Chapter 1

General introduction and background.

1. Drug-induced liver injury: a major problem in drug development.

1.1. The liver.

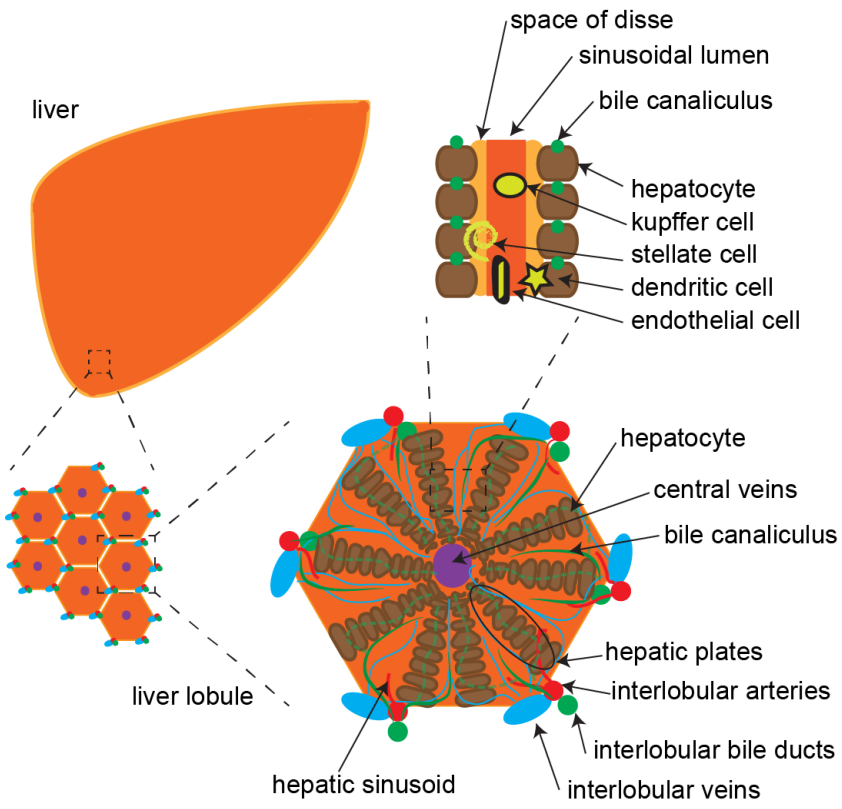
The liver is the main metabolizing organ and the first organ encountered by nutrient rich blood delivered by the hepatic vein which originates from the gastrointestinal tract. The liver consists of millions of functional subunits: individual hexagon lobules made up of the parenchym surrounded by portal triads (Fig. 1). The portal triads consist of bile ducts, hepatic arteries and portal veins. The liver consists of the parenchyme and non-parenchyme cell types. More than 70% of the liver is composed of the parenchyme: the hepatocytes. Hepatocytes are responsible for most of the liver metabolizing capacity and contain the drug metabolizing enzymes and transporters responsible for xenobiotic transformation. Non parenchymal cells are cholangiocytes (bile duct epithelial cells), sinusoidal cells (which form triad walls) and several immune related cells such as the stellate cells, Kupfer cells and pit cells. Because hepatocytes are by far the most abundant cell type and responsible for the liver metabolizing capacity and detoxification/xenobiotic metabolism these cells are typically used for in vitro toxicity testing.

1.1. Drug-induced liver injury.

Drug-induced liver injury (DILI) constitutes liver injury as the result of drug treatments with potentially fatal adverse events. DILI can be classified into a hepatic, cholestatic or mixed phenotype. Hepatic DILI involves damage and cell death of the parenchyme which are the hepatocytes. Cholestasis involves perturbations due to altered bile acid metabolism and transport culminating in accumulation of bile in the liver or damaged bile canaliculi. Additional DILI pathophysiological phenotypes include: steatosis, the accumulation of fat droplet in the liver due to fatty acid metabolism; phospholipidosis, the accumulation of phospholipids in the liver; inflammation, the infiltration of leucocytes in the liver; and fibrosis, the increase of scarred fibrotic tissue in the liver.

1.2. Societal impact of DILI.

The market for pharmaceutical compounds in the EU is estimated at a yearly turnover of 205 billion Euro (Eurostat data 2015) with an estimated average of 0.6 % (1.2 billion) spent on drug safety (European Commission pharmacovigilance report, 2008). However the social financial impact of preventable adverse drug reactions (ADR's) in the same report is estimated at 24 billion annually for the European population. This estimated cost is based on a 5 % incidence of ADR's during hospitalization [1], a 5% ADR related hospital admission [2] and drug-related morbidity and mortality other than hospital admission or prolonged hospitalization [3]. Moreover, the public



health importance of ADR's is estimated at 200,000 deaths annually in the EU alone [1]. In addition ADR's lead to preclinical discontinuation and post market withdrawals.

