem* **29**, 730-741.

Bhattacharya, S., Zhang, Q., Carmichael, P. L., Boekelheide, K., and Andersen, M. E. (2011). Toxicity testing in the 21 century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *PloS one* **6**, e20887.

Blaauboer, B. J. (2010). Biokinetic modeling and in vitro-in vivo extrapolations. *J.Toxicol.Environ.Health B Crit Rev.* **13**, 242-252.

Boekelheide, K., and Andersen, M. E. (2010). A mechanistic redefinition of adverse effects - a key step in the toxicity testing paradigm shift. *Altex* **27**, 243-252.

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.

Boobis, A. R., Cohen, S. M., Dellarco, V., McGregor, D., Meek, M. E., Vickers, C., Willcocks, D., and Farland, W. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit Rev Toxicol* **36**, 781-792.

Boobis, A. R., Doe, J. E., Heinrich-Hirsch, B., Meek, M. E., Munn, S., Ruchirawat, M., Schlatter, J., Seed, J., and Vickers, C. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit Rev Toxicol* **38**, 87-96.

Hartung, T., and McBride, M. (2011). Food for Thought ... on mapping the human toxome. *Altex* **28**, 83-93.

Jaworska, J., and Hoffmann, S. (2010). Integrated Testing Strategy (ITS) - Opportunities to better use existing data and guide future testing in toxicology. *Altex* **27**, 231-242.

Judson, R. S., Houck, K. A., Kavlock, R. J., Knudsen, T. B., Martin, M. T., Mortensen, H. M., Reif, D. M., Rotroff, D. M., Shah, I., Richard, A. M., and Dix, D. J. (2010). In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project. *Environmental Health Perspectives* **118**, 485-492.

Julien, E., Boobis, A. R., and Olin, S. S. (2009). The Key Events Dose-Response Framework: a cross-disciplinary mode-of-action based approach to examining dose-response and thresholds. *Crit Rev Food Sci Nutr* **49**, 682-689.

Kavlock, R., and Dix, D. (2010). Computational toxicology as implemented by the U.S. EPA: providing high throughput decision support tools for screening and assessing chemical exposure, hazard and risk. *J.Toxicol.Environ.Health B Crit Rev.* **13**, 197-217.

Kleinstreuer, N. C., Judson, R. S., Reif, D. M., Sipes, N. S., Singh, A. V., Chandler, K. J., Dewoskin, R., Dix, D. J., Kavlock, R. J., and Knudsen, T. B. (2011). Environmental impact on vascular development predicted by high-throughput screening. *Environmental health perspectives* **119**, 1596-1603.

Krewski, D., Westphal, M., Al-Zoughool, M., Croteau, M. C., and Andersen, M. E. (2011). New directions in toxicity testing. *Annu Rev Public Health* **32**, 161-178.

Martin, M. T., Knudsen, T. B., Reif, D. M., Houck, K. A., Judson, R. S., Kavlock, R. J., and Dix, D. J. (2011). Predictive model of rat reproductive toxicity from ToxCast high throughput screening. *Biol Reprod* **85**, 327-339.

OECD (2013). Guidance document on developing and assessing adverse outcome pathways. Series on Testing and Assessment No. 184, ENV/JM/MONO(2013)6.

NRC (2007). Toxicity Testing in the 21st Century: A Vision and a Strategy. The National Academies Press, Washington, DC.

Reif, D. M., Martin, M. T., Tan, S. W., Houck, K. A., Judson, R. S., Richard, A. M., Knudsen, T. B., Dix, D. J., and Kavlock, R. J. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environmental health perspectives* **118**, 1714-1720.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., Lecluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., Sochaski, M. A., Kavlock, R. J., Boellmann, F., Martin, M. T., Reif, D. M., Wambaugh, J. F., and Thomas, R. S. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicological sciences : an official journal of the Society of Toxicology* **117**, 348-358.

