Universiteit Leiden
The Netherlands

Adverse outcome pathways: opportunities, limitations and open questions

Leist, M.; Ghallab, A.; Graepel, R.; Marcahn, R.; Hassan, R.; Hougaard Bennekou, S.; ... ; Hengstler, J.G.

Citation

Leist, M., Ghallab, A., Graepel, R., Marcahn, R., Hassan, R., Hougaard Bennekou, S., ... Hengstler, J. G. (2017). Adverse outcome pathways: opportunities, limitations and open questions. *Archives Of Toxicology*, *91*(11), 3477-3505. doi:10.1007/s00204-017-2045-3

Version:	Publisher's Version
License:	Licensed under Article 25fa Copyright Act/Law (Amendment Taverne)
Downloaded from:	https://hdl.handle.net/1887/3202104

Note: To cite this publication please use the final published version (if applicable).

Arch Toxicol (2017) 91:3477–3505
DOI 10.1007/s00204-017-2045-3

REGULATORY TOXICOLOGY

Adverse outcome pathways: opportunities, limitations and open questions

Marcel Leist[1] · Ahmed Ghallab[2,3] · Rabea Graepel[4] · Rosemarie Marchan[2] · Reham Hassan[2,3] · Susanne Hougaard Bennekou[5] · Alice Limonciel[6] · Mathieu Vinken[7] · Stefan Schildknecht[1] · Tanja Waldmann[1] · Erik Danen[4] · Ben van Ravenzwaay[8] · Hennicke Kamp[8] · Iain Gardner[9] · Patricio Godoy[2] · Frederic Y. Bois[10] · Albert Braeuning[11] · Raymond Reif[2] · Franz Oesch[12] · Dirk Drasdo[13,14] · Stefan Höhme[15] · Michael Schwarz[16] · Thomas Hartung[17] · Thomas Braunbeck[18] · Joost Beltman[4] · Harry Vrieling[19] · Ferran Sanz[20] · Anna Forsby[21,38] · Domenico Gadaleta[22] · Ciarán Fisher[9] · Jens Kelm[23] · David Fluri[23] · Gerhard Ecker[24] · Barbara Zdrazil[24] · Andrea Terron[25] · Paul Jennings[26] · Bart van der Burg[27] · Steven Dooley[28] · Annemarie H. Meijer[29] · Egon Willighagen[30,31] · Marvin Martens[30] · Chris Evelo[30,31] · Enrico Mombelli[10] · Olivier Taboureau[32,33] · Alberto Mantovani[34] · Barry Hardy[35] · Bjorn Koch[29] · Sylvia Escher[36] · Christoph van Thriel[2] · Cristina Cadenas[2] · D. Kroese[37] · Bob van de Water[4] · Jan G. Hengstler[2]

Received: 3 July 2017 / Accepted: 21 August 2017 / Published online: 19 October 2017
© Springer-Verlag GmbH Germany 2017

Abstract Adverse outcome pathways (AOPs) are a recent toxicological construct that connects, in a formalized, transparent and quality-controlled way, mechanistic information to apical endpoints for regulatory purposes. AOP links a molecular initiating event (MIE) to the adverse outcome (AO) via key events (KE), in a way specified by key event relationships (KER). Although this approach to formalize mechanistic toxicological information only started in 2010, over 200 AOPs have already been established. At this stage, new requirements arise, such as the need for harmonization and re-assessment, for continuous updating, as well as for alerting about pitfalls, misuses and limits of applicability.

✉ Marcel Leist
marcel.leist@uni-konstanz.de

✉ Jan G. Hengstler
hengstler@ifado.de

1 In Vitro Toxicology and Biomedicine, Department inaugurated by the Doerenkamp-Zbinden Foundation, University of Konstanz, Box 657, Universitaetsstr. 10, 78457 Konstanz, Germany

2 Leibniz Research Centre for Working Environment and Human Factors, Technical University Dortmund, Dortmund, Germany

3 Department of Forensic Medicine and Toxicology, Faculty of Veterinary Medicine, South Valley University, Qena, Egypt

4 Research Division of Drug Discovery and Safety, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

5 The Danish EPA, Copenhagen, Denmark

6 Division of Physiology, Department of Physiology and Medical Physics, Medical University of Innsbruck, 6020 Innsbruck, Austria

7 Department of In Vitro Toxicology and Dermato-Cosmetology, Vrije Universiteit Brussel, Laarbeeklaan 103, 1090 Brussels, Belgium

8 BASF SE, Experimental Toxicology and Ecology, Ludwigshafen am Rhein, Germany

9 Simcyp (A Certara Company), Sheffield, UK

10 INERIS, DRC/VIVA/METO, Parc ALATA BP2, 60550 Verneuil en Halatte, France

11 Department of Food Safety, German Federal Institute for Risk Assessment, Max-Dohrn-Str. 8-10, 10589 Berlin, Germany

12 Institute of Toxicology, University of Mainz, Mainz, Germany

13 INRIA, Unit Rocquencourt, B.P.105, 78153 Le Chesnay Cedex, France

14 Laboratoire Jacques-Louis Lions, France Université of Paris 06, CNRS, UMR 7598, 4 pl. Jussieu, Paris, France

15 Institute for Computer Science, University of Leipzig, Haertelstraße 16-18, 04107 Leipzig, Germany

16 Department of Toxicology, Institute of Experimental and Clinical Pharmacology and Toxicology, Eberhard Karls University, Tübingen, Germany

17 Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

In this review, the history of the AOP concept and its most prominent strengths are discussed, including the advantages of a formalized approach, the systematic collection of weight of evidence, the linkage of mechanisms to apical end points, the examination of the plausibility of epidemiological data, the identification of critical knowledge gaps and the design of mechanistic test methods. To prepare the ground for a broadened and appropriate use of AOPs, some widespread misconceptions are explained. Moreover, potential weaknesses and shortcomings of the current AOP rule set are addressed (1) to facilitate the discussion on its further evolution and (2) to better define appropriate vs. less suitable application areas. Exemplary toxicological studies are presented to discuss the linearity assumptions of AOP, the management of event modifiers and compensatory mechanisms, and whether a separation of toxicodynamics from toxicokinetics including metabolism is possible in the framework of pathway plasticity. Suggestions on how to compromise between different needs of AOP stakeholders have been added. A clear definition of open questions and limitations is provided to encourage further progress in the field.

Keywords Regulatory toxicology · Systems biology · Multi-scale integration · Computational toxicology · Interspecies extrapolation · Metabolism · Pathway unidirectionality · Liver fibrosis · Paracetamol · CCl_4 · Vinyl acetate · Tumor promotion · Binning of events · Multiple hit events · Proof of non-toxicity · Prioritization of compounds

Introduction

Outline of the AOP concept

The adverse outcome pathway (AOP) concept (Ankley et al. 2010) has attracted a large amount of attention in recent years, with more than 200 articles published on the issue in less than a decade. An AOP is defined as a sequence of events that begins with a molecular initiating event (MIE; e.g., inhibition of an enzyme or formation of DNA adducts) (Fig. 1). It progresses through a series of key events (KE) that define given cellular states. Such key events can be alterations of metabolic pathway, signaling events as well as modifications of organelles or cell functions (such as cell cycle control). Notably, an AOP may also be defined even

[18] Aquatic Ecology and Toxicology Section, Centre for Organismal Studies, University of Heidelberg, Im Neuenheimer Feld 504, 69120 Heidelberg, Germany

[19] Department of Human Genetics, Leiden University Medical Center, Postal Zone S4-P, PO Box 9600, 2300 RC Leiden, The Netherlands

[20] Research Programme on Biomedical Informatics (GRIB), Department of Health and Life Sciences, Hospital del Mar Medical Research Institute (IMIM), Universitat Pompeu Fabra, Barcelona, Spain

[21] Unit of Toxicology Sciences, Swetox, Karolinska Institutet, Forskargatan 20, SE-151 36 Södertälje, Sweden

[22] Laboratory of Environmental Chemistry and Toxicology, Department of Environmental Health Sciences, IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Via la Masa 19, 20156 Milano, Italy

[23] InSphero AG, Wagistrasse 27, CH-8952 Schlieren, Switzerland

[24] Pharmacoinformatics Research Group, Department of Pharmaceutical Chemistry, University of Vienna, Althanstraße 14, 1090 Vienna, Austria

[25] European Food Safety Authority (EFSA), Pesticide Unit, via Carlo Magno 1 A, 43126 Parma, Italy

[26] Division of Molecular and Computational Toxicology, Amsterdam Institute for Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, De Boelelaan 1108, 1081 HZ Amsterdam, The Netherlands

[27] BioDetection Systems b.v. (BDS), Science Park 406, 1098 XH Amsterdam, The Netherlands

[28] Section Molekular Hepatology, II. Medical Clinic, Medical Faculty Mannheim, University of Heidelberg, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany

[29] Institute of Biology, Leiden University, 2333 CC Leiden, The Netherlands

[30] Department of Bioinformatics, BiGCaT, NUTRIM, Maastricht University, PO Box 616, 6200 MD Maastricht, The Netherlands

[31] Open PHACTS Foundation, Cambridge, UK

[32] Inserm UMR-S973, Molécules Thérapeutiques in silico, Paris Diderot University, Paris, France

[33] Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[34] Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Roma, Italy

[35] Douglas Connect GmbH, Bärmeggenweg 14, Zeiningen, 4314 Aargau, Switzerland

[36] Fraunhofer Institute of Toxicology and Experimental Medicine (ITEM) Chemical Risk Assessment Group Manager Structure Activity Relationships/databases, and expert systems Nikolai-Fuchs-Strasse 1, 30625 Hannover, Germany

[37] TNO (Department of Risk Analysis of Products in Development), Zeist, The Netherlands

[38] Department of Neurochemistry, Stockholm University, Svante Arrhenius väg 16, SE-106 91 Stockholm, Sweden

Fig. 1 Schematic adverse outcome pathway (AOP) diagram. The schematic has been reproduced from the "Users' handbook supplement to the guidance document for developing and assessing AOPs (2017)" *MIE* molecular initiating event, *KE* key event, *KER* key event relationship, *n* number of KE of the AOP

though the MIE is not exactly known, as long as there is at least one clearly defined KE. Finally, the AOP leads to an adverse outcome (AO), i.e., a classical apical end point as used traditionally for hazard evaluation and risk assessment. According to OECD guidance (Guidance document on developing and assessing adverse outcome pathways 2017; Users' handbook supplement to the guidance document for developing and assessing AOPs 2017), each AOP should have one MIE and one AO, but there is no limitation in the number of KEs. Examples of AO include liver fibrosis, defined types of cancer, skin sensitization or parkinsonian-type motor deficits (OECD guidance document 2013). In an ecotoxicological context, typical AOs address reduced generalized fitness or a decline in reproductive success. One of the main purposes of AOPs is to connect existing knowledge on mechanisms of toxicity (in the form of KE) in a controlled and transparent weight of evidence process to apical end points (AO). The overall objective is to support regulatory decision-making, such as hazard identification and risk assessment, by formally describing the key events leading to an AO. This may pave the way to the use of KE-related end points for decision-making. These surrogate end points may be used in addition to classical apical end points to strengthen the rationale of predictions and decisions or they may in future even be used instead of classical end points, as soon as the causal relationship of KEs and AO is sufficiently established. As such, the AOP concept can be regarded as a knowledge management tool that facilitates the use of mechanistic knowledge for regulatory purposes in toxicology and ecotoxicology. In this article, members of the EU-ToxRisk consortium and other experts in the field collected and analyzed information on the history of the AOP concept, the most prominent strengths, advantages and limitations, as well as possible ways to overcome such limitations.

Conceptions and misconceptions concerning AOPs

The AOP concept is a relatively new tool and as such provides numerous opportunities (Table 1). As always in such situations, it is important to have a good understanding of the concept to make the right use of it, i.e., to use it in situations in which it is well suited and to be careful of not overusing it in situations that have shortcomings. The concept also allows the development and optimization of new applications on the basis of both the understanding of the strengths and limitations of the AOP construct and on the experience gained from its use. To exemplify this, one may think of an AOP as a pair of plyers. They may be used to pull out a nail. A completely different, but also perfect use would be to turn, e.g., a nut. As the plyers are made of steel, they may also be used to hammer in a nail, or as a lever to wedge something away, but this would be less ideal and there would be better tools for such applications. Plyers also look a bit like scissors, but they obviously have a different function.

AOPs are not signaling pathways. AOPs look to some extent like a drawing of a signaling pathway and share some of its features, but there are also important differences. One difference is the perspective taken. A classical signaling pathway starts at the signal initiation, and then often leads to signal amplification and splitting to arrive at multiple targets. Likewise, many academic researchers in the field of molecular toxicology are also interested in the initial target of a toxicant (e.g., the Ah receptor) and then follow the various possible downstream consequences, but the exact link to an AO is often unclear. The AOP concept takes the opposite perspective. The anchoring point is the AO and each KE is examined concerning its relevance and necessity for the AO. A major intention is to link earlier KEs of an AOP to the AO, even if the starting point is unknown or some intermediate KEs are not (yet) fully understood.

AOPs are compound agnostic. AOPs describe a sequence of biological events, and the transitions from event to event that lead to an AO. This is independent of any specific chemical. Thus, AOPs do not describe the mode of action of a chemical. However, a chemical may activate a specific AOP by provoking the MIE, or it may be used to provide empirical evidence for the existence of an AOP.

AOPs describe toxicodynamics, not toxicokinetics and therefore not risk. As AOPs are 'compound agnostic' they cannot incorporate a toxicokinetic phase. ADME properties and toxicokinetics cannot be uncoupled from specific chemicals, while AOPs do not describe specific chemicals. Thus, AOPs as such cannot be used for risk assessment, but they describe a potential hazard, given that a specific compound at the concentration, which is reached at a relevant site, triggers the MIE that activates the AOP. To use AOP information for risk assessment, it should be imbedded together with

Table 1 Opportunities, shortcomings and open questions of the AOP concept

Characteristics of AOP	Comments, examples and next steps
Opportunities/strengths	
Systematic collection of weight of evidence	Introduction of 'systematic reviews' into toxicology (contrary to 'narrative reviews' and 'cherry picking' of evidence and mechanisms
Sorting of information across multiple levels of organization	The concept goes beyond classical 'signaling' pathways: bridging scales of organization from organelle to tissue/organism
Link of mechanistic end points to apical end points	Provides anchoring of mechanistic information to known apical end points: rationale for use of mechanistic data in regulatory contexts
Convenient structure to visualize gaps of knowledge	Useful to direct research needs; important for definition of uncertainty of mechanistic information
Consensus and certification process for canonical knowledge	Knowledge aggregation by a broad community (Wiki-process; multiple review levels); provides plausibility to epidemiological data
Allows to examine plausibility of epidemiological data	Toxicological relevance of epidemiological data is strengthened, if robust AOPs explain how exposure to an agent can lead to an AO
Guidance for use and design of mechanistic test methods	The defined (causal, plausible, in part quantitative) links of key events (and respective assays) to adverse outcomes allow the use of key event tests for regulatory filtering, prioritization and decisions
Shortcomings (with respect to some potential applications)	
Linearity and unidirectionality rules	Biology frequently uses feedforward and feedback
Binning of large groups of events	The MIE and KE of AOP often comprise dozens or even hundreds of individual biochemical events (e.g., inflammation, oxidative stress)
Dealing with multiple hit events	Several major pathologies require multiple hits and cannot be explained by one MIE
How to deal with effect duration	The time component is not reflected in the AOP scheme: it does not allow for a display of effect duration (or time window) required
Neglect of event modifiers in graphical displays	Although considered in the AOP handbook, and in the text parts, event modifiers cannot be displayed in the AOP scheme
Dealing with compensatory mechanisms	Considerations of the counter-regulation would need to be part of an AOP. Although considered for KER, display possibilities are limited
Alerts/open questions/misconceptions	
Defining non-toxicity	Will the AOP system get comprehensive enough? Can we exclude with sufficient certainty activation of an AOP?
Quantification of KE/KER; threshold setting	Few thresholds fully known; hard to translate between in vitro and in vivo systems
Review process	The classical peer review process of journals is hardly suited for the review of such large documents. How can peer review be achieved?
Quality assurance and validation	How can AOP-based test batteries be validated? How good are the literature data used for AOP construction?
One procedure for many different complexity levels	Can one apply the same tools for different concepts as 'aromatase inhibition' (biochemical step) vs 'steatosis' (complex pathology)?
Can toxicodynamics be separated from toxicokinetics?	Is generation of reactive electrophiles part of ADME or of an AOP? Can the ADME phase always be separated from the AOP phase?
How to deal with pathway plasticity	In pathological situations, pathways themselves may change (bypass of KE, activation of steps not normally present)
How to deal with species differences	AOP describe species effects, but in practice most data are derived from rodent systems; species specificity is not always known
AOPs are a snapshot of knowledge at a given time	How is continuous updating ensured? Who decides on updates?
Applicability fields	AOPs are working tools, made for specific purposes. They cannot be used blindly outside their main field of application
AOP vs systems biology models	AOP can be more or less quantitative, but the concept is fundamentally different from systems biology models. AOP describe a transition of system states (not computable relationships between system modules)

ADME considerations in an integrated approach to testing and assessment (IATA) of potential toxicants (Tollefsen et al. 2014).

AOPs are multi-scale data integration tools. Unlike many signaling pathways, AOPs span many levels of biological organization, from molecular over sub-cellular, cellular to organ and organism or, especially in the case of ecotoxicology, even to the population level. The anchoring point is the AO relevant for the entire organism, while the MIE always is a molecular-level interaction. The AOP provides a scaffold to integrate data at such different scales, all with a general anchoring toward predictions of adverse outcomes. Notably, the rules of AOP construction not only allow integrating the data, but also provide a scaffold for sorting and weighing of contradictory or inconsistent data through a weight-of-evidence process. In parallel, multi-scale quantitative models of toxicity are being developed (Obiol-Pardo et al. 2011).