Figure 1: **Cell types within a liver and structure of a liver lobule.** The smallest functional unit of a liver is the lobule which each has a set of interlobular -veins, -arteries and -bile ducts attributed to it and a single central vein. The liver is built up of parenchymal and several non-parenchymal cell types.

Hepatic and cardiac toxicity has contributed disproportionately to drug withdrawals: of 47 drugs withdrawn during the period 1975 - 2007, 21 were related to cardiotoxicity and 15 involved hepatic toxicity [4]. DILI is responsible for 30% of drug withdrawals from the market [5] and non-approvals by regulatory authorities.

Often no changes in hepatocellular toxicity parameters such as alanine or aspartate aminotransferase (ALT/AST) levels or increased total bilirubin are found in pre-clinical settings and drugs are marketed until more than 1 in 10,000 drug users demonstrate signs of liver failure [6].

Severe DILI in the clinical setting is most typically caused by a so called idiosyncratic reaction which by definition means dose independent and rare (<1:10,000) and thus highly unpredictable. Preclinical studies often miss these rare idiosyncratic reactions due to limitation on the feasible number of test animals and species specific differences. In addition to idiosyncratic reactions also 'normal' DILI found in preclinical animal studies or during human trials is difficult to predict and as such causes many compounds to be terminated and thus imposes significant costs for the pharmaceutical industry [7]. For these reasons major efforts are being made by industry and academia to obtain better biomarkers to better understand and predict DILI pre-clinically.

2. In vitro toxicity testing methods.

2.1. Safety assessment and the Adverse Outcome Pathway (AOP) paradigm.

The current preclinical testing paradigm has improved drug safety over the last 30 years; it is estimated that 70% of human toxicity is predicted pre-clinically [8]. Typically animal experiments are performed pre-clinically to rule out as much toxic compounds as possible. For the liver these animal studies are used to determine histopathological endpoints after acute or repeated dosing of the drug. Pathologists then determine if the liver toxicity shows signs of acute, cholestatic, steatotic, necrotic or fibrotic phenotypes. However such endpoints unfortunately do not provide

much information on the mechanism of the observed drug-induced liver toxicity. In addition to this scientific shortcoming, public pressure against animal testing, exemplified by the ban on animal testing for cosmetics in Europe, and increasing societal momentum to implement more modern scientific methods for pre-clinical testing is apparent. Several multimillion funded consortia have worked on human relevant alternatives such as SEURAT-1 and MIP-DILI in Europe and ToxCast and Tox21 in the U.S. to work on these issues. Thus there is a strong incentive for consortia to focus on implementing modern scientific methodologies which are focused on human-relevant concepts. These methodologies are founded on developing a mechanistic understanding of adverse drug reactions in *in vitro* based models for evidence and read across-based approaches for risk assessment. One such methodology developed by the OECD in 2012 being the initiation and population of the Adverse Outcome Pathway (AOP) framework. The AOPs framework is described as a "sequential chain of causally linked key events at different levels of biological organization that together culminate in an adverse health outcome or ecotoxicological effect". While several AOPs have been established, a next important step is to translate AOP-related mechanistic understanding in advanced, preferably quantitative, high throughput assays that reflect pathways essential in target organ toxicity. Such an AOP framework is highly suited as an evidence-gathering-based approach in which affected (adaptive stress) pathways can be included as a fundamental cellular response following biochemical perturbation due to chemical exposure, also branded as the molecular initiating events (see Fig. 2).

2.2. Overall in vitro methods for safety assessment.

What is the diversity of *in vitro* methods that are available? Firstly, several *in vitro* and *in silico* methods currently exist that can predict the pharmacokinetic properties and clearance of compounds [9]. Compounds that are metabolized or affect liver metabolic enzymes can be identified by *in vitro* methods. Secondly, cytotoxic endpoint assays have been developed to determine *in vitro* EC50-values for apoptosis and necrosis using colorimetric assays.