Rusyn, I., and Daston, G. P. (2010). Computational toxicology: realizing the promise of the toxicity testing in the 21st century. *Environmental health perspectives* **118**, 1047-1050.

Sipes, N. S., Martin, M. T., Reif, D. M., Kleinstreuer, N. C., Judson, R. S., Singh, A. V., Chandler, K. J., Dix, D. J., Kavlock, R. J., and Knudsen, T. B. (2011). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicological sciences : an official journal of the Society of Toxicology* **124**, 109-127.

Sonich-Mullin, C., Fielder, R., Wiltse, J., Baetcke, K., Dempsey, J., Fenner-Crisp, P., Grant, D., Hartley, M., Knaap, A., Kroese, D., Mangelsdorf, I., Meek, E., Rice, J. M., and Younes, M. (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul Toxicol Pharmacol* **34**, 146-152.

Thomas, R. S., Black, M. B., Li, L., Healy, E., Chu, T. M., Bao, W., Andersen, M. E., and Wolfinger, R. D. (2012). A comprehensive statistical analysis of predicting in vivo hazard using high-throughput in vitro screening. *Toxicological sciences : an official journal of the Society of Toxicology* **128**, 398-417.

Thomas, R. S., Philbert, M. A., Auerbach, S. S., Wetmore, B. A., Devito, M. J., Cote, I., Rowlands, J. C., Whelan, M. P., Hays, S. M., Andersen, M. E., Meek, M. E., Reiter, L. W., Lambert, J. C., Clewell, H. J., 3rd, Stephens, M. L., Zhao, Q. J., Wesselkamper, S. C., Flowers, L., Carney, E. W., Pastoor, T. P., Petersen, D. D., Yauk, C. L., and Nong, A. (2013). Incorporating New Technologies into Toxicity Testing and Risk Assessment: Moving from 21st Century Vision to a Data-Driven Framework. *Toxicological sciences : an official journal of the Society of Toxicology*.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J., 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., Allen, B., Judson, R. S., Singh, R., Kavlock, R. J., Richard, A. M., and Thomas, R. S. (2011). Integration of Dosimetry, Exposure and High-Throughput Screening Data in Chemical Toxicity Assessment. *Toxicological sciences : an official journal of the Society of Toxicology*.

Zhang, Q., Bhattacharya, S., Andersen, M. E., and Conolly, R. B. (2010). Computational systems biology and dose-response modeling in relation to new directions in toxicity testing. *J.Toxicol.Environ.Health B Crit Rev.* **13**, 253-276.

European Commission
EUR 26389 – Joint Research Centre – Institute for Health and Consumer Protection

Title: Thresholds for Endocrine Disrupters and Related Uncertainties.
Report of the Endocrine Disrupters Expert Advisory Group (ED EAG)

Authors: Maurice Whelan, Melvin Andersen

Luxembourg: Publications Office of the European Union

2013 – 41 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593 (print), ISSN 1831-9424 (online)

ISBN 978-92-79-34910-2 (pdf)
ISBN 978-92-79-34911-9 (print)

doi: 10.2788/49626

Abstract
The US NRC (2007) report, "Toxicity Testing in the 21st Century: A Vision and Strategy", represented a re-orientation of thinking surrounding the risk assessment of environmental chemicals. The key message is that by understanding Toxicity Pathways (TP) we could profile the potential hazard and assess risks to humans and the environment using intelligent combinations of computational and in vitro methods. However five years on, key questions surrounding a TP based approach to chemical risk assessment remain. In fact the very idea of 'Toxicity Pathway' still needs refinement to be useful in defining for example, how in vitro toxicity testing tools for chemical risk assessment. However, the issue is not that we need a better or clearer definition, but in fact we need to invest more time and energy in making sure that the concepts and consequences of TP thinking are clearly communicated and understood in the wider community. Only then can we expect more consistent use of the term, more comprehensive TP descriptions that address the key attributes that a TP should possess, and more conviction in the development and application of truly TP-based test systems for application in chemical risk assessment.

LB-NA-26389-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-34910-2