AOPs are not intended to be molecularly detailed. It is often not clearly understood that AOPs are meant to describe a series of system states with some information on the transition between these states. Importantly, AOPs do not intend to comprehensively describe and quantify all system parameters. For practical reasons, often relatively superordinate processes, like mitochondrial function/dysfunction or intact/disturbed calcium ion homeostasis, are used to describe the system states. This allows measurement of the system state with relatively simple and accessible methods, and thus has large practical advantages for predictive toxicology. In this sense, the concept differs from molecularly defined pathways of toxicity that require complex and comprehensive tools for their quantification (Kleensang et al. 2014; Sauer et al. 2015; Rahnenführer and Leist 2015). A mapping of molecular pathways of toxicity (Hartung and McBride 2011) by omics technologies instead of the literature compilation typically used for AOP was pioneered by the Human Toxome Project (Bouhifd et al. 2015; Hartung 2016).

AOPs are not systems biology models. While AOPs span multiple levels of biological organization, most systems biology models still have large difficulties integrating different scales of organization, although some projects have recently achieved this goal (e.g., Ghallab et al. 2016; Schliess et al. 2014; Drasdo et al. 2014; Carbonell et al. 2017). One of the practical limitations of systems biology models is that their establishment is usually labor intensive. A systems biology model often has a network character, while AOPs are linear. Likewise, AOP networks, defined as sets of AOPs sharing at least one common element, are different from systems biology networks. AOP networks are intended to represent potential chemical effects. They are the tools for real-life application, unlike for individual AOPs. They may be focused on a single effect (e.g., steatosis), but describe all the pathways leading to this effect (e.g., multiple MIEs, crossing of KE). They provide information on interactions between AOPs and have the potential to reveal previously unknown links between biological pathways. Currently, it remains challenging to translate AOPs into mathematical models with ordinary differential equations, but research activities in this field are ongoing (Obiol-Pardo et al. 2011). Some mathematical models of AOPs appear possible and they may be regarded as transition stages to systems biology models (https://www.effectopedia.org/). It should be taken into account that system biology and AOPs are complementary tools. The AOP is intended for regulatory use, minimizing complexity to serve pragmatic decisions that so far are mainly based on animal studies. A future perspective is that systems biology models would be integrated into AOPs and informing risk assessment.

AOPs can provide plausibility for statistical associations. Most toxico-epidemiological research deals with statistical associations of chemical exposures and adverse outcomes. In most cases, it is difficult or impossible to clarify the causality of such correlations. One step toward this direction would be to provide a plausible link between exposure and adverse outcomes. AOPs, coupled to exposure and toxicokinetic information may provide such links by coupling exposure to the triggering of an MIE. This approach may support environmental risk assessment.

AOPs can identify knowledge gaps, indicate testing flaws and may guide testing. An important practical application of AOPs is that they can guide the development of novel, mechanism based and often animal-free methods of hazard assessment. In this sense, the AOP framework constitutes a basis for the development of new non-animal test methods (i.e., in vitro methods anchored to KEs) and integrated testing strategies (Hartung et al. 2013a, b; Rovida et al. 2015; Clippinger et al. 2016). It also provides biological context for mechanistic information from existing assays. In many cases, AOPs may be combined in AOP networks that share at least one KE. Analyses of these AOP networks can aid in the prioritization of assay development, whether the goal is to develop a single assay with predictive ability for multiple outcomes or development of assays that are highly specific for a particular mode of action. Appropriate IATA or other types of integrated testing strategies (ITS), based on AOP (AOP networks) and toxicokinetic information, may in future use KE information to develop mechanistic assays (ideally based on human cells) to predict adverse effects for humans at high throughput and with low uncertainty related to interspecies extrapolations (Tollefsen et al. 2014; Jacobs et al. 2016; Paparella et al. 2017).

The history of the AOP concept

The AOP concept, introduced in 2010 as a tool to support risk assessment in ecotoxicology, is based on the long-standing principle that understanding the mechanisms linking

initial events to adverse effects is a key aspect of (predictive) toxicology (Ankley et al. 2010; Villeneuve et al. 2014a, b). Does this concept have a history beyond the short time span of the last 7 years? For more than 100 years, toxicologists have been well aware of the importance of understanding the mechanisms linking initial events, e.g., DNA damage or protein adducts, to adverse effects, such as cancer or fibrosis. Before 2010, such mechanistic explanations have not been termed AOP. Before that date, they were often mentioned as toxicity mechanism, mode of action or toxicity pathway and these terms included several key aspects of the current AOP concept. Likewise, the `biomarker` concept originally designed in the 1990s (Beliaeff and Burgeot 2002; Depledge 1994; Huggett et al. 1992; McCarthy and Shugart 1990) contained major aspects of the current understanding of KEs. In fact, such thinking even goes back to the nineteenth century. For example, Louis Lewin (1850–1929) was aware of the advantages of classifying toxic compounds based on their pharmacologic or toxic mechanisms that led to adverse effects, as demonstrated by his grouping of psychoactive drugs into five categories (Lewin 1885, 1924). Although in Lewin's time it was not yet feasible to study molecular mechanisms, the idea that a mechanism which explains an adverse effect allows a better classification than a simple description of toxic symptoms was already born. Since then, the improvement of risk evaluation by understanding of the toxicological mechanisms has been a focus of generations of toxicologists. A major milestone was the establishment and regulatory acceptance of the Ames test in the 80s of the last century. This test method would nowadays be coined as a KE assay for genotoxicity that predicts the AO of genotoxic carcinogenesis. A further key process was the work on KE tests relevant to skin sensitization. This resulted in the first AOP accepted for regulatory purposes, but the conceptual and experimental work was mostly planned and executed before the term AOP was coined. A further event preparing the ground for AOP was then the report of the US National Academy of Sciences in 2007, which suggested a reorientation of toxicology, away from apical end points, moving toward mechanistic principles and introducing the toxicity pathways as a guiding principle (Leist et al. 2008; NRC 2007).

In the light of this historical perspective, the question arises: What was new about the AOP concept? Three key aspects need to be mentioned here:

A first aspect is that AOP aims for a more formalized approach to risk assessment than previous mechanistic toxicology. Even the toxicity pathways of the NAS strategy paper were still relatively little defined (Andersen and Krewski 2009, 2010). AOPs are by definition systematic, structured, quality controlled and weight of evidence based. One objection may be that this should apply to any toxicological science and can therefore hardly be considered a novelty. Examples from other fields may help to convey the concept that novelty can indeed lie in the methodology and its stringency: Henry Ford was a great innovator in the automobile world, as was Steve Jobs for consumer electronics. They both mainly perfected methods, based on already known basic knowledge. By analogy, there is a chance that the AOP concept will be pivotal for the application of mechanistic principles in risk assessment. So far, mechanistic data have rarely been considered in the assessment of industrial chemicals (with the notable exception of genotoxicity) and this has prevented the use of large amounts of available information by regulators.

A second important novelty is that large regulatory agencies have adopted the concept of mechanistic toxicology (Leist et al. 2014; Edwards et al. 2016). The AOP concept is supported by the OECD, and this greatly facilitates its implementation into regulatory thinking. The consequence will be an integration of existing knowledge on toxic mechanisms into regulatory procedures. A secondary consequence is that the European Commission and other funding agencies have realized the need to support research projects that define and implement AOPs (Daneshian et al. 2016; Daston et al. 2015).

This leads to a third important issue: new ways of safety assessment are supported by the AOP concept. While mechanistic information traditionally has largely been seen as a supplementary for better explaining and rationalizing apical end points of toxicity, the AOP concept now facilitates the use of a conclusive chain of mechanistic information alternative to a classical toxicological end point. This also means that new test systems, including in vitro approaches, are about to reach a new level of importance and recognition. It is very likely that an in vitro test battery or *in silico* methods, based on, or linked to key events in a relevant AOP, will be more likely to gain regulatory acceptance than tests not linked to or supported by an AOP. A first example of this was seen with the testing for dermal sensitization.

Further developments

As introduced here, the AOP concept has seen a dynamic development and as of June 2017 the AOP knowledge base maintained by the OECD lists 204 AOPs (with 1477 key events). An easy access to these AOPs is provided by the AOP-wiki (https://aopwiki.org/), where one may also follow the development stages, which are overseen by the Advisory Group on Molecular Screening and Toxicogenomics (EAGMST). While these are positive developments, there is still a need to further advance and improve the AOP concept. It is essential to understand and clearly communicate how an AOP is defined and what it can be used for. On this foundation, the AOP concept can be further refined and adapted to future needs that are not yet covered. Most importantly, open

questions are highlighted to guide future development and appropriate solutions to the limitations (Table 1).

Strengths and advantages of AOPs

Systematic collection of weight of evidence

The assembly of information on a scientific topic is often based on reviews written by experts in the field. Such reviews can be qualified according to their (1) contents, (2) form and (3) quality control procedure. Classical narrative reviews with contents similar to what may be found nowadays in an AOP have been common in toxicology for decades. An advantage of the AOP concept is that it now also provides very clear guidance on form and quality control. Strong emphasis is put on weight-of-evidence approaches to prevent personal bias. The large impact of such a formal process has been demonstrated in the field of medicine by the Cochrane collaboration and their 'systematic reviews and protocols' on health care. This concept has been adopted for toxicology by the Evidence-Based Toxicology Collaboration (Hoffmann and Hartung 2006), which addresses the quality scoring of existing studies (Samuel et al. 2016) and systematic reviews (Stephens et al. 2016; Hoffmann et al. 2017). The AOP guidelines have particularly optimized this procedure for toxicity pathways. Adherence to the rules should guarantee and may possibly improve the description of mechanisms of toxicity.

Certification process for canonical knowledge

In many disciplines of biomedical research, including toxicology, the definition of canonical knowledge is an implicit process, i.e., it is embedded in the discipline without defined procedures. The collection of such knowledge is a slow process that involves opinion leaders, textbook authors and others. The definition and updating of the state of the art can take relatively long periods of time to reach acceptance, and this process has worked well for basic science. However, for applied and regulatory science, a more explicit and definitive process is desirable. For regulators it is of high importance to have clarity on what is considered the scientific gold standard. Here, the process of AOP establishment and validation can be extremely helpful. The weight-of-evidence procedure, coupled to multiple levels of peer review, helps to define, certify and update canonical knowledge on a global (or at least OECD-wide) scale.

The view of an AOP database as a growing repository of canonical knowledge has implications for several areas. An application that springs to mind immediately is the support of mode-of-action information in regulatory dossiers, which will become more and more important. However, there are other areas of interest. For instance, an additional level of information can be added to read across approaches (Patlewicz et al. 2014; Ball et al. 2016) by comparing not only chemical structures of a group of compounds of interest, but also their cellular responses, i.e., by examining whether similar molecular targets and key events are involved (Zhu et al. 2016).

Sorting of information across multiple levels of organization

AOPs are particularly designed to allow integration of information spanning from the molecular level, over sub-cellular, cellular and organ levels to the AO at the level of entire organisms (e.g., humans) or even populations or ecosystems (e.g., in ecotoxicology). Many other approaches have particular strengths and track records in dealing with information from a single organizational level, and past experience has shown that bridging of levels of complexity in organization is a particularly difficult scientific task, be it in physics, biology or toxicology. Although AOPs do not claim to provide all elements of a level-spanning pathway, they have the ambition to cover the main steps and to provide a clear rationality for the plausibility, causality and strength of linkage.

Link of mechanistic end points to apical end points

The issue above implies that mechanistic end points are connected to apical end points. On first sight, this does not appear to be a particularly new concept and advantage. However, taking into account the practical situation of mechanistic toxicology, this apparently small point represents one of the most advantageous features of AOPs: There are some (but few) examples that mechanistic end points have been used in the past by regulators together with, or instead of, classical apical end points. They mostly refer to the area of genotoxicity and this area has shown how successful this concept can be. Countless 2-year rodent carcinogenicity bioassays have been avoided by the use of the Ames test or similar genotoxicity assays. These test methods are well accepted, because there is a well-accepted chain of events that links the mechanistic end point of DNA mutations to the apical end point of carcinogenesis. In contrast, in most other domains of toxicology, mechanistic end points have been used to a very limited extent for regulatory purposes. One of the reasons of this limited use is that the interpretation has often been difficult and the linkage of the mechanistic end point to the apical end point has not been sufficiently defined. For some areas, like developmental neurotoxicity (DNT), it has been widely assumed that mechanistic end points and in vitro models can never cover the in vivo complexity (Kuegler et al. 2010; van Thriel et al. 2011). The

AOP concept now provides quality-controlled and formalized tools to generate such linkages and therefore opens entirely new approaches to hazard assessment and eventually risk assessment. The possibility to regulate on the basis of mechanistic end points is a major driving force for in vitro toxicology and methods development.

Examination of the plausibility of epidemiological data

A particular version of the link of mechanistic end points and adverse outcomes has been examined by the European Food Safety Authority (EFSA; Ockleford et al. 2017; Pelkonen et al. 2017). In general, it is difficult to deal with and to react to epidemiologic data that suggest there is a statistical correlation between exposure to a chemical and certain adverse health outcomes. The causality of effects can hardly ever be proven on the basis of epidemiological approaches. However, one step ahead is the examination of whether there is at all a plausible link between a given chemical and the purported health outcome. The plausibility of such links is confirmed if there is an AOP that connects MIE triggered by chemicals at question to a relevant AO. In a further step, information from the AOPs may then be used (1) to examine whether the observed exposure would result in sufficiently high internal concentrations to trigger such an AOP, (2) to develop or recruit test methods that can quantify early KE or the MIE to provide testing based on easily accessible mechanistic end points and (3) to define gaps of knowledge that may be filled by dedicated research programs.

Visualization of knowledge gaps

The issue of research gaps can be even further generalized, as it does not only apply to epidemiological data. Many AOPs have been put together from the published literature. Mechanistic toxicology publications often examine one mechanistic step (i.e., one KER) or the link of one particular KE to an AO. However, few publications address an entire AOP, and thus there are often gaps of knowledge between some KEs of an AOP. Therefore, AOP developers are recommended to consider the tailored Bradford Hill criteria while assessing newly postulated AOPs (Becker et al. 2015). Obviously, these are also important gaps that limit the application of AOPs for hazard assessment, as the overall scientific confidence (strength) of the AOP is reduced, if not all connections are supported by strong and preferably quantitative exemplar data. This situation highlights the overall gaps in our knowledge and can guide future research in several areas: first, the causal relationship between the respective $KE_{upstream}$ and $KE_{downstream}$, starting from the MIE and leading to the AO. The quantitative aspects of these KERs at specific exposure levels will gain more and more importance in the future. Second, based on such quantitative data, the similarity between compounds with respect to AOP activation can be determined. Such similarity along an AOP can form an important basis for read across, as it adds valuable information to that of purely chemical similarity. Third, a new wave of research will go beyond the KE as such and address the relative importance of modulatory events. In all likelihood, there will be a fruitful interaction between the improved construction of an AOP and the clarification of general gaps of knowledge concerning modulatory events of AOP KEs. Filling these knowledge gaps is an important step toward the construction of quantitative AOPs (qAOP).

Design of mechanistic test methods

An important practical advantage of AOPs is that they allow to define test methods for KEs and that these results can in principle be related to classical apical end points of hazard assessment (AO) via the AOP definition, if sufficient knowledge is available. Such tests provide, for example, a relevant way for the prioritization of relevant chemicals. This is of interest in early phases of drug development (Hartung 2017) and also in large-scale toxicity screens of chemicals. The data obtained in this way can be used as a decision base for further testing, for further development or for risk management strategies. The vision is that MIE/KE-based test batteries would substitute at some point classical end point testing. This strategy has been applied for instance in the area of skin sensitization and it is widely used in the area of genotoxicity (Basketter et al. 2012). An example from ecotoxicology is the use of vitellogenin modulation as a conclusive indicator of estrogen-related endocrine disruption (Matozzo et al. 2008).

The concept of mechanistic tests as filters for toxicological decisions (also called ‘pre-tests’) has a long history. Pre-tests were introduced more than 50 years ago in specific fields of toxicology, e.g., based on the knowledge that mutations represent a key event on the path toward developing cancer (Bhattacharya 1948; Ames and Whitfield 1966). In the past decades, this approach has allowed for the waiving of numerous unnecessary in vivo carcinogenicity studies. For instance, when bacterial and mammalian cell mutagenicity tests result in clear-cut positive results, it is usually not necessary to perform a 2-year carcinogenicity study in rats or mice, because of the high probability of a positive result (Basketter et al. 2012).

Genotoxicity tests are an example for a selection of a KE common to many AOPs. Multiple MIE converge on the KE of DNA mutation or chromosome breaks. The design of tests for such KE is particularly efficient for early identification of safety hazards, as large groups of chemicals can be captured. It is likely that the first larger KE-based test batteries will strongly focus on such ‘promiscuous’ or ‘joint’ KEs (i.e., KE

that are common to various AOP). Another typical example is mitochondrial toxicity. It represents an easy-to-measure KE and is clearly linked to various AOs (Hamon et al. 2014; Wilmes et al. 2013; Limonciel et al. 2011). If a chemical reaches the toxicity target cells in a tissue at concentrations that compromise mitochondrial function, then there is a high likelihood that adverse effects would ensue at the cell and organ level. For this reason, mitochondrial dysfunction tests are being developed in many projects. A relevant goal of the immediate future is to define a comprehensive set of KEs measureable in vitro, which have adverse consequences when present in vivo. Examples include predictive systems of developmental toxicity mediated by chemicals that inhibit migration of neural crest cells (Zimmer et al. 2012, 2014; Smirnova et al. 2014; Nyffeler et al. 2017) that block neurite growth (Stiegler et al. 2011; Krug et al. 2013b) or compromise the regulation of genes normally up- or downregulated during normal embryonic development (Krug et al. 2013a; Rempel et al. 2015). Other important examples are interferences with nuclear receptors and developmental signaling pathways (van der Burg et al. 2015a, 2015b; Bal-Price et al. 2015a, b, 2017), the generation of oxidative stress (Schildknecht et al. 2009, 2011, 2017; Krug et al. 2014) and endoplasmic reticulum stress.