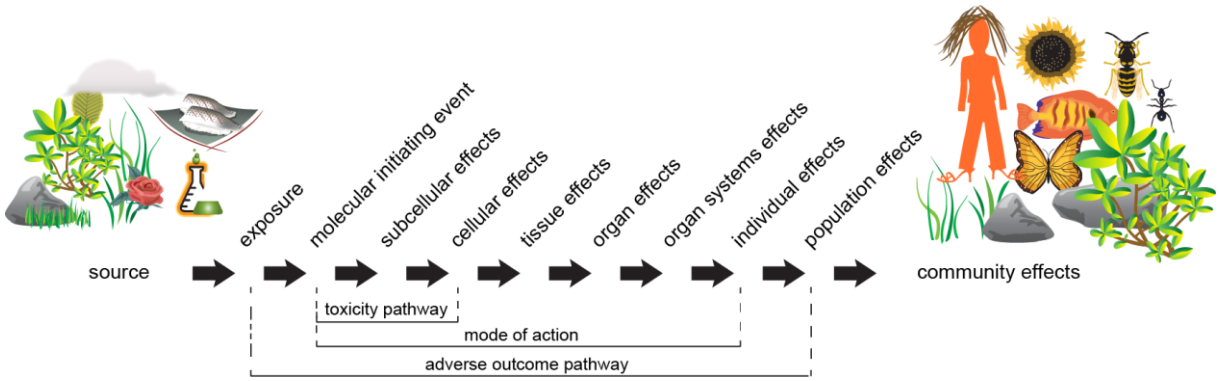


Figure 2: **The AOP framework**. An evidence gathering based approach starting from the source to the final adverse effects. Figure adapted from K. Crofton 2010, OECD AOP Meeting Definitions

Thirdly, more mechanism-based *in vitro* methods exist that can give insight into the type of toxicity in cells such as the MTT assay which reflects NAD(P)H dependent metabolic activity of cells, MitoSOX-red which is a mitochondrial superoxide indicator [10], lipid dyes which visualize phospholipid accumulation e.g. Bodipy 493/503, and the fluorescent phospholipid probe NBD-PE [11] or Fluo-4 AM to detect cytoplasm free-calcium levels [12]. In addition certain dyes are

capable of selectively binding to certain organelles or substructures such as the mitochondria, golgi, chromosomal DNA or cell membranes and perturbation of these structures after chemical exposure can be detected in this manner. Antibodies are often used in *in vitro* toxicity assays to detect perturbations of protein levels or to visualize organelles and cellular structures such as actin to detect cell-cell junctions or bile canaliculi. Polymerase chain reaction (PCR) or luciferase assays are used to monitor changes in transcript levels of certain key transcripts such as the nuclear receptors [13], anti-oxidant response element transcripts or phase I, II and III drug metabolism enzymes. Fourthly, most *in vitro* assays using the above approaches are based on monolayer cultures that do not fully mimic the *in vivo* context. More complex *in vitro* testing strategies are in development which try to mimic the 3D structure [14], multiple cell type interactions and mechanical flow and shear stress based on microfluidic reactors [15, 16].

Development of all these functional assays, largely based on an improved understanding of the cellular pathways involved on chemical induced toxicity, is increasingly based on mechanistic concepts. Within the development of these assays one of two broad routes can be defined: on the one hand the more physiological relevant liver models and on the other hand the more detailed mechanism-based predictive models. The physiological mimicking of the liver has obtained increased interest in recent years, and models such as liver slices [17], 3D models [18], pluripotent stem cell-derived differentiated cell lines [19] and liver bioreactors [16] are being developed in many investigations and some are already utilized in safety testing strategies. Mechanism-based toxicological readouts and throughput from these models are often more limited due to several technical limitations such as the inability to perform high resolution high content imaging or single cell type quantification of proteins or metabolites. Often general macro-phenotypic profiling, measurements in culture supernatant of metabolites or secretion and leakage of specific proteins from these models is performed. The more detailed mechanism-based models under development focus more on detailed mechanistic insight of molecular networks and the molecular initiating events as well as key events that eventually lead to cellular adversity (see the AOP framework earlier). More robust models that can be cultured easily and are highly reproducible are used, including the cancer-derived cell lines HepG2 and HepaRG. These models also enable functional genomics assessment of molecular mechanisms using e.g. small interfering RNA knock down approaches and the generation of genetically modified (often knock-out or tagged) cell lines; this is not possible with primary human hepatocytes. It is the combination of these two different approaches that will take science to the next level of understanding chemical-induced toxicity. This thesis focuses on the second approach: application of genetically modified hepatoma cell lines to unravel mechanism of action in the context of DILI.