In the context of research on new chemical entities (drugs, pesticides, etc.), such early testing has been used to guide chemical design strategies: For instance, it implied the development of alternative structures, which have the same desired effect, while avoiding the toxic effect identified in mechanistic in vitro tests, an approach called Green Toxicology (Maertens et al. 2014; Crawford et al. 2017). A challenge of these is the adequate integration of metabolic activation and detoxification into the testing strategy to perform adequate in vitro-to-in vivo extrapolations (IVIVE). As a consequence, there is a need for relevant kinetic models to determine target tissue concentrations of chemicals (Tsaioun et al. 2016). If these issues are solved, organ toxicity may be predicted and animal experiments may be avoided.

Anchor point for computational toxicology predictions

There are multiple ways to use the information embedded in AOPs for toxicity predictions, and the development of such methods represents an emerging field. One exemplar approach is the identification of experimental test methods that are associated with KEs and in particular with MIEs. Then, large databases (e.g., Tox21) can be screened for detecting compounds that are hits in such assays. Such compounds are of particular concern, as they are likely to activate the respective AOP. A variant of this approach is the *in silico* screening of all new compounds in an industry (e.g., cosmetics and consumer products) against their interaction with all known MIEs of AOPs. According to this, compounds of unknown hazard can be assigned to various concern levels (e.g., 'high concern' if they are likely to strongly activate several MIEs; or 'low concern' if no known MIE is activated). Based on this knowledge, chemicals could be classified according to their biological fingerprint (profile of biological activation). Notably, this approach will gain momentum and impact with increasing numbers of AOPs, as then a larger fraction of all possible hazards will be covered. A second example for exploitation of AOPs for computational toxicology may be based on transcriptome data. Most MIEs and KEs are associated to certain genes, e.g., if an MIE is 'binding to the androgen receptor (NR3C4)', then it is associated with the corresponding androgen receptor gene. Transcriptome mapping of toxicant responses also yields lists of genes. Computational and statistical methods can be used to probe for overrepresentation of certain genes among toxicant-induced transcriptome changes. This approach can be used to identify AOP-associated genes and develop new AOPs or improve existing ones, or to determine a potential mode of action from transcriptome data.

Recently, so-called cpAOPs (computationally predicted AOPs) have been described, which are built by integrating openly available HTS in vitro data and information on the disease phenotype (Oki and Edwards 2016). Further, mapping gene expression data onto disease data revealed interesting new causalities via the use of network analyses. This approach aims to unravel mechanistic relationships merely from the available data, without any prior knowledge about a certain AOP. Thus, in the near future data science will likely offer a wealth of orthogonal strategies to biological approaches for building AOPs.

Improved predictions on toxic responses for humans by interspecies extrapolation

As mentioned above, the underlying mechanistic concept of AOPs has been in use for decades, although a different nomenclature was used (Leist et al. 2008). An important application in the past was to improve predictions on toxic responses for humans by interspecies extrapolation. The mechanistic understanding was also used to better understand the basis of interindividual differences within human populations. An important example for the former point is the explanation why humans are relatively resistant to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD)-mediated toxicity, compared to certain mouse strains like C57BL6/N and certainly not as susceptible as guinea pigs (Hengstler et al. 1998). More recent research has also shown that this species difference applies to developmental toxicity due to altered expression of the target for the MIE (Baumann et al. 2016; Bal-Price et al. 2015a; Gassmann et al. 2010). The concept has also been used to explain why humans are several orders of magnitude less susceptible to tamoxifen-induced

carcinogenesis than rats and mice, and why humans are more susceptible to digitoxin compared to mice due to MIE interspecies differences regarding the inhibition of the Na/K-ATPase (all cases reviewed in Hengstler et al. 1999). These examples demonstrate that understanding the mechanisms causing toxicity helped to extrapolate from animal models to humans. They also helped to understand the molecular basis of interindividual differences regarding the susceptibility of humans to toxic substances, e.g., by definition of genetic variants of genes relevant at various positions of the AOP chain (Hengstler et al. 1999). Important cases of species differences are found in the field of endocrine disrupters. Here, the targets are not altogether different, but relative expression levels and fine modulations of regulation show pivotal differences. In this field, a large step has been the design of human cell-based test batteries, involving an AOP structure and computational modeling to substitute existing animal tests, such as the uterotrophic assay (Casey 2016; Kleinstreuer et al. 2016, 2017).

Limitations and shortcomings at the present stage of AOP development

It was only 10 years ago that in 2007 toxicity pathways were suggested by the National Academy report (NRC 2007, Leist et al. 2008, Andersen and Krewski 2009) as a basis for risk assessment. Several related concepts, such as pathway of toxicity, the human toxome or biomarkers of toxicity (Blaauboer et al. 2012) arose from this idea. The specific AOP concept itself was only presented as recently as in 2010 by Ankley et al. (2010; Becker et al. 2015; Knapen et al. 2015; Wittwehr et al. 2017). Despite the rapid development since then, it should not be surprising that the AOP concept, and its handling by regulatory bodies, is still in a dynamic phase. This implies continuous optimizations and necessary adaptations, also including a learning-by-doing process. To support the further development, it is important to highlight the advantages, but also to pinpoint areas that should be reconsidered and optimized. Finding the appropriate balance requires a continued discussion process and revision of the respective current state of knowledge and demands to the system. Awareness of potential weaknesses and understanding of the respective reasons for them allows the optimal handling of the tool by various users. For this reason, some present weaknesses, perhaps to be amended soon, as well as others to be considered on the long term are listed below.

The assumptions of unidirectionality and linearity

Typically, AOPs are presented as a unidirectional chain of events from MIE over KEs to AO (Fig. 1) (Ankley et al. 2010). Moreover, the preferred mode of graphical presentation is a linear array of KEs. This has practical advantages, which are extensively discussed, e.g., by Villeneuve and colleagues (Villeneuve et al. 2014a, b). It has been clear to the developers of the concept that this unidirectional linearity is not always in agreement with biology. Especially degenerative processes often bear features of vicious circles (Schildknecht et al. 2017; Nicotera et al. 1999) or require multiple hits (Hengstler et al. 2012). Although users and developers of AOPs are aware of the problems, the ideal solution has not yet been found. Further developments will be necessary to make the AOP concept useful for all applications in regulatory toxicology. As a discussion starter, it is important to review briefly the alternatives to a linear and unidirectional AOP:

1. A toxicity pathway may involve positive or negative feedback loops. This situation is found quite frequently. For instance, one of the most important principles of metabolic regulation is negative feedback: the negative control of early steps by products of late steps. Positive feedback loops are typically found in the regulation of signaling pathways or in proteolytic cascades. This leads to an avalanche effect, once the cascade has been triggered;
2. Non-linearity may occur in at least two major variants. Either the AOP bifurcates, or another pathway joins the AOP. An example of the former case is the recently established AOP on mitochondrial complex I inhibition, leading to parkinsonian motor deficits, where a bifurcation leads to neuroinflammation as positively enhancing feedforward event (https://aopwiki.org/aops/3). Examples of the latter case are multi-hit events, where, e.g., ongoing DNA damage triggered by an electrophilic compound/metabolite, or by oxygen radicals, joins the path of events triggered by a tumor promotor;
3. An AOP may appear to be linear and unidirectional, but their nodes (KEs) get input not only from the upstream KEs, but also from various modulatory events. These inputs are integrated in complex (often not simply additive) ways (e.g., by allosteric regulation or by repair processes against damage), and the final outcome determines whether the next (downstream) KE is ever activated.

Especially, feedforward loops and pathway synergy can be hard to handle in a quantitative manner. Systems biology has created tools to address these issues appropriately, but such approaches require quantitative modeling, and they are not reflected by simple arrow schemes. Non-linear multi-step processes are not only found in carcinogenesis, where they are well documented, but are also relevant to organ toxicity. Besides steatosis, and other forms of liver damage, many inflammation-related reactions are

non-linear because of structural non-linearities (in addition to possible non-linearities in the quantitative relationship between two successive KEs). For instance, the experimental liver toxicant concanavalin A can activate lymphocytes. Lymphocytes can then activate Kupffer cells to produce tumor necrosis factor (TNF) and Kupffer cells can further activate lymphocytes. This sequence can lead to a strong enhancement of TNF release compared to direct Kupffer cell stimulation, and this huge augmentation has a key role in the ensuing liver damage (Gantner et al. 1996).

First, attempts have been undertaken to address the issues of linearity. For instance, a positive feedback loop has been incorporated into the complex I-parkinsonism AOP (Fig. 2) and a negative feedback loop into the cholestasis AOP (AOP-wiki Nr. 27) from deregulated transporter proteins back to bile acid accumulation. It has been acknowledged that a feedback response can be causally linked to the AO. If it is measureable and fits a linear sequence, the present rules allow it to be included as KE in the AOP (see e.g., https://aopwiki.org/aops/23); in other cases, new solutions are required. One of these novel approaches has been chosen by the developers of the protein alkylation-liver fibrosis AOP (Horvat et al. 2017). For the core AOP, they suggest a canonical linear and unidirectional AOP (Fig. 3a), but also discuss that several steps are modulated by oxygen radicals and inflammation. Then they present a more complex scheme with feedback loops for the general overview (Fig. 3b). Perhaps, one solution for the future will be a demand-driven practical approach to produce a series of AOP variants with differing levels of complexity, adapted to the respective application.

The new version of the AOPs User's Handbook (in progress, but not yet published as of July 2017) will also open some new possibilities. It is stipulated that representation of branching in an AOP will be acceptable, especially when there are multiple KEs, causally linked to the MIE and AO that occur concurrently and likely act in concert to drive the downstream effects. In such cases, the various KEs do not have to be necessarily placed neatly into a single temporal sequence, because they effectively occur simultaneously. Likewise, it cannot necessarily be determined which of the concurrent KEs are most essential or critical, because there are multiple KEs (measurable biological changes) contributing jointly in an additive manner such that it cannot be effectively determined whether one could cause the pathway to progress without the other. This is contrasted with cases where KEs act independently such that one event or the other, alone, would allow progression toward the outcome.

Neglect of event modifiers in graphical displays

As detailed above, downstream KEs do not only get input from the upstream KEs, but from multiple other directions. In initial versions of the AOP guidance, there was more emphasis on such modulatory events, and the graphical display of modifiers as an integral part of the AOP was considered. The removal of modulatory events in graphical representations, as suggested by the current handbook and website procedures, has led to a simplified display, and this provides a faster and easier overview. However, the downside is that major biological information is lost from the display. It is stipulated in the handbook that this information can be added as text. For instance, the information is

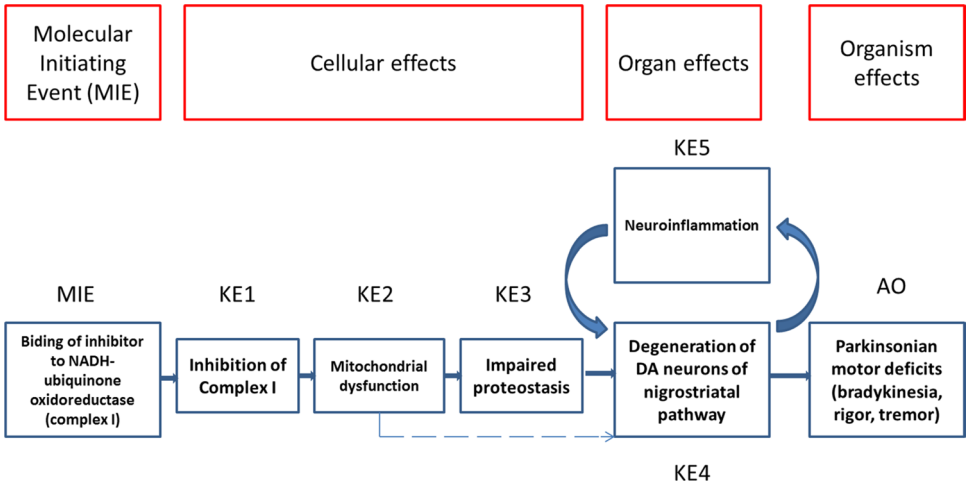

Fig. 2 Example of a recently proposed AOP aiming to describe critical steps from binding of a toxicant to mitochondrial complex I to parkinsonian motor deficits. The schematic has been reproduced from the AOP website (http://aopwiki.org/aops/3). It exemplifies how a positive feedback loop has been incorporated into an AOP (KE5 feeding back on KE4) that is thus non-linear. This was introduced as it is not possible on the basis of available data to clearly define whether KE4 or KE5 is more upstream. Strong scientific evidence supports a vicious circle at this position of the AOP

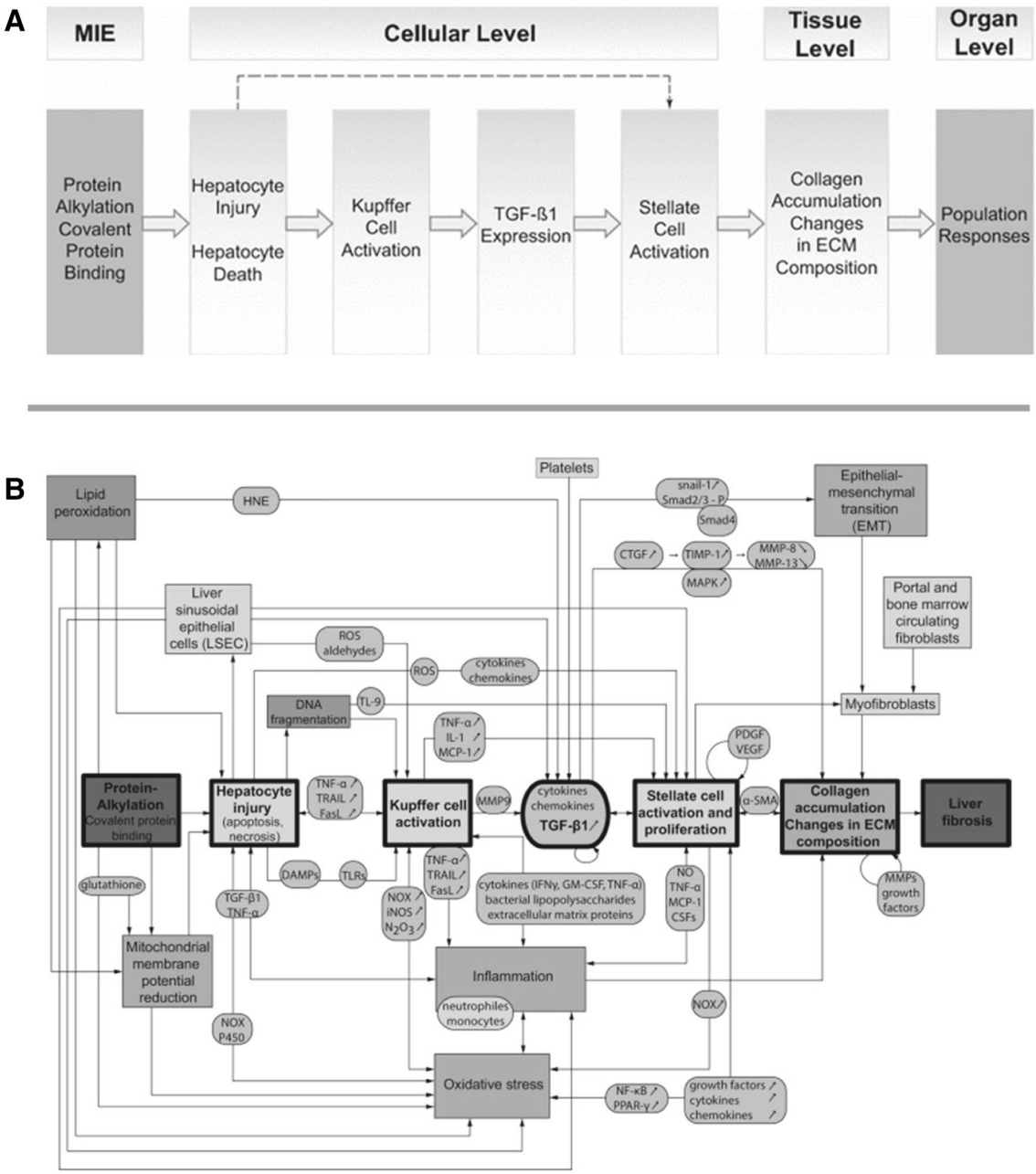

Fig. 3 Comparison of a linearized AOP and a more complex scheme of the underlying biology. **a** Example of a recently proposed AOP aiming to describe critical steps from protein alkylation to liver fibrosis (reproduced from Horvat et al. 2017). The simple linearized form provides a clear and rapid overview of important events. However, modulatory events cannot be seen from such graphical presentations. For instance, there are several mechanisms that remove activated stellate cells and thus affect KE4. Liver fibrosis only occurs when these anti-fibrotic mechanisms are overwhelmed. In line with this, there is evidence from rodent experiments that all steps shown in this AOP, from protein alkylation over hepatocyte death, stellate cell activation and increased collagen deposition can be active without leading to liver fibrosis as the adverse outcome. This may have important implications for some regulatory applications of such an AOP. **b** Suggestion of the same authors on how a more complex network of events may be visualized. Such steps may be taken when an AOP is further developed as quantitative AOP (qAOP)

captured in the "Quantitative understanding of the linkage" section of the KERs. The more completely and quantitatively the KERs and their potential modulating factors are described, the greater is the capacity to predict whether the severity of perturbation (in terms of dose, duration, timing, etc.) at the MIE will be sufficient to elicit the AO. However, it is nevertheless not optimal if pivotal inputs and control mechanisms of high relevance for an AO cannot be displayed

graphically. It is often argued that the description of KE and KER leaves all necessary options to add complex information. However, this argument does not consider the fact that the choice between what is displayed graphically (and what most people will use as the main source of information) and what is presented 'in text form only' is a prioritization of information. Here, the likely user behavior must be considered; and it is probable that with the availability of hundreds of AOPs, the largest majority of users will not read all text (thousands of pages).

The graphical display rules convey the impression that a KE is always more important than a modifier. This assumption can be biologically wrong, and the various modulatory factors may become more important than the KE itself. For instance, TNF may play a role in hepatotoxicity models (KE: TNF-induced apoptosis), but modifiers have such a huge impact that the toxicity is augmented 1000-fold by pretreatment with galactosamine (Leist et al. 1996, 1997a, c, b), or the KE never takes place (independent of the amount of TNF and hepatocyte pre-sensitization) if, e.g., hepatocyte ATP is depleted (Latta et al. 2000). Another example is excitotoxicity triggered by *N*-methyl-D-aspartate (NMDA) receptors, which is modulated by many orders of magnitude by the mitochondrial function and cellular energy state (Volbracht et al. 2001a, b; Leist et al. 1997a, b, 1998a; Leist and Nicotera 1997).

The concept of modulatory factors has been explored in depth in toxicology and pathology concerning the importance of genetic variants relevant to an AOP. Two examples illustrate that these can be more important than the actual threshold event. In carcinogenesis, a chemical can lead to DNA damage (KE) without further pathology (no AO). Notably, DNA damage is ongoing continuously without exposure to chemicals due to endogenously formed oxygen radicals. Yet, the high rate of DNA damage is not sufficient to trigger tumor formation. The activation of 'modulatory' oncogenes or the inactivation of tumor suppressor genes plays a major role in carcinogenesis, and this is not necessarily part of the AOP (as downstream KE); rather, sometimes a chemical-independent modulatory event is determined by, e.g., genetic factors. The same importance of modulatory events is observed in an entirely different pathological setting, involving mitochondrial failure. Toxicants can affect mitochondrial function, but also many endogenous events present a continuous challenge to mitochondria. Modulatory factors that change the threshold of mitochondrial failure, such as the large family of B cell lymphoma 2 (Bcl-2)-related proteins, determine in many cases whether toxicity will be induced or not. In the future, knowledge on mechanisms, together with detailed knowledge on modulatory factors that affect KE thresholds will be instrumental in extrapolating data from in vitro systems to the human situation. This issue is currently tackled, although there is still a lot of work required to understand the complex stress signaling networks and other pathways affecting cell physiology in different culture conditions (Daneshian et al. 2016). Inevitably, KERs and modulatory factors affecting them will have to be experimentally assessed and confirmed by experimentation, using gene knockdown or chemical inhibitors.

Dealing with compensatory mechanisms

Compensatory mechanisms are a specific variant of modulatory factors, but they are so important in biology that they are highlighted here with particular emphasis. AOPs deal with damaging cascades, but are not optimized to deal with repair processes, regeneration and counter-regulation. Notably, the guidance handbook stipulates that all such processes are considered and described (in the ´Quantitative understanding of the linkage' part of the KER judgment) during the establishment of the AOP. Especially in the consideration of KERs, strong emphasis is to be put on the fact that an MIE progresses to the first KE only when all compensatory mechanisms have been overwhelmed and the AO is only reached when this is true for all steps. Once the AOP has been reviewed and established, such considerations are only accessible to those reading the background text in detail. For other users, all considerations of repair and compensation are lost in the graphical display. However, they play a crucial role, especially for chronic toxicities (Hengstler et al. 2012). The prediction of the balance between damage and compensation will be key to the issue of defining adversity from simple test systems, and this research area therefore needs to be strongly fostered and supported. The potential issues will be illustrated using the example of the recently published, and above described, AOP protein alkylation leading to liver fibrosis. This AOP has been published on the AOP-wiki (https://aop.wiki.org/aops/38), and similar figures have been presented (Edwards et al. 2016; Oki and Edwards 2016) (Fig. 2). Briefly, this AOP states that protein alkylation leads to a sequence of hepatocyte death, Kupffer cell and stellate cell activation, increased production of the extracellular matrix and finally to bridging fibrosis.

The relevance of stellate cell activation for liver fibrosis was first reported more than 30 years ago and repeatedly reviewed (Bataller and Brenner 2005; Wynn 2008; Godoy et al. 2013). It is an AOP KE. According to the AOP, a sufficient level of protein alkylation in hepatocytes would suffice to trigger the series of critical events ultimately leading to liver fibrosis. However, this simple sequence of subsequent steps, one triggering the next, does not predict the AO with sufficient accuracy. The flaw can be illustrated by a simple experiment: administration of a hepatotoxic dose of protein alkylating compounds, e.g., acetaminophen (APAP, 300 mg/kg, i.p), (Fig. 4) or carbon tetrachloride (CCl_4, 1 g/kg, i.p.), (Fig. 5a) causes pericentral hepatocyte

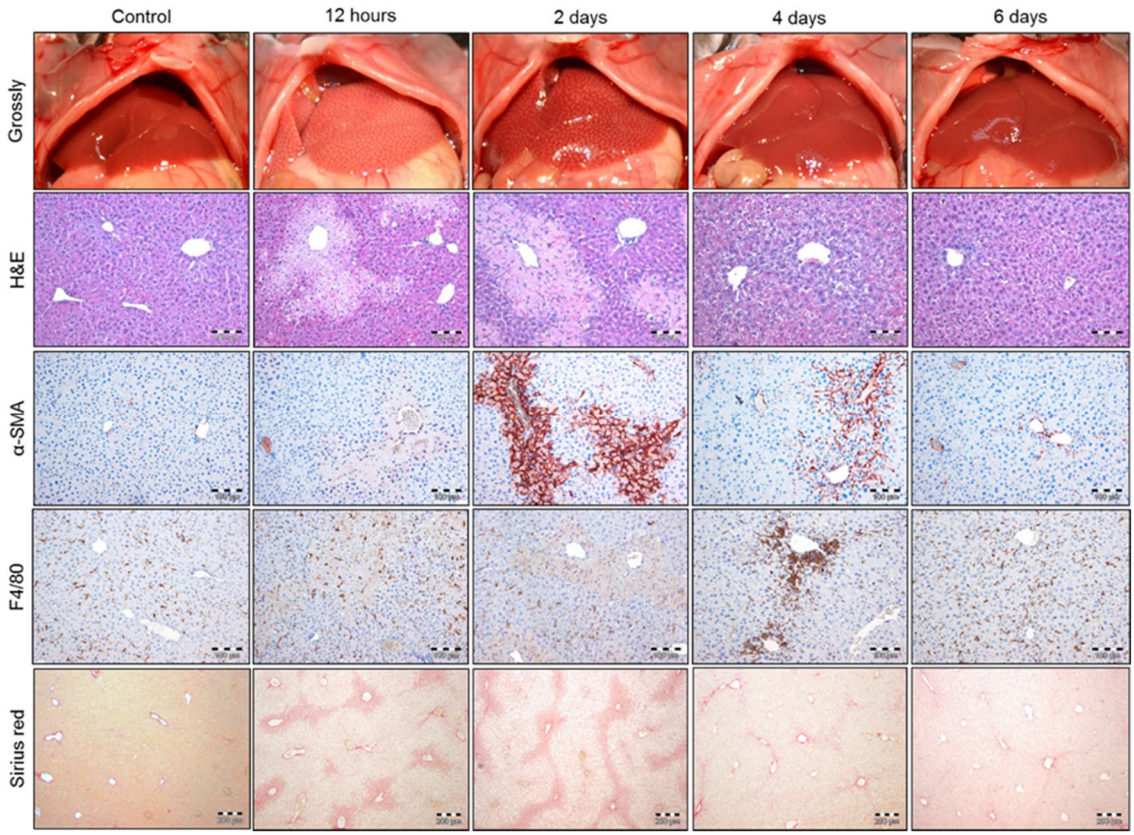

Fig. 4 Occurrence of massive hepatic death and stellate cell activation without ensuing necrosis. The images show livers of C57BL6/N mice at various time periods after a single intraperitoneal administration of 300 mg/kg paracetamol (APAP). The *panels* show the macroscopic appearance of the livers (*Grossly*), hematoxylin and eosin staining (*H&E*), immunostaining with antibodies against α-smooth muscle actin (*α-SMA*) as a marker for activated stellate cells, immunostaining with antibodies against F4/80 as a marker for infiltrating and tissue resident macrophages also named 'Kupffer cells' (*F4/80*) and Sirius red staining to visualize fibrosis (*Sirius red*). APAP causes massive pericentral necrosis visible in the *H&E*-stained slides at 12 h and 2 days. Two days after APAP intoxication, the entire pericentral necrotic area is occupied by activated stellate cells visualized by *α-SMA*. However, 4 days after intoxication, the number of activated stellate cells is reduced and the tissue is completely cleared from activated stellate cells after 6 days. Particularly at day 4, infiltration of F4/80-positive macrophages is seen, which play a role in removing activated stellate cells. Sirius red staining demonstrates that no fibrosis is formed. The key message of this experiment is that massive hepatocyte death and massive stellate cell activation do not necessarily lead to liver fibrosis

death, followed by massive stellate cell activation and increased collagen production. Therefore, all key events shown in the AOP scheme (Fig. 3), hepatocyte death, Kupffer cell activation, and stellate cell activation are triggered, and their activation occurs to a massive extent. Approximately, 40% of all hepatocytes are killed and at least a similar fraction of all stellate cells becomes activated. However, despite the massive activation of all key events proposed in the recently published AOP, no liver fibrosis is induced (Fig. 5). Instead, a perfect regeneration of functional tissue architecture is observed (Hoehme et al. 2007, 2010). Similarly, no liver fibrosis is seen following repeated injections of the same doses of APAP or CCl_4, if the intervals between the individual doses are long enough, e.g., 4 weeks, even if massive hepatocyte death and stellate cell activation are induced after each individual compound administration. What is the reason for the apparent discrepancy of such experimental findings and the published AOP? There are two important answers: (1) from the mechanistic perspective, the reason is that the AOP does not account for regeneration and compensation. In fact, the activated stellate cells are rapidly removed during liver regeneration (Fig. 6). While a maximal number of activated stellate cells is seen around day 3 after the induction of liver damage, it takes only until day 6 for all of them to be removed. It is important to consider that removal of activated stellate cells represents an active process. The responsible mechanisms have been reported previously and include clearance by immune cells, e.g., macrophages (Duffield et al. 2005; Ramachandran et al. 2012), apoptosis which can be induced by the activity of natural killer cells infiltrating the regions where stellate

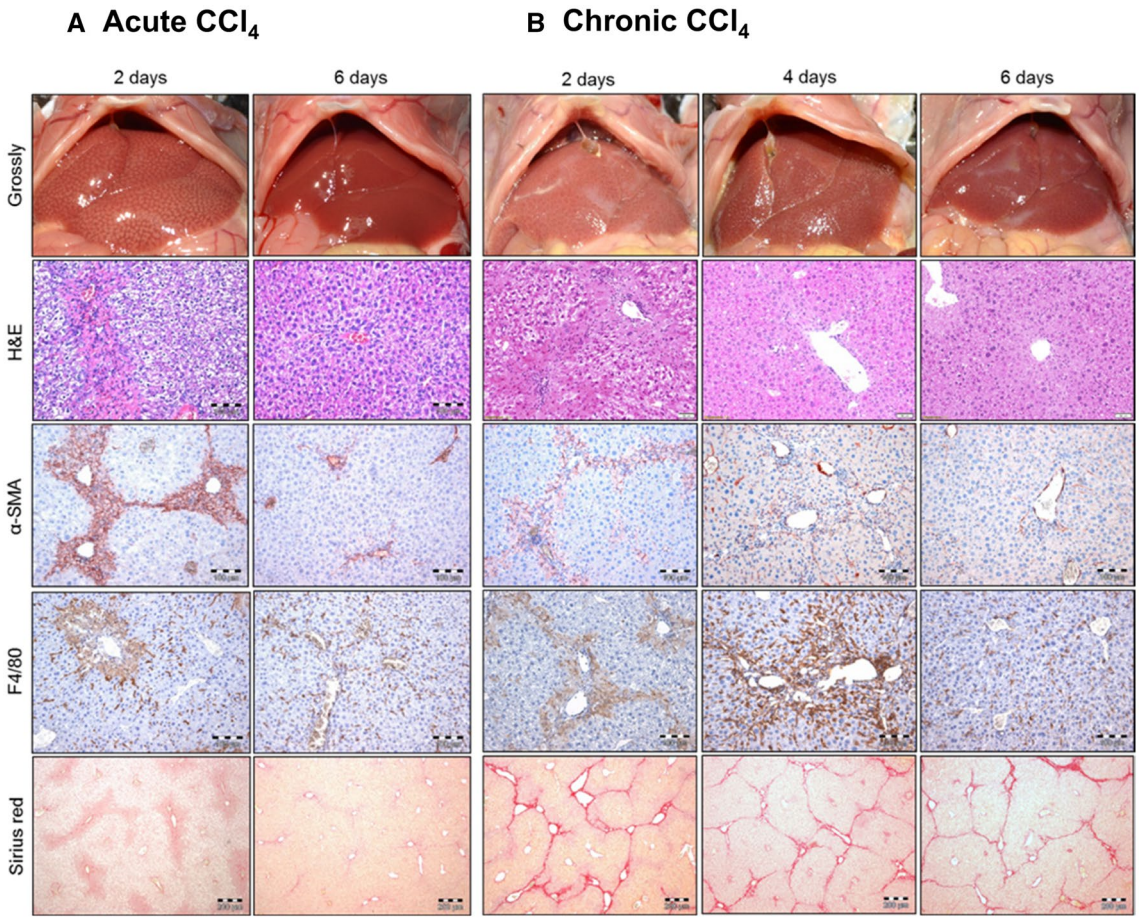

Fig. 5 Lack of correlation of stellate cell activation and the degree of fibrosis. **a** The images show livers of C57BL6/N mice at various time periods after a single intraperitoneal administration of 1 g/kg CCl_4. Despite massive pericentral cell death (*H&E*) and stellate cell activation (*α-SMA*) at day 2, all activated stellate cells are removed until day 6 and no fibrosis is formed (Sirius red). Therefore, the situation is similar to that after acute APAP intoxication shown in Fig. 4a. **b** Administration of 1 g/kg CCl_4 twice weekly for 2 months. The images were taken 2, 4 and 6 days after the last dose of CCl_4. Interestingly, stellate cell activation is much weaker compared to the acute situation. Nevertheless, liver fibrosis occurs (Sirius red), which is not reversible within the studied time period and persists although the activated stellate cells are removed (*α-SMA*). The experiments show that formation of liver fibrosis is not sufficiently explained by the extent of activated stellate cells. A correct model must take into account whether anti-fibrotic mechanisms, such as removal of activated stellate cells and lysis of deposited extracellular matrix, are overwhelmed or still able to compensate pro-fibrotic mechanisms

cells have been activated (Radaeva et al. 2006; Tian et al. 2013; Fasbender et al. 2016), senescence of activated stellate cells which results in a phenotype with reduced fibrogenic activity and increased sensitivity to killing by immune cells (Krizhanovsky et al. 2008; Kim et al. 2013), as well as reversal of activated, collagen-producing stellate cells to a quiescent state (Kisseleva et al. 2012). Moreover, numerous studies have been published that describe the removal of the deposited extracellular matrix by proteases (Duffield et al. 2005; Ramachandran et al. 2012; Ramachandra and Iredale 2012; Tacke and Zimmermann 2014); (2) from a risk assessment point of view, the AOP should assume a certain time frame of repeated exposure (frequent, continued exposure, short intervals). Only in this case, e.g., twice weekly administration of hepatotoxic doses (Fig. 5b), fibrosis is formed.

Thus, the AOP may sometimes be right and sometimes it delivers inadequate predictions. A detailed examination of strengths and weaknesses, as exemplified here, may help to make the AOP more comprehensive. The studies displayed here suggest that a more universally applicable AOP of liver fibrosis must include pro- and anti-fibrotic mechanisms as well as exposure and timing conditions. An adverse outcome, such as fibrosis, will occur only when pro-fibrotic processes exceed and exhaust anti-fibrotic mechanisms for a critical period of time.

In conclusion, those examples illustrate that an all too simple (linear, non-modulated) AOP concept may not agree

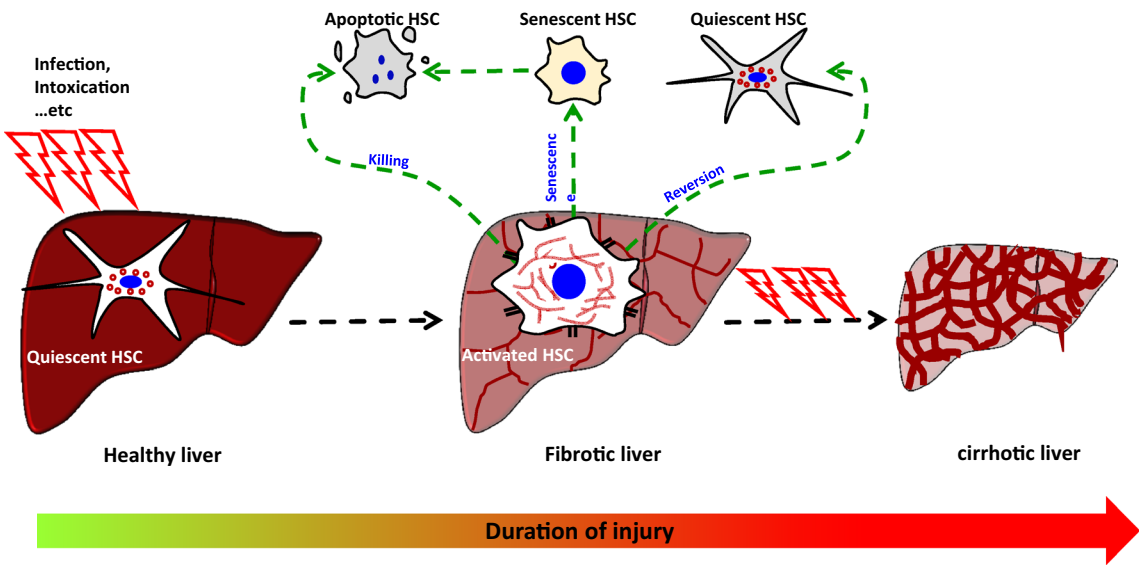

Fig. 6 Display of anti-fibrotic mechanisms. Removal of activated stellate cells by immune cell-mediated killing, senescence and reversion to a quiescent state represent some of the most important anti-fibrotic mechanisms

with experimental data. It is important to take into account that hepatocyte death and stellate cell activation represent important key events, although per se not inevitably causing liver fibrosis. It is out of question that humans should be protected from any exposure that causes hepatocyte death and stellate cell activation. Therefore, establishing in vitro systems that reliably predict these key events is of high relevance. On the other hand, an AOP with "stellate cell activation" as a key event should not be interpreted in a way that lack of stellate cell activation guarantees non-toxicity, in this case no fibrosis. The descriptive AOP concept (as detailed above) may not be up to this challenge. For this reason, the concept of quantitative AOP (qAOP; a concept which is much more akin to systems biology) is being developed in parallel, and some of the descriptive AOPs of nowadays may be converted into qAOPs later on (Table 1).

Binning of events

Due to the limited information on what set of molecular events really contributes essentially to a given KE (like mitochondrial dysfunction), many KEs in AOPs are defined very broadly (synaptic dysfunction, hepatocyte injury, stellate cell activation, etc.). Thus, many different biochemical/molecular events are put into 'large bins'. Such bins can be as large as the whole field of neuroinflammation or mitochondrial failure, or altered intracellular proteolysis, each comprising hundreds of players and being covered by thousands of articles. A KE is meant to define a given cellular state, but if the exact molecular change is not well defined, it actually comprises several cellular states and this may be problematic for the subsequent KER. The binning of several cell states into one KE is a helpful simplification for many purposes, but it makes quantification and computation in the sense of systems biology very difficult. Therefore, it will be important to explore in much more detail how far AOPs with large sets of KEs can be used for qualitative or semi-quantitative hazard estimates. Technical artifacts are another consequence of the 'binning' approach in AOP networks. Apparently identical KEs (e.g., inflammation or mitochondrial dysfunction in various AOPs) may in fact be based on different biochemical mechanisms that have been put into the same bin, and this would lead to very different KERs. Binning is not only done for KEs, but also for MIEs. For instance, protein alkylation is considered as an MIE for liver fibrosis. However, there are several fundamentally different ways to alkylate protein's nucleophilic structures. Some agents preferentially target cysteine residues, while others prefer other targets, and the type of cellular damage and its downstream consequences are therefore different. Moreover, of course, other mechanisms besides protein alkylation may cause hepatocyte death followed by stellate cell activation and liver fibrosis. Therefore, the absence of experimental evidence for protein alkylation should not be misinterpreted in a way that the substance of interest will not cause fibrosis in vivo.

Dealing with multiple hit events

The linearity of AOPs suggests that only one MIE initiates the chain of events that finally results in an AO. It is unclear how this concept can account for two or multi-hit

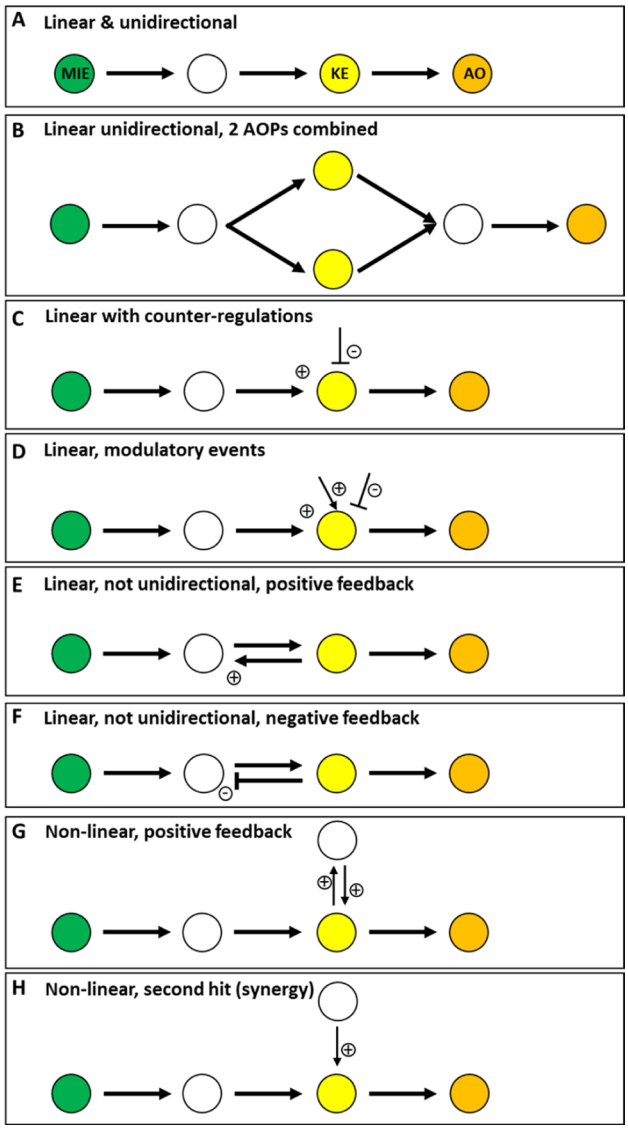

Fig. 7 Illustration of linearity, unidirectionality, feedforward, feedback and modulatory events. **a** A linear AOP consists at least of one MIE (*green*), one KE (here two are shown; in *white* and *yellow*) and the adverse outcome (*orange*). **b** Linear AOPs can be combined to AOP networks with shared MIE and AO, as shown here, or with disparate MIE/AO and only one shared KE. **c**, **d** Some KEs may be strongly affected by modulatory factors (positive or negative). These are not usually displayed graphically. **e** An AOP-like sequence may appear linear on graphical display, but a positive feedback loop may indeed break unidirectionality. For modeling applications, this corresponds to a bifurcation (at a given KE, the continuation goes two directions (backwards and forwards). Note that in the graphical display, the AOP still looks like a linear (one-dimensional) array of events. Such features may, e.g., be observed in cell death signaling. **f** Another AOP-like sequence may contain a negative feedback loop. Such features are found in many metabolic pathways. **g** A positive feedback loop may occur in a non-linear sequence, as exemplified in Fig. 2. **h** An AOP may take a non-linear appearance, if it joins an endogenous, toxicant-independent event that is continuously ongoing and important for the AO to happen. For instance, tumor promotion may occur by phorbol esters, acting on protein kinase C (MIE), but they require continuously ongoing DNA damage. Several events in toxicology require more than one hit (e.g., classical initiation-promotion-progression model of carcinogenesis) and are difficult to display as linear event with a single MIE at least if both events are triggered by the same chemical, or if one of the event is chemically independent (color figure online)

mechanisms, or the initiation–promotion model in carcinogenesis. A further example, acute-on-chronic organ failure (most frequently of the liver or kidney), represents an increasingly recognized entity encompassing an acute deterioration of organ function in response to an acute event, which occurs only under conditions of a chronically stressed tissue (Jalan et al. 2016). All situations in which a KE is not a downstream consequence of an MIE, but an independent event, essential for pathology to happen, cannot easily be captured by an AOP, and the concept needs to be opened for such important pathological sequences (see Fig. 7 for overview). A typical scenario where multi-hit events occur is when one compound generates more than one toxicologically relevant metabolite. In such situations, each metabolite may induce different KEs. This may become critical, when those KEs synergize and contribute to the same AO. Especially, when the simultaneous occurrence of two KEs is a precondition to induce a specific AO, it will be challenging to represent this by a linear AOP. Likewise, AOP networks may be difficult to construct, if there is not an overlap of KEs, but rather a KE of one AOP acts as strong modifier of a KE in another AOP. An illustrative example is that of the nasal tumors induced by vinyl acetate in rats (reviewed in Hengstler et al., 2003). After inhalation, vinyl acetate is metabolized to acetic acid and acetaldehyde at sites of contact, such as the olfactory epithelium. Acetic acid causes a pH decrease that, when exceeding critical thresholds, leads to cell death and replacement proliferation. The second metabolite, acetaldehyde, induces genotoxic effects, such as DNA–protein cross-links, chromosomal aberrations and sister chromatid exchanges, but not point mutations. Extensive experimental work has shown that the genotoxic effect synergizes with pH-induced cell death, and this mechanistic knowledge has supported regulatory toxicology and identification of adequate threshold values. Long-term exposure to 600 ppm vinyl acetate causes a reduction in pH by 0.49 units, which leads to olfactory (basal cell) degeneration and an increase in basal cell proliferation that is more than twofold above control levels (Hengstler et al. 2012). Physiologically based pharmacokinetic (PBPK) models indicate that inhalation of 600 ppm vinyl acetate leads to about 12 µg/ml of acetaldehyde in olfactory epithelial cells. This is clearly above the concentrations that cause sister chromatid exchanges and chromosomal aberrations in in vitro tests. This matches the results of a carcinogenicity study where a significant increase in nasal tumors was observed in rats at 600 ppm vinyl acetate. At 200 ppm of vinyl acetate, the pH

in the olfactory epithelium is reduced by approximately 0.25 units, which leads to only mild olfactory degeneration and very weak replacement proliferation. Acetaldehyde concentrations in olfactory epithelial cells are 5.4 µg/ml according to the PBPK model and thereby still above concentrations that cause genotoxicity in vitro. In the carcinogenicity study, only one nasal tumor was observed, which was not statistically significant. 50 ppm vinyl acetate can be considered a NOAEL for both mechanisms—replacement proliferation as well as genotoxicity. This concentration causes a pH reduction of 0.08 units, which results in neither cell death nor replacement proliferation. The basal cell acetaldehyde concentration is modeled to be 1.7 µg/ml, which is below concentrations reported to cause genotoxicity in vitro. No nasal tumor is induced at this exposure level.

Acetaldehyde is a natural constituent of the body and is formed in endogenous metabolic processes, a fact that is relevant for risk evaluation. Normal, blood concentrations are in the range of 0.3 µg/ml (Hengstler et al. 2003). Therefore, a very conservative argumentation would be that vinyl acetate exposure should be below levels at which acetaldehyde metabolically generated from vinyl acetate exceeds the endogenous concentrations of 0.3 µg/ml. This would be in the range of 1 ppm vinyl acetate. The mechanistic knowledge from basic science briefly summarized in this section provides regulators with a basis to make informed decisions on adequate threshold values.

Trying to force this and similarly complex scenarios, where different MIE synergize to induce an AO into the schematic of a linear AOP (Fig. 1), does not seem to be adequate. Presenting two MIE triggered by two metabolites that induce two different KEs may be more appropriate. Importantly, both KEs synergize with respect to the AO. Besides vinyl acetate, situations where different mechanisms have to be active simultaneously to induce an AO occur also outside the field of carcinogenicity (Jalan et al. 2016).

How to deal with effect duration

The text description of AOP elements allows for considerations of effect duration or for delayed effects. However, the construction elements available do not visualize this, and the ideas on how this needs to be considered for semi-quantitative or quantitative AOPs are still at a relatively basic level. There are many biological examples of the extreme importance of stress duration, independently of the overall stress intensity threshold. This principle was already formulated at the beginning of the twentieth century in the form of Haber's rule (toxic effect = concentration of toxicant × exposure duration). An illustrative example has also been presented here (above, Fig. 4) with the experiments on liver fibrosis triggered by acetaminophen (APAP) or CCl_4. Single doses, or repeated doses with sufficient recovery time, do not lead to AO. However, liver fibrosis occurs if a certain frequency of hepatotoxic doses is exceeded; for example, 2 doses of 1 g/kg CCl_4 per week will exceed this threshold (Fig. 5b). For applications of AOPs in risk assessment, formalization of the timing aspect appears important.

A specific feature of the timing aspect is not only the length of exposure (relevant in repeated dose toxicity), but especially in the area of developmental and reproductive toxicity (DART) the window of exposure, i.e., the exact phase of development in which exposure occurs. This is exemplified by the recently published AOP leading from topoisomerase inhibition and subsequent mixed lineage leukemia (MLL) gene rearrangement to infant leukemia (Pelkonen et al. 2017). In the suggested AOP, the MIE is called "in utero topoisomerase poisoning" to note the developmental period. Incorporation of such timing information into graphical displays should be further encouraged for all DART-relevant AOPs. An exact documentation of the timing aspect is also provided by work on stem cell-based developmental toxicity models, where small changes of the exposure window cause large differences ranging from pronounced adverse effects to no effect (Balmer et al. 2012, 2014; Balmer and Leist 2014; Shinde et al. 2016).

Open questions to be addressed and communicated concerning construction, assessment and handling of AOPs

With a complex system such an AOP, which is still at the beginning of its development, two types of question often arise. First, there are still open issues that need to be discussed, consolidated, or developed further. Second, some concept features are easily missed or misunderstood, or are applied and promoted by stakeholders in ways originally not foreseen. Both issues are important for harmonized use and understanding. These issues are not necessarily strengths or weaknesses, but discussing them is helpful to appropriately use and implement the AOP concept in an in vitro-based hazard evaluation. Some of these features are briefly addressed in this chapter.

The issue of non-toxicity

AOPs are a tool for hazard assessment. It is widely accepted that a compound likely to activate an AOP is of higher concern than compounds activating no AOP. However, going one step further, the issue is less clear. What does this mean in quantitative terms? What does concern really mean? Also very important: is absence of evidence really evidence of absence concerning toxicity? The answers to these questions are far beyond the scope of this review, but the discussion

appears important and two of the arguments (coming from very different angles) are briefly highlighted here.

In some discussions, the real or purported flaws of AOPs are used to increase doubt about the potential toxicity of compounds. Some discrepancies, matters of incompleteness, or species issues in AOPs are highlighted. Moreover, the apparent lack of absorption, distribution, metabolism, and excretion (ADME) information is criticized and any such 'AOP imperfections' are used to suggest that association of a given compound with an AOP is not relevant to its hazard. Such arguments are found for instance in public consultations on new AOPs, mostly promoted by representatives of the chemical industry.

An entirely different type of discussion arises from the question of which conclusions can be drawn from the activation or non-activation of AOPs. In this context, it is particularly important to avoid the misconception that failure to activate an AOP, a group of AOPs, or even all documented AOPs means non-toxicity of a compound. The AOP concept states that an individual AOP makes no statements on 'non-toxicity', and this question is thus neither a strength nor weakness of AOPs. This must also be understood in the light of defining criteria of AOPs: They are compound agnostic and do not involve the toxicokinetic phase. For statements on non-toxicity, a defined compound always needs to be considered, and ADME plays a key role. Thus, at present, the arguments on this issue are pretty clear. The situation may however change in the future, when many more AOPs will be established, and the toxicological community will feel that a large percentage of all possible AOPs in a given area, or in the entire field of toxicology, have been described. Tests for all MIEs/KEs may become available, and these may allow a comprehensive experimental access to the question whether a compound activates any AOPs at a relevant concentration. In such a situation, failure of a chemical to activate any of the many tests may be interpreted as a low likelihood to be hazardous to humans. However, at present, the state of the art is far away from such a situation.

Notably, this issue is not to be confused with the use of AOPs as a basis to develop an integrated testing strategy (ITS), or an integrated approach to testing and assessment (IATA). In this case, testing of AOP KEs is combined with other information, and the KE tests are chosen in such a way as to broadly cover many modes of action leading to a widely accepted and relatively general adverse outcome. The most prominent example is testing for skin sensitization (NTP 2017), and similar approaches are in preparation for testing of estrogenicity. In such cases, negative results would have toxicological and regulatory importance, although only within the small range of questions covered by the ITS/IATA.

An interesting question is how to handle AOPs that are driven by threshold or non-threshold mechanisms. For threshold mechanisms, the situation is quite clear. If an MIE occurs at levels that are too low to induce a KE, no AO can be expected. The scenario becomes more challenging when no threshold mechanism has been demonstrated, which is for example the case for some types of DNA adducts (MIE), mutations (KE) and tumors (AO). In this case, one has to assume that in the presence of MIEs all other steps (KE) are also active leading to the respective AO, or at least an increased risk of AO. This underlines the importance of differentiating whether an AO is caused by threshold or non-threshold mechanisms (Bolt et al. 2004; Oesch et al. 2000, 2001; Hengstler et al. 2003).

Quantification of KEs/KERs; threshold setting

The example of the first practically applied AOPs, i.e., the AOP developed for skin sensitization, shows how difficult it is to move from a qualitative AOP (as it has been deposited) to something with clear quantitative thresholds that can be used in a prediction model for risk assessment and quantification. Until now, there is no consensus prediction model available, although the KE tests have been extensively validated and a vast amount of quantitative work has been invested in this area. One of the difficulties in such a quantitative use of AOPs is the setting of appropriate thresholds, or the establishment of respective mathematical rules that connect the sufficient perturbation of $KE_{upstream}$ with the triggering of $KE_{downstream}$. As sequences are not all linear, and as they are subject to modulatory factors, complex non-linear relationships can arise, and the parametrization of such models, as well as their validation, is both difficult and requires large data sets.

An illustrative example has been given above for vinyl acetate: the threshold setting for a KE assay of one metabolite would depend on the pH, and the latter is determined by another metabolite, in addition to multiple other factors. In this situation, an AOP-guided risk assessment would most likely be useful in cases of very high compound exposure (far beyond a likely threshold), but uncertainties would rise if lower concentrations, relevant to human exposure are considered, and the tool may not be suitable for regulatory decisions in complex situations of synergistically interacting pathways.

Another example is found above, with fibrosis triggered by alkylating agents. The extent of hepatocyte death is a KE threshold of relevance. This is illustrated by the fact that healthy livers continuously experience low levels of hepatocyte death events, which leads to permanent low-level replacement proliferation. If for example, a healthy adult C57BL6/N mouse receives a single injection of bromodeoxyuridine (BrdU) 1 h prior to preparation, approximately 0.5% of all hepatocytes stain BrdU positive. This illustrates that hepatocyte death and replacement proliferation occur

continuously to a low extent in healthy livers. These hepatocyte death events are quantitatively below critical thresholds, since mice (and humans) do not develop liver fibrosis throughout an entire lifetime unless affected by liver diseases. Thus, this KE is always ongoing without triggering the AO, unless additional modulatory factors change, and a lot of work will be required for the exact threshold setting.

Can toxicodynamics be separated from toxicokinetics and metabolism?

The AOP concept states that AOPs only predict hazard, but not risk. Moreover, the aspects of ADME are explicitly not considered by AOPs. This may seem logic within the AOP system, as they are 'compound agnostic': Only compounds can have toxicokinetic behaviors, but not a pathway. However, in the weight-of-evidence process to support AOP KERs, example chemicals are essential, and for such examples it is hard to neglect their individual pharmacokinetic properties. Considering this, it is not clear whether the toxicokinetic and toxicodynamic phase can always be clearly separated. For instance, one of the best-known neurotoxicants is MPTP. Its selective toxicity to neuronal subpopulations, but not to astrocytes, can only be explained by metabolic and distribution steps (Schildknecht et al. 2015). Vice versa, the MPTP metabolite MPP + is toxic in vitro to neurons (Schildknecht et al. 2013), but hepatotoxic (and not neurotoxic) in vivo.

Another complicated situation arises, if one compound generates several metabolites, which trigger various interacting AOPs. Moreover, some AOP KEs may affect the metabolism/distribution of a compound that triggers the respective AOP; or in other words, activating or detoxifying metabolism does not necessarily have to occur exclusively prior to the first KE (although this is often the case), but can be between two KEs and represent a necessary precondition for a downstream KE. Examples are tumor precursor cells (e.g., preneoplastic nodules of the liver) that express higher levels of detoxifying enzymes and give preneoplastic cells a selection advantage which facilitates the next steps on the path to cancer. A clinically relevant example is 'acute-on-chronic liver failure'. Here, chronic administration of a compound (or a chronic disease state) may alter the levels of metabolizing enzymes that render the organ susceptible to 'second hits'. A special example is toxic cholestasis. Here, the initial KE is damage to cholangiocytes, leading to obstructions of the biliary tract. This leads to increased bile salt concentrations in the blood. As a mechanism of this KE, it has been suggested that obstructions in parts of the biliary tract induce an increase of the inner surface area of interlobular bile ducts by surface corrugation as well as looping and branching, which increases their reabsorption capacity for bile salts (Vartak et al. 2016; Jansen et al. 2017). As an adaptive response to increased bile salt concentrations in blood, hepatocytes downregulate enzymes involved in bile salt synthesis and transporters that convey bile salts into hepatocytes (Jansen et al. 2017). This is accompanied by down-regulation also of xenobiotic metabolizing enzymes, which may render the liver susceptible to 'second hits'. Moreover, also with molecules undergoing metabolism and complex transport processes, a strict separation of the ADME phase and AOP may be somewhat artificial; strictly (and theoretically) speaking, the AOP begins with an MIE triggered by the concentration of the respective compound or its metabolite at the target site of the MIE, but in many cases such concentrations are hard to measure. Moreover, great care needs to be taken to avoid establishment and exemplification of AOPs on the basis of chemicals used at unrealistically high concentrations (e.g., in in vitro systems). At present, the experience gained with AOPs is too limited to decide whether the strict separation between AOP and toxicant ADME can and should be maintained. The question needs to be borne in mind and the situation monitored accordingly. It should also be considered that separation of metabolism and toxicokinetics from toxicodynamics has been introduced because of practical circumstances. Since specialized tools are required, ADME is usually analyzed in different departments than toxicodynamics. However, from a biochemical or cell biology point of view, this separation is rather arbitrary. A glutathione molecule binding to a substrate or an epoxide binding to the N7 position of guanine both represent nucleophilic additions and it may be considered as arbitrary to accept only one of them as KE. Therefore, one potential alternative may be to include metabolic activation or inactivation of parent compounds in the form of a special KE. Such processes of drug metabolism are measurable not only in vivo, but relatively reliably also in vitro, determine the outcome of toxicity in a decisive way and describe important system states. The technically relatively easy introduction of metabolism as a special KE into AOPs would open attractive additional possibilities. On the one side, the compound agnostic nature of the AOPs would not be compromised if metabolic activation/detoxification is introduced as general KE. Additionally, important compound-related aspects and interspecies differences could be correctly described; importantly, a KE 'metabolism/detoxification' could be analyzed in vitro. Such currently untapped potential of the AOP concept is exemplified by the case of the heterocyclic amine MeIQx (2-amino-3,8-dimethylimidazo[4,5-*f*]quinoxaline), which is present in cooked and fried meat. MeIQx is carcinogenic in rats, but does not induce hepatocellular cancer in cynomolgus monkeys (reviewed in: Hengstler et al. 1999). Without knowledge of the responsible molecular mechanisms, this discrepancy between the two animal species made human risk evaluation difficult. The relevant metabolism to generate the genotoxic electrophile has been explored in depth:

MeIQx is metabolically activated to a nitrenium ion. The latter then binds covalently to DNA. This MIE initiates a complex series of KE, such as mutations, activation of oncogenes, inactivation of tumor suppressor genes and compromised cell cycle control, which finally leads to cancer (AO) (reviewed in Hengstler et al. 1999). Knowledge of the metabolism and of the hazard cascade (AOP) could now be combined in two different ways.

The solution suggested by most relevant stakeholders of the AOP concept would be to combine the different types of information in an IATA to be used for risk assessment or for species comparison and extrapolation. In this context, the information could be used that human hepatocytes form the phase I metabolite hydroxylamine (a precursor of the nitrenium ion) in a similar way to hepatocytes isolated from rat strains that are susceptible to MelQx carcinogenicity; in cynomolgus monkeys, this metabolic pathway shows only little activity. Therefore, humans have to be expected to be similarly susceptible to MeIQx-induced carcinogenicity as some rat strains, but differ from cynomolgus monkeys that are resistant. In the past, such an observation would have puzzled scientists, since the phylogenetic proximity of humans to monkeys may intuitively lead to the opposite expectation.

Another solution would be to add a KE of 'electrophile formation' before the MIE or it could be integrated as part of the MIE. This would at present be against the rules and concept of AOP. It needs to be observed and compared for many cases to come in the near future whether there would be an advantage in a more overlapping approach of ADME and hazard assessment, especially considering the issue that AOPs do not refer to specific compounds while ADME somehow has to be compound specific.

Review process and updating

There are two quality control issues to be considered: (1) the establishment of the AOPs on the OECD website; (2) the publications on specific AOPs in scientific journals. The journal peer review process has mainly been designed for primary original research work. It puts a high time constraint on prominent researchers and only works since this is considered a mutual service, linked to the benefit of knowing at an early time point what the competitors are doing. These incentives are not given for the reviewing of a large and complex document as an AOP. Therefore, it is unlikely that the major experts in the field will be willing to invest sufficient time in it. This will reduce the quality of the review process. Moreover, the one-way, single-blind review system is not suitable for reviewing extremely complex weight-of-evidence information as happens in AOP. Discussion and feedback are required, just as the involvement of broad sets of experts. This makes the publication of AOPs in scientific journals challenging and problematic. Another problem with journal publications is that the articles are immovable, while knowledge continuously advances. Both, for the review and for updating of content, a blog-like system would be more suited. Concerning journal publications, new solutions will need to be found, as it is important to inform the broad scientific community also in journals on key advances in the field of AOPs. Some review journals have already implemented a discussion process of reviewers and authors, and few of them have also documented this discussion.

Similar considerations apply to the AOP website (AOP-wiki). An AOP developed within a project of the AOP work program at the OECD undergoes a rigorous and transparent internal and external review process. If successfully completed, this ends with endorsement and publication in the OECD series on AOPs (Delrue et al. 2016). The establishment and accreditation of an AOP is thus already now a multi-stage process with several discussion and feedback rounds. However, it is difficult to track version history and discussion arguments later on. There is also no clear plan for scheduled updating of AOPs and for conflict management. At present, the highly controlled process is very structured, but this may prevent important and valuable input from many scientists who shy away from the administrative effort. It remains also unclear how the hundreds of AOPs would be kept updated. A radical, but clean solution would be to tag AOP with an expiry date and if they are not reviewed, e.g., within a 5-year period, they would be automatically labeled as obsolete.

Quality assurance and validation of AOPs

Closely linked to the above question is the issue of quality assurance and validation of an AOP in its function as a toxicological tool (e.g., as a basis for a testing strategy). At present, AOPs are based on scientific literature. However, the current "reproducibility crisis" describes the increasing realization that scientific literature is too often not reproducible (Hartung 2013), may contain contradictive data, or may not be sufficiently transparent and clear to the reader (Leist et al. 2010, 2012). Moreover, most of the relevant data are derived from animals, and the transferability to human biology is often not clear (Leist and Hartung 2013). This stresses the importance of using elements of evidence-based toxicology (Hartung 2009) such as quality scoring, systematic review, risk-of-bias analysis, and meta-analysis for the quality control and validation of AOPs. There is not yet agreement on how to apply and formalize this; some discussion can be found in the context of mechanistic validation (Hartung et al. 2013a).

One procedure for simple steps and complex events

Some AOPs are very straightforward and the information contained is relatively 'trivial'. For instance, an aromatase inhibitor leads to aromatase inhibition and thus to a lack of hormones generated by aromatase (http://aopwiki.org/aops/25). Similarly, an estrogen receptor agonist increases systemic estrogen signaling. These examples may help to break the ice and promote the field. However, it is not yet clear whether such examples will really help us in defining and solving the problems of building a complex test strategy. Note that 'trivial' is not meant as a derogatory term here, but refers only to the complexity of the construct.

It is important to recall here that one major goal of the AOP framework is to provide an overarching concept to improve our in vitro testing strategies. It is well established that aromatase inhibition will compromise estrogen synthesis and may therefore lead to reproductive dysfunction. Therefore, it is clear that aromatase inhibitors may cause reproductive dysfunction, when present at sufficiently high concentrations at their target sites. However, besides aromatase inhibition, many more mechanisms may cause reproductive dysfunction. A challenge for the future will be to establish batteries of in vitro assays that cover as many as possible relevant mechanisms that besides aromatase inhibition may also cause reproductive dysfunction. The more complex the mechanisms that lead to an adverse effect, e.g., the development of steatosis or neurodevelopmental disturbances, the larger will this challenge become. In these cases, more complex forms of presenting AOPs and AOP networks may be required. Alternatively, it could be that dealing with more complex pathologies is something entirely different and cannot be attained by simple upscaling of linear small steps. It may require completely different approaches and concepts. We do not know this at present, and it is important to better define areas where AOPs have particular strengths and areas, where it is not clear how well the concept works. In this context, it is important to remember that the concept always has to be seen in combination with a defined application, if its suitability is judged. AOPs for complex events have their justification also now already for instance to sort the available information (Hartung et al. 2012, 2017).

First evidence has been presented that high-throughput transcriptomics may be helpful to populate sAOPs with data that allow decisions and predictions (Jennings 2013). For instance, it has been found that specific AOs are associated with relatively similar gene expression responses, even if the MIE are diverse. An example is hepatotoxicity, where incubation of primary human hepatocytes with relatively high numbers of compounds ($n = 148$) has shown that a stereotypical expression response (of the same set of up- or down-regulated genes) is induced by numerous compounds known to be hepatotoxic in humans, even if these compounds act by different molecular mechanisms (Grinberg et al. 2014). For development of test methods, this observation may be of critical relevance; instead of analyzing a multitude of MIEs, it may be sufficient to measure a set of 'biomarker genes' if their deregulation is strictly associated with a set of distinct KEs that all lead to the same AO. However, implementation of this concept still requires a better understanding of the causal relationships between molecular initiating events, deregulation of 'biomarker genes' for KEs and the occurrence of AO.

How can the hazard-based system of AOP be integrated with toxicokinetics in overall risk assessment?

In the late 1920s, JW Trevan's work aided the introduction of the lethal dose 50 (LD_{50}) test (Trevan et al. 1927). Although considered crude today, it was a major step at the time, as it provided the first quantitative measure to compare the toxicity of different substances. Quantitative dose–response relationships remain a core principle of toxicology, where the dose (estimated or quantified) is related to a particular quantifiable adverse response, the most severe being death, as in the LD_{50} assay. When cell cultures began to be applied to chemical testing in the 1960s, the LD_{50} test was recapitulated using various types of cytotoxicity assays, from trypan blue exclusion to the MTT assay. Later, this became a bit more sophisticated in that we could discriminate between apoptosis and necrosis. However, none of it really helped with the prediction of human risk in real-world exposure scenarios as: (1) very little mechanistic insight was forthcoming, (2) there was often a serious over- or underestimation of the actual exposure concentration, (3) often the actual tissue concentrations remained unknown and (4) many of the in vitro systems did not represent the normal physiology of the parent tissue, which can affect the uptake, extrusion, metabolism, or molecular targets. Today, the LD_{50} in in vivo experiments plays only a minor role and more quantile and information-rich measurements are used, such as tissue structure, behavior and biomarkers. The in vitro counterpart is almost exclusively involved in discovery and verification of the mode of (adverse) action of compounds.

Regardless of the chemical (or nanoparticle) that can cause an adverse effect, it first has to reach the target tissue and reach concentrations high enough to cause a perturbation, which could also be an activation or inhibition of a process, and the perturbation must be of a certain strength and duration to cause an imbalance in cellular homeostasis. Only if the initial perturbation (or MIE) is substantial enough, it will trigger adaptive and/or maladaptive responses, which will alter cellular biology and in time lead to an adverse outcome.

Thus, there are two major challenges for the use of AOPs in risk assessment: (1) quantification of the kinetics (what

the biological system does to the chemical over time) to determine the internal concentration at the MIE site; and (2) quantification of KE thresholds that lead to downstream KEs (what the chemical does to the system over time) (Wilmes et al. 2013, 2015). In this context, it is important to note that KEs may be measured indirectly through biomarkers. Many of these biomarkers may indicate a KE, but are not implicated in the AOP progression themselves and thus are by definition no KEs.

Concerning the issue of toxicokinetics, one challenge that remains is to be able to model the oral doses that lead to these critical concentrations, because the possibilities of pharmacokinetic modeling are still limited due to uncertainties. In this context, modeling and simulation approaches offer a way to bridge the relative chemical exposure seen in the in vitro system and in the target organ in vivo (Fisher et al. 2017). In the in vitro system, a number of factors need to be considered to relate the nominal (applied) concentration to the concentration that is actually driving the toxic response. These include binding of the chemical to lipids and proteins in the cell culture medium, partition between the medium and the headspace of volatile chemicals, partitioning into and binding to cellular constituents, as well as binding to plastic in the cell culture system. All of these processes will act to decrease the free (unbound) concentration in the in vitro system. A number of mathematical models have been developed to consider these various factors and allow the free concentration in the in vitro system to be calculated. PBPK models have been used to predict the concentrations of chemicals achieved in the different tissues of the body following exposure by different routes (e.g., oral, dermal and inhalation; Bois et al. 2010). Traditionally, PBPK models have been built to describe the measured tissue concentration versus time profile in animals and then adapted to predict concentration in human tissues. Over recent years, the development of in vitro–in vivo extrapolation (IVIVE) techniques to predict drug clearance together with mechanistic models to predict tissue distribution (Poulin and Theil 2002) has enabled direct prediction of concentrations in humans and has led to a more widespread use of PBPK models (Rostami-Hodjegan 2012), including the use of such models to support regulatory decision-making in the pharmaceutical industry. Applying a systems approach to PBPK models where information on the biological system and chemical-specific properties are separated makes it possible to use the PBPK model to simulate exposure scenarios in different human subpopulations. For instance, the physiology can be changed to reflect the expected physiology (e.g., tissue composition, blood flow, enzymatic makeup of chemical metabolizing enzymes) in individuals of different ages (i.e., pediatric or geriatric subjects) and the simulated exposures compared to those expected in a healthy, young adult population (Polasek et al. 2013).

Demonstration that the techniques applied to pharmaceutical compounds can also be successfully used to predict the in vivo tissue and plasma concentrations of chemicals in humans from a broader chemical space together with an assessment of the uncertainty associated with the predictions would enable these IVIVE-PBPK techniques to be more widely applied in chemical risk assessment. Currently, all pharmacokinetics and quantitative hazard information, including modulatory factors and considerations of quantitative thresholds, is integrated in the context of the IATA concept (Tsaioun et al. 2016).

How to deal with species differences

AOPs consider taxonomic applicability, but in practice much of the weight-of-evidence information is derived from various animal models and the species transferability (extrapolation to human situation) is in most cases not exactly known (Leist and Hartung 2013). However, this lack of knowledge of species transferability is not an inherent limitation of the AOP concept; it is rather a consequence of a lack of adequate data. Rather than being a disadvantage, the establishment of MIEs and KEs is a precondition of a thorough interspecies comparison. If in vitro systems with human and animal cells are available that allow analysis and generation of in vivo relevant data of the most critical MIE or KE, a comparison between species becomes possible. In the past, this principle has already been used extensively, for example to understand interspecies differences in the toxicity of TCDD, tamoxifen or others. Given the multitude and variability of target species, interspecies differences are even more pronounced in ecotoxicology.

How to deal with pathway plasticity

The AOP concept aims to link MIE via a strictly defined series of KEs to AO. Then, the overall idea is that chemical hazard would be evaluated, based on knowledge of which AOP is triggered by a given substance. Although MIE are not known for the majority of chemicals, such information may be obtained using comprehensive test batteries. Although this concept appears straightforward, the complexity underlying the different possible responses of organisms to chemicals should not be underestimated, as they often do not follow a single pre-defined path. The most troublesome confounder is that the path may change depending on pathological situations. For this reason, not only simple screen systems are used, but also complex in vitro models that allow several pathological features to be monitored at the same time. On this basis, the objective of the project is to build new AOPs/AOP networks, ideally including quantitative information and attempting to fill important gaps of knowledge. For instance, interlobular bile ducts respond to

cholestasis by elongation and branching, as well as corrugation of their inner surface and upregulation of bile salt carriers, which altogether enhance their capacity to reabsorb bile salts from the duct lumen and transport them into the neighboring portal vein (Vartak et al. 2016). This protects cholangiocytes from too high concentrations of bile salts in the lumen of ducts, while simultaneously leading to increased blood concentrations of bile salts, an effect that is further enhanced by a blockage in the basolateral uptake of bile salts from the blood of liver sinusoids into hepatocytes during cholestasis (Jansen et al. 2017). Overall, the liver is protected from bile salt toxicity at the expense of higher blood concentrations, which after some time may lead to kidney toxicity. Similar diversions of pathways are known from the cell death field (Leist and Jäättelä 2001). For instance, a standard AOP may have caspase activation as KE (Gerhardt et al. 2001; Schierle et al. 1999; Hansson et al. 2000). When caspases are inhibited, the AOP does not stop, but new connections, not initially active, are formed, and e.g., calpains, cathepsins or the proteasome are activated as alternative proteases (Foghsgaard et al. 2001; Volbracht et al. 1999, 2001a, b; Hirt et al. 2000). This may involve the switch from mitochondria to lysosomes as sub-cellular target and is modified by multiple factors such as cellular chaperones (Fehrenbacher et al. 2004; Hansson et al. 2003; Nylandsted et al. 2002). Apart from altered proteolytic pathways, stress response pathways are responsible for altering the transcriptional program in an attempt to redress the original perturbation and reinstate cellular homeostasis (Jennings et al. 2013). Thus, a certain KE may cause negative feedback loops due to induction of stress response pathways, and the cellular regulation and the basis for AOP may be strongly altered.

From AOPs to quantitative AOPs (qAOP)

As stated above, AOP have different uses, many of them not requiring quantitative data. A future perspective is that next-generation AOPs may also be applied in risk assessment of chemicals. For this purpose, quantitative thresholds need to be known. It is at present not clear how this will exactly be achieved. The AOP handbook states that some regulatory applications "will require that quantitative relationships between KEs be defined". Quantifying certain KERs might prove to be a complex task, especially if the KEs in question are in different biological domains, for example if one is at the transcript level and the other is at an enzyme activity level or organelle function. More detailed experimental evidence might be needed to better define the quantitative relationship of the two KEs. When the molecular interactions will have been sufficiently elucidated, simple equations or more complex mathematical models could be applied. Indeed, such activity is being supported by another closely related OECD initiative, the Effectopedia. Effectopedia is defined as 'an open-knowledge and structured platform able to display quantitative information on AOPs. It represents a collaborative tool designed to facilitate the interdisciplinary efforts for delineating AOPs in an encyclopedic manner with greater predictive power.' While many AOPs will be developed, it is likely that KEs and KERs will be shared and thus will be interchangeable modular components. Thus, the development of mathematical models for KERs can be deployed across several AOPs, increasing their value. One such example is the development of systems biology models for the Nrf2 pathway (Hamon et al. 2014).

An important precondition for qAOP is knowledge on basic homeostatic processes, which are supposed to be perturbed by toxicant exposure. In a dynamic process such as liver steatosis, for example, there is no hope of developing a correct qAOP if we do not have a detailed understanding and description of fatty acid balance in hepatocytes. This is where toxicology meets general systems biology and cooperation with other disciplines is particularly helpful. An approach of only exposing cells to toxicants is not sufficient to feed the information requirements of qAOPs. In addition, information on gene and metabolic regulatory networks and signaling pathways will be required. In conclusion, moving from AOPs to qAOPs would represent an important progress, but its implementation remains a challenging future perspective.

Perspectives and outlook: the low and the high hanging fruits

The AOP concept represents an important research opportunity and has helped to integrate molecular mechanisms into the field of regulatory toxicology. First applications of AOPs are emerging, but it appears as if there is still a long way to go until AOP-based hazard assessment based on in vitro MIE and KE testing will replace the current toxicological approaches. Currently, it cannot be expected that AOPs assembled only on the basis of historical research not specifically designed for AOP development will be sufficient for the regulation of chemicals. A dedicated research program to specifically confirm and validate AOPs and to add quantitative information on KERs is required. There is however an expectation that the AOP conceptual framework would help in pushing through a mechanistic shift in risk assessment by making use of all available information in a weight of evidence approach and/or forming the basis for a mechanistically driven testing strategy. In the short term, it is also expected that the framework will be used as a scaffold for developing guidance for complex hazard characterization, providing a mechanistic contextualization of apical end points of concern.

There are relatively low hanging fruits in this field of research that can be readily reached. These are mainly in the areas of priority setting and pre-screening to provide alerts for downstream testing or implementing tiered approaches in the risk assessment of chemical substances instead of relying exclusively on large, expensive and sometime poorly predictive in vivo regulatory studies. For this purpose, full quantitative chains of events are not necessarily needed. What is required is individual KEs with a high plausibility value (with respect to a link to an AO) that can be studied in vitro in a concentration-dependent manner. Examples are compromised migration of neuronal precursors or compromised mitochondrial functions. Alternatively, interference with important nuclear receptors may be probed with a comprehensive test battery. In a regulatory context, the information from such KE-based testing can then be used to prioritize chemicals for further, more complex testing. Similarly, such mechanistic tests are often very informative for in-house decision-making in pharmaceutical and chemical industries to discard toxic compounds early in development. Moreover, first information can be obtained on relevant internal concentrations associated with adverse effects.

The next years are likely to bring profound changes and progress in this field. It will be exciting to see how work on the AOP platform can help to both fill in data (on existing and new AOPs) and shape the concept. In this sense, a revised version of this review published at regular intervals in the future could be a valuable approach to document and follow the developments.

Acknowledgements This work was supported by the EU-ToxRisk project (An Integrated European "Flagship" Program Driving Mechanism-Based Toxicity Testing and Risk Assessment for the 21st Century) funded by the European Commission under the Horizon 2020 programme (Grant Agreement No. 681002). This review is a joint activity of members of the EU-ToxRisk project and additionally incorporates ideas and suggestions of several colleagues who are not members of this consortium. We thank in particular Brigitte Landesmann, Maurice Whelan, Anna Bal-Price, and Christian Desaintes for their valuable discussion. Some work on examples was funded by the German Federal Ministry of Education and Research (BMBF) (LiSyM, SysDT, NeuriTox, LivSys, Lebersimulator projects). We thank M. Turajski and B. Schanze for valuable bibliographic support, and B. Barton for excellent support in handling the manuscript.

References

Ames BN, Whitfield HJ Jr (1966) Frameshift mutagenesis in Salmonella. Cold Spring Harb Symp Quant Biol 31:221–225

Andersen ME, Krewski D (2009) Toxicity testing in the 21st century: bringing the vision to life. Toxicol Sci 107:324–330

Andersen ME, Krewski D (2010) The vision of toxicity testing in the 21st century: moving from discussion to action. Toxicol Sci 117:17–24

Ankley GT, Bennett RS, Erickson RJ et al (2010) Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem 29:730–741

Ball N, Cronin MTD, Shen J et al (2016) Toward good read-across practice (GRAP) guidance. ALTEX 33:149–166

Balmer NV, Leist M (2014) Epigenetics and transcriptomics to detect adverse drug effects in model systems of human development. Basic Clin Pharmacol Toxicol 115:59–68

Balmer NV, Weng M, Zimmer B et al (2012) Epigenetic changes and disturbed neural development in a human embryonic stem cell-based model relating to the fetal valproate syndrome. Hum Mol Genet 21:4104–4114

Balmer NV, Klima S, Rempel E et al (2014) From transient transcriptome responses to disturbed neurodevelopment: role of histone acetylation and methylation as epigenetic switch between reversible and irreversible drug effects. Arch Toxicol 88:1451–1468

Bal-Price A, Crofton K, Leist M et al (2015a) International STakeholder NETwork for developmental neurotoxicity (ISTNET): creating a developmental neurotoxicity (DNT) testing roadmap for regulatory purposes. Arch Toxicol 89:269–287

Bal-Price A, Crofton KM, Sachana M et al (2015b) Putative adverse outcome pathways relevant to neurotoxicity. Crit Rev Toxicol 45:83–91

Bal-Price A, Lein PJ, Keil KP et al (2017) Developing and applying the adverse outcome pathway concept for understanding and predicting neurotoxicity. Neurotoxicology 59:240–255

Basketter DA, Clewell H, Kimber I et al (2012) A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing—t4 report. ALTEX 29:3–91

Bataller R, Brenner DA (2005) Liver fibrosis. J Clin Investig 15:209–218

Baumann J, Gassmann K, Masjosthusmann S et al (2016) Comparative human and rat neurospheres reveal species differences in chemical effects on neurodevelopmental key events. Arch Toxicol 90:1415–1427

Becker RA, Ankley GT, Edwards SW et al (2015) Increasing scientific confidence in adverse outcome pathways: application of tailored Bradford-Hill considerations for evaluating weight of evidence. Regul Toxicol Pharmacol 72:514–537

Beliaeff B, Burgeot T (2002) Integrated biomarkers response: a useful tool for ecological risk assessment. Environ Toxicol Chem 21:1316–1322

Bhattacharya S (1948) A test for mutagenicity of methylcholanthrene. Nature 162:573

Blaauboer BJ, Boekelheide K, Clewell HJ et al (2012) The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans. Altex 29:411–425

Bois FY, Jamei M, Clewell HJ (2010) PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. Toxicology 278:256–267

Bolt HM, Foth H, Hengstler JG, Degen GH (2004) Carcinogenicity categorization of chemicals-new aspects to be considered in a European perspective. Toxicol Lett 151(1):29–41

Bouhifd M, Andersen ME, Baghdikian C et al (2015) The human toxome project. ALTEX 32:112–124

Casey WM (2016) Advances in the development and validation of test methods in the United States. Toxicol Res 32:9–14

Carbonell P, Lopez O, Amberg A, Pastor M, Sanz F (2017) Hepatotoxicity prediction by systems biology modeling of disturbed metabolic pathways using gene expression data. ALTEX 34(2):219-234

Clippinger AJ, Hill E, Curren R et al (2016) Bridging the gap between regulatory acceptance and industry use of non-animal methods. ALTEX 33:453–458

Crawford SE, Hartung T, Hollert H et al (2017) Green toxicology: a strategy for sustainable chemical and material development. Environ Sci Eur 29:16

Daneshian M, Kamp H, Hengstler J et al (2016) Highlight report: launch of a large integrated European in vitro toxicology project: EU-ToxRisk. Arch Toxicol 90:1021–1024

Daston G, Knight DJ, Schwarz M et al (2015) SEURAT: safety evaluation ultimately replacing animal testing—recommendations for future research in the field of predictive toxicology. Arch Toxicol 89:15–23

Delrue N, Sachana M, Sakuratani Y et al (2016) The adverse outcome pathway concept: a basis for developing regulatory decision-making tools. Altern Lab Anim 44:417–429

Depledge MH (1994) The rational basis for the use of biomarkers as ecotoxicological tools. In: Fossi MC, Leonzio C (ed.) Nondestructive biomarkers in vertebrates, pp. 271-295

Drasdo D, Hoehme S, Hengstler JG (2014) How predictive quantitative modelling of tissue organisation can inform liver disease pathogenesis. J Hepatol 61:951–956

Duffield JS, Forbes SJ, Constandinou CM et al (2005) Selective depletion of macrophages reveals distinct, opposing roles during liver injury and repair. J Clin Invest 115:56–65

Edwards SW, Tan YM, Villeneuve DL et al (2016) Adverse outcome pathways-organizing toxicological information to improve decision making. J Pharmacol Exp Ther 356(1):170–181

Fasbender F, Widera A, Hengstler JG et al (2016) Natural killer cells and liver fibrosis. Front Immunol 7:19

Fehrenbacher N, Gyrd-Hansen M, Poulsen B et al (2004) Sensitization to the lysosomal cell death pathway upon immortalization and transformation. Cancer Res 64:5301–5310

Fischer FC et al (2017) Modeling exposure in the Tox21 in vitro bioassays. Chem Res Toxicol 30(5):1197–1208

Foghsgaard L, Wissing D, Mauch D et al (2001) Cathepsin B acts as a dominant execution protease in tumor cell apoptosis induced by tumor necrosis factor. J Cell Biol 153:999–1010

Gantner F, Leist M, Küsters S et al (1996) T cell stimulus-induced crosstalk between lymphocytes and liver macrophages results in augmented cytokine release. Exp Cell Res 229:137–146

Gassmann K, Abel J, Bothe H et al (2010) Species-specific differential AhR expression protects human neural progenitor cells against developmental neurotoxicity of PAHs. Environ Health Perspect 118:1571–1577

Gerhardt E, Kügler S, Leist M et al (2001) Cascade of caspase activation in potassium-deprived cerebellar granule neurons: targets for treatment with peptide and protein inhibitors of apoptosis. Mol Cell Neurosci 17:717–731 **(PubMed PMID: 11312607)**

Ghallab A, Cellière G, Henkel SG et al (2016) Model-guided identification of a therapeutic strategy to reduce hyperammonemia in liver diseases. J Hepatol 64:860–871

Godoy P, Hewitt NJ, Albrecht U et al (2013) Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Arch Toxicol 87:1315–1530

Grinberg M, Stöber RM, Edlund K et al (2014) Toxicogenomics directory of chemically exposed human hepatocytes. Arch Toxicol 88(12):2261–2287

Guidance document on developing and assessing adverse outcome pathways (2017) Environment directorate joint meeting of the chemicals committee and the working party on chemicals, pesticides and biotechnology; Series on Testing and Assessment No. 184. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2013)6&doclanguage=en. Accessed 4 Sept 2017

Hamon J, Jennings P, Bois FY (2014) Systems biology modeling of omics data: effect of cyclosporine a on the Nrf2 pathway in human renal cells. BMC Syst Biol 8:76

Hansson O, Castilho RF, Kaminski Schierle GS et al (2000) Additive effects of caspase inhibitor and lazaroid on the survival of transplanted rat and human embryonic dopamine neurons. Exp Neurol 164:102–111

Hansson O, Nylandsted J, Castilho RF et al (2003) Overexpression of heat shock protein 70 in R6/2 Huntington's disease mice has only modest effects on disease progression. Brain Res 970:47–57

Hartung T (2009) Food for thought… on evidence-based toxicology. ALTEX 26:75–82

Hartung T (2013) Look back in anger—what clinical studies tell us about preclinical work. ALTEX 30:275–291

Hartung T (2016) Making big sense from big data in toxicology by read-across. ALTEX 33:83–93

Hartung T (2017) Utility of the adverse outcome pathway concept in drug development. Expert Opin Drug Metab Toxicol 13:1–3

Hartung T, McBride M (2011) Food for thought… on mapping the human toxome. ALTEX 28:83–93

Hartung T, van Vliet E, Jaworska J et al (2012) Systems toxicology. ALTEX 29:119–128

Hartung T, Luechtefeld T, Maertens A et al (2013a) Integrated testing strategies for safety assessments. ALTEX 30:3–18

Hartung T, Stephens M, Hoffmann S (2013b) Mechanistic validation. ALTEX 30:119–130

Hartung T, FitzGerald R, Paul J et al (2017) Systems toxicology—real world applications and opportunities. Chem Res Toxicol 30:870–882

Hengstler JG, Arand M, Herrero ME et al (1998) Polymorphisms of *N*-acetyltransferases, glutathione S-transferases, microsomal epoxide hydrolase and sulfotransferases: influence on cancer susceptibility. Recent Results Cancer Res 154:47–85

Hengstler JG, Van der Burg B, Steinberg P et al (1999) Interspecies differences in cancer susceptibility and toxicity. Drug Metab Rev 31:917–970

Hengstler JG, Bogdanffy MS, Bolt HM et al (2003) Challenging dogma: thresholds for genotoxic carcinogens? The case of vinyl acetate. Annu Rev Pharmacol Toxicol 43:485–520

Hengstler JG, Marchan R, Leist M (2012) Highlight report: towards the replacement of in vivo repeated dose systemic toxicity testing. Arch Toxicol 86:13–15

Hirt U, Gantner F, Leist M (2000) Phagocytosis of non-apoptotic cells dying by caspase-independent mechanisms. J Immunol 164:6520–6529

Hoehme S, Hengstler JG, Brulport M et al (2007) Mathematical modelling of liver regeneration after intoxication with CCl(4). Chem Biol Interact 168:74–93

Hoehme S, Brulport M, Bauer A et al (2010) Prediction and validation of cell alignment along microvessels as order principle to restore tissue architecture in liver regeneration. Proc Natl Acad Sci USA 107:10371–10376

Hoffmann S, Hartung T (2006) Towards an evidence-based toxicology. Hum Exp Toxicol 25:497–513

Hoffmann S, de Vries RBM, Stephens ML et al (2017) A primer on systematic reviews in toxicology. Arch Toxicol 91(7):2551–2575

Horvat T, Landesmann B, Lostia A et al (2017) Adverse outcome pathway development from protein alkylation to liver fibrosis. Arch Toxicol 91:1523–1543

Huggett RJ, Kimerle RA, Mehrle PM, Bergman HL (1992) Biomarkers. Biochemical, physiological, and histological markers of anthropogenic stress. Lewis Publishers, Boca Raton, p 347

Jacobs MN, Colacci A, Louekari K et al (2016) International regulatory needs for development of an IATA for non-genotoxic carcinogenic chemical substances. ALTEX 33:359–392

Jalan R, Moreau R, Kamath PS, Arroyo V (2016) Acute-on-chronic liver failure: a distinct clinical condition. Semin Liver Dis 36(2):107–108

Jansen PL, Ghallab A, Vartak N et al (2017) The ascending pathophysiology of cholestatic liver disease. Hepatology 65(2):722–738
Jennings P (2013) Stress response pathways, toxicity pathways and adverse outcome pathways. Arch Toxicol 87:13–14
Jennings P, Limonciel A, Felice L et al (2013) An overview of transcriptional regulation in response to toxicological insult. Arch Toxicol 87:49–72
Kim KH, Chen CC, Monzon RI et al (2013) Matricellular protein CCN1 promotes regression of liver fibrosis through induction of cellular senescence in hepatic myofibroblasts. Mol Cell Biol 33:2078–2090
Kisseleva T, Cong M, Paik Y et al (2012) Myofibroblasts revert to an inactive phenotype during regression of liver fibrosis. Proc Natl Acad Sci USA 109:9448–9453
Kleensang A, Maertens A, Rosenberg M et al (2014) Pathways of toxicity. ALTEX 31:53–61
Kleinstreuer NC, Ceger PC, Allen DG et al (2016) A curated database of rodent uterotrophic bioactivity. Environ Health Perspect 124:556–562
Kleinstreuer NC, Ceger P, Watt ED et al (2017) Development and validation of a computational model for androgen receptor activity. Chem Res Toxicol 30:946–964
Knapen D, Vergauwen L, Villeneuve DL et al (2015) The potential of AOP networks for reproductive and developmental toxicity assay development. Reprod Toxicol 56:52–55
Krizhanovsky V, Yon M, Dickins RA et al (2008) Senescence of activated stellate cells limits liver fibrosis. Cell 134:657–667
Krug AK, Balmer NV, Matt F et al (2013a) Evaluation of a human neurite growth assay as specific screen for developmental neurotoxicants. Arch Toxicol 87:2215–2231
Krug AK, Kolde R, Gaspar JA et al (2013b) Human embryonic stem cell-derived test systems for developmental neurotoxicity: a transcriptomics approach. Arch Toxicol 87:123–143
Krug AK, Gutbier S, Zhao L et al (2014) Transcriptional and metabolic adaptation of human neurons to the mitochondrial toxicant MPP(+). Cell Death Dis. 5:e1222
Kuegler PB, Zimmer B, Waldmann T et al (2010) Markers of murine embryonic and neural stem cells, neurons and astrocytes: reference points for developmental neurotoxicitiy testing—a review by the Transatlantic Think Tank for toxicology (t4). ALTEX 27:17–42
Latta M, Künstle G, Leist M et al (2000) Metabolic depletion of ATP by fructose inversely controls CD95- and tumor necrosis factor receptor 1-mediated hepatic apoptosis. J Exp Med 191:1975–1985
Leist M, Hartung T (2013) Inflammatory findings on species extrapolations: humans are definitely no 70-kg mice. Arch Toxicol 87:563–567
Leist M, Jäättelä M (2001) Four deaths and a funeral: from caspases to alternative mechanisms. Nat Rev Mol Cell Biol 2:589–598
Leist M, Nicotera P (1997) Calcium and neuronal death. Rev Physiol Biochem Pharmacol 132:79–125
Leist M, Gantner F, Künstle G et al (1996) The 55-kD tumor necrosis factor receptor and CD95 independently signal murine hepatocyte apoptosis and subsequent liver failure. Mol Med 2:109–124
Leist M, Fava E, Montecucco C, Nicotera P (1997a) Peroxynitrite and nitric oxide donors induce neuronal apoptosis by eliciting autocrine excitotoxicity. Eur J Neurosci 9:1488–1498
Leist M, Gantner F, Naumann H et al (1997b) Tumor necrosis factor-induced apoptosis during the poisoning of mice with hepatotoxins. Gastroenterology 112:923–934
Leist M, Volbracht C, Kühnle S et al (1997c) Caspase-mediated apoptosis in neuronal excitotoxicity triggered by nitric oxide. Mol Med 3:750–764
Leist M, Gantner F, Künstle G et al (1998a) Cytokine-mediated hepatic apoptosis. Rev Physiol Biochem Pharmacol 133:109–155
Leist M, Volbracht C, Fava E et al (1998b) 1-Methyl-4-phenylpyridinium induces autocrine excitotoxicity, protease activation, and neuronal apoptosis. Mol Pharmacol 54:789–801
Leist M, Hartung T, Nicotera P (2008) The dawning of a new age of toxicology. ALTEX 25:103–114
Leist M, Efremova L, Karreman C (2010) Food for thought on considerations and guidelines for basic test method descriptions in toxicology. ALTEX 27:309–317
Leist M, Hasiwa N, Daneshian M et al (2012) Validation and quality control of replacement alternatives—current status and future challenges. Toxicol Res 1:8–22
Leist M, Hasiwa N, Rovida C et al (2014) Consensus report on the future of animal-free systemic toxicity testing. ALTEX 31:341–356
Lewin L (1885) Lehrbuch der Toxikologie: für Aerzte, Studierende und Apotheker (translation: textbook of Toxicology for physicians, students and pharmacists). Urban & Schwarzenberg, Wien
Lewin L (1924) Phantastica. Die betäubenden und erregenden Genussmittel. Für Ärzte und Nichtärzte (translation: Phantastica. Anaesthetic and stimulating drugs for physicians and laymen). Verlag von Georg Stilke, Berlin
Limonciel A, Aschauer L, Wilmes A et al (2011) Lactate is an ideal non-invasive marker for evaluating temporal alterations in cell stress and toxicity in repeat dose testing regimes. Toxicol In Vitro 25:1855–1862
Maertens A, Anastas N, Spencer PJ et al (2014) Green Toxicology. ALTEX 31:243–249
Matozzo V, Gagné F, Marin MG, Ricciardi F, Blaise C (2008) Vitellogenin as a biomarker of exposure to estrogenic compounds in aquatic invertebrates: a review. Environ Int 34:531–545
McCarty JF, Shugart LR (1990) Biomarkers of environmental contamination. Lewis Publishers, CRC Press, Boca Raton, p 457
Nicotera P, Leist M, Manzo L (1999) Neuronal cell death: a demise with different shapes. Trends Pharmacol Sci 20(2):46–51
NRC (2007) Toxicity testing in the 21st century: a vision and a strategy. The National Academies Press, Washington
NTP (2017) https://ntp.niehs.nih.gov/pubhealth/evalatm/test-method-evaluations/immunotoxicity/nonanimal/index.html. Accessed 4 Sept 2017
Nyffeler J, Karreman C, Leisner H et al (2017) Design of a high-throughput human neural crest cell migration assay to indicate potential developmental toxicants. ALTEX 34:7555–7594
Nylandsted J, Wick W, Hirt UA et al (2002) Eradication of glioblastoma, and breast and colon carcinoma xenografts by Hsp70 depletion. Cancer Res 62:7139–7142
Obiol-Pardo C, Gomis-Tena J, Sanz F, Saiz J, Pastor M (2011) A multiscale simulation system for the prediction of drug-induced cardiotoxicity. J Chem Inf Model 51:483–492
Ockleford C et al (2017) EFSA panel on plant protection products and their residues (PPR) investigation into experimental toxicological properties of plant protection products having a potential link to Parkinson's disease and childhood leukaemia. EFSA J 15:4691
Oesch F, Herrero ME, Hengstler JG, Lohmann M, Arand M (2000) Metabolic detoxification: implications for thresholds. Toxicol Pathol 28(3):382–387
Oesch F, Herrero ME, Lohmann M, Hengstler JG, Arand M (2001) Sequestration of biological reactive intermediates by trapping as covalent enzyme-intermediate complex. Adv Exp Med Biol 500:577–586
Oki NO, Edwards SW (2016) An integrative data mining approach to identifying adverse outcome pathway signatures. Toxicology 28(350–352):49–61
Paparella M, Colacci A, Jacobs MN (2017) Uncertainties of testing methods: what do we (want to) know about carcinogenicity? ALTEX 34(2):235–252

Patlewicz G, Ball N, Becker RA et al (2014) Read-across approaches—misconceptions, promises and challenges ahead. ALTEX 31:387–396

Pelkonen O, Terron A, Hernandez AF et al (2017) Chemical exposure and infant leukaemia: development of an adverse outcome pathway (AOP) for aetiology and risk assessment research. Arch Toxicol 91(8):2763–2780

Polasek TM, Patel F, Jensen BP, Sorich MJ, Wiese MD, Doogue MP (2013) Predicted metabolic drug clearance with increasing adult age. Br J Clin Pharmacol 75(4):1019–1028

Poulin P, Theil FP (2002) Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. J Pharm Sci 91:129–156

Radaeva S, Sun R, Jaruga B et al (2006) Natural killer cells ameliorate liver fibrosis by killing activated stellate cells in NKG2D-dependent and tumor necrosis factor-related apoptosis-inducing ligand-dependent manners. Gastroenterology 130:435–452

Rahnenführer J, Leist M (2015) From smoking guns to footprints: mining for critical events of toxicity pathways in transcriptome data. Arch Toxicol 89:813–817

Ramachandran P, Iredale JP (2012) Macrophages: central regulators of hepatic fibrogenesis and fibrosis resolution. J Hepatol 56:1417–1419

Ramachandran P, Pellicoro A, Vernon MA et al (2012) Differential Ly-6C expression identifies the recruited macrophage phenotype, which orchestrates the regression of murine liver fibrosis. Proc Natl Acad Sci USA 109:E3186–E3195

Rempel E, Hoelting L, Waldmann T et al (2015) A transcriptome-based classifier to identify developmental toxicants by stem cell testing: design, validation and optimization for histone deacetylase inhibitors. Arch Toxicol 89:1599–1618

Rostami-Hodjegan A (2012) Physiologically based pharmacokinetics joined with in vitro–in vivo extrapolation of ADME: a marriage under the arch of systems pharmacology. Clin Pharmacol Ther 92(1):50–61

Rovida C, Alépée N, Api AM et al (2015) Integrated testing strategies (ITS) for safety assessment. ALTEX 32:171–181

Samuel GO, Hoffmann S, Wright R et al (2016) Guidance on assessing the methodological and reporting quality of toxicologically relevant studies: a scoping review. Environ Int 92–93:630–646

Sauer JM, Hartung T, Leist M et al (2015) Systems toxicology: the future of risk assessment. Int J Toxicol 34:346–348

Schierle GS, Hansson O, Leist M et al (1999) Caspase inhibition reduces apoptosis and increases survival of nigral transplants. Nat Med 5:97–100

Schildknecht S, Pöltl D, Nagel DM et al (2009) Requirement of a dopaminergic neuronal phenotype for toxicity of low concentrations of 1-methyl-4-phenylpyridinium to human cells. Toxicol Appl Pharmacol 241:23–35

Schildknecht S, Pape R, Müller N et al (2011) Neuroprotection by minocycline caused by direct and specific scavenging of peroxynitrite. J Biol Chem 286:4991–5002

Schildknecht S, Karreman C, Pöltl D et al (2013) Generation of genetically-modified human differentiated cells for toxicological tests and the study of neurodegenerative diseases. ALTEX 30:427–444

Schildknecht S, Pape R, Meiser J et al (2015) Preferential extracellular generation of the active parkinsonian toxicant MPP+ by transporter-independent export of the intermediate MPDP+. Antioxid Redox Signal 23:1001–1016

Schildknecht S, Di Monte DA, Pape R et al (2017) Tipping points and endogenous determinants of nigrostriatal degeneration by MPTP. Trends Pharmacol Sci 38:541–555

Schliess F, Hoehme S, Henkel SG et al (2014) Integrated metabolic spatial-temporal model for the prediction of ammonia detoxification during liver damage and regeneration. Hepatology 60:2040–2051

Shinde V, Hoelting L, Perumal SS et al (2016) Definition of transcriptome-based indices for quantitative characterization of chemically disturbed stem cell development—introduction of the STOP-Toxukn and STOP-Toxukk t. Arch Toxicol 91:839–864

Smirnova L, Hogberg HT, Leist M et al (2014) Developmental neurotoxicity—challenges in the 21st century and in vitro opportunities. ALTEX 31:129–156

Stephens ML, Betts K, Beck NB et al (2016) The emergence of systematic review in toxicology. Toxicol Sci 152:10–16

Stiegler N, Krug A, Matt F et al (2011) Assessment of chemical-induced impairment of human neurite outgrowth by multiparametric live cell imaging in high-density cultures. Toxicol Sci 121:73–87

Tacke F, Zimmermann HW (2014) Macrophage heterogeneity in liver injury and fibrosis. J Hepatol 60:1090–1096

Tian Z, Chen Y, Gao B (2013) Natural killer cells in liver disease. Hepatology 57:1654–1662. doi:10.1002/hep.26115

Tollefsen KE, Scholz S, Cronin MT et al (2014) Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). Reg Toxicol Pharmacol 70:629–640

Trevan JW (1927) The error of determination of toxicity. Proc R Soc Lond B Biol Sci 101(712):483–514

Tsaioun K, Blaauboer BJ, Hartung T (2016) Evidence-based absorption, distribution, metabolism, excretion (ADME) and its interplay with alternative toxicity methods. ALTEX 33:343–358

Users' handbook supplement to the guidance document for developing and assessing AOPs (2017) http://www.oecd-ilibrary.org/environment/users-handbook-supplement-to-the-guidance-document-for-developing-and-assessing-adverse-outcome-pathways_5jlv1m9d1g32-en. Accessed 4 Sept 2017

van der Burg B, Pieterse B, Buist H et al (2015a) A high throughput screening system for predicting chemically-induced reproductive organ deformities. Reprod Toxicol 55:95–103

van der Burg B, Wedebye EB, Dietrich DR et al (2015b) The ChemScreen project to design a pragmatic alternative approach to predict reproductive toxicity of chemicals. Reprod Toxicol 55:114–123

van Thriel C, Westerink R, Beste C et al (2011) Translating neurobehavioural endpoints of developmental neurotoxicity tests into in vitro assays and readouts. NeuroToxicol 33:911–924

Vartak N, Damle-Vartak A, Richter B et al (2016) Cholestasis-induced adaptive remodeling of interlobular bile ducts. Hepatology 63:951–964

Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M (2014a) Adverse outcome pathway (AOP) development I: strategies and principles. Toxicol Sci 142:312–320

Villeneuve DL, Crump D, Garcia-Reyero N, Hecker M, Hutchinson TH, LaLone CA, Landesmann B, Lettieri T, Munn S, Nepelska M, Ottinger MA, Vergauwen L, Whelan M (2014b) Adverse outcome pathway development. II: best practices. Toxicol Sci 142:321–330

Vinken M (2015) Adverse outcome pathways and drug-induced liver injury testing. Chem Res Toxicol 28:1391–1397

Volbracht C, Leist M, Nicotera P (1999) ATP controls neuronal apoptosis triggered by microtubule breakdown or potassium deprivation. Mol Med. 5:477–489

Volbracht C, Fava E, Leist M et al (2001a) Calpain inhibitors prevent nitric oxide-triggered excitotoxic apoptosis. NeuroReport 12:3645–3648

Volbracht C, Leist M, Kolb SA et al (2001b) Apoptosis in caspase-inhibited neurons. Mol Med 7:36–48

Wilmes A, Limonciel A, Aschauer L et al (2013) Application of integrated transcriptomic, proteomic and metabolomic profiling for the delineation of mechanisms of drug induced cell stress. J Proteomics 79:180–194

Wilmes A, Bielow C, Ranninger C et al (2015) Mechanism of cisplatin proximal tubule toxicity revealed by integrating transcriptomics, proteomics, metabolomics and biokinetics. Toxicol In Vitro 30:117–127

Wittwehr C, Aladjov H, Ankley G et al (2017) How adverse outcome pathways can aid the development and use of computational prediction models for regulatory toxicology. Toxicol Sci 155:326–336

Wynn TA (2008) Cellular and molecular mechanisms of fibrosis. J Pathol 214:199–210

Zhu H, Bouhifd M, Kleinstreuer N et al (2016) Supporting read-across using biological data. ALTEX 33:167–182

Zimmer B, Lee G, Balmer NV et al (2012) Evaluation of developmental toxicants and signaling pathways in a functional test based on the migration of human neural crest cells. Environ Health Perspect 120:1116–1122

Zimmer B, Pallocca G, Dreser N et al (2014) Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery. Arch Toxicol 88:1109–1126