3. From molecular mechanisms to DILI prediction.

3.1. Mechanisms of DILI.

For only several drugs the underlying mechanisms leading to DILI is by and large elucidated. The best known example is acetaminophen, which induces acute liver injury after overdosing. The cytochrome-P450 enzyme system metabolizes acetaminophen into more hydrophilic and reactive metabolites including N-acetyl-p-benzoquinone imine (NAPQI) [20]. At lower concentrations NAPQI does not accumulate because antioxidant molecules such as glutathione act as a redox buffer clearing the reactive NAPQI before it can bind and damage macromolecules and elicit

further significant cellular damage. However after an overdose acetaminophen, glutathione becomes depleted and cellular damage caused by NAPQI leads to acute liver damage. The ultimate cell death involves many cellular perturbations, including mitochondrial injury, increased reactive oxygen formation, activation of cellular stress pathways, that all individually contribute to the onset of cell death [21, 22].

A second well-known example is the dose-dependent toxicity of valproic acid. Valproic acid is a simple fatty acid and therefore a substrate of the β-oxidation pathway. Valproic acid and its metabolites can cause interference with e.g. mitochondrial β-oxidation, oxidative phosphorylation, and depletion of CoA culminating in the accumulation of fatty acids in liver cells eventually resulting in microvesicular steatosis [23-25].

Unfortunately for most compounds that induce severe DILI the exact molecular mechanisms are not fully understood and various plausible mechanisms for individual drugs have been proposed that could contribute to DILI. Below various molecular mechanisms that contribute to DILI are discussed in more detail with a focus on several key adaptive stress response programs.

3.2. Mitochondria injury, oxidative stress & apoptosis: an intricate interplay.

Mitochondrial dysfunction is considered a key component to the overall mechanism of many DILI related drugs [26]. In the literature many examples can be found involving drug-induced impairment of mitochondrial fatty acid oxidation (see above the example of valproic acid), electron transfer within the respiratory chain, oxidative phosphorylation and mitochondrial DNA damage [27]. The mitochondria are central to cellular metabolism as the reduction of oxygen to water fueled by organic catabolic processes provides the necessary energy in the form of ATP. With limited ATP supply the overall cell function will be severely affected as most cellular metabolic processes require energy; eventually cells will die. Importantly, mitochondrial function is linked to the cellular redox state of the cell as the mitochondrial respiratory chain is critical in the $NAD(P)^+$/NAD(P)H ratio. Therefore, mitochondrial dysfunction will inherently lead to oxidative stress and overproduction of reactive oxygen species and lipid peroxidation [28].

In addition, mitochondria function as a central hub in programmed cell death. Cell death signals, such as mediated by the mitochondrial translocation of the pro-apoptotic Bcl2 family member Bax, leads to the formation and opening of mitochondrial permeability transition (MPT) which leads to the release of mitochondrial calcium and loss of mitochondrial membrane potential and therefore ATP production [29] as well as release of cytochrome c. A second mechanism related to mitochondrial control of programmed cell death is the formation of mitochondrial outer membrane pores (MOMP) that do contain BAX/BAK dimers. These pores mediate the release of mitochondrial proteins such as cytochrome c and other pro-apoptotic factors such as apoptosis inducing factor (AIF), and Smac/Diablo [30]. The release of the pro-apoptotic proteins defines the formation of the apoptosome, containing cytochrome c, APAF1 and pro-caspase-9 as well as dATP, resulting in activation of caspase-9 followed by downstream activation of caspase-3. Since ATP is required for apoptosis, severe impairment of cellular metabolism caused by chemical-induced mitochondrial damage will inhibit the onset of programmed cell death. In this more severe and uncontrolled manner cell death will occur by necrosis. Following necrosis so-called damage activation molecular patterns (DAMPs) are released from cells into the extracellular space resulting in attraction and activation of innate immune cells which in turn will release cytokines

such as TNFα. These cytokines in turn act as a signal to recruit additional immune cells. Additionally, TNFα will activate intracellular NF-κB signalling in a paracrine manner on the hepatocytes, which may also affect overall cell survival.

In contrast to mitochondrial impairment that may indirectly cause oxidative stress, direct chemically-induced oxidative stress either in the form of pro-oxidants, alkylating agents or reactive metabolites, can directly affect the antioxidant status in cells by reducing the free (non)-protein thiol status and initiating oxidative stress. The decreased cellular antioxidant status, as mentioned earlier, can impair the mitochondrial oxidative phosphorylation; reactive oxygen species (ROS) can also damage mitochondrial DNA. ROS in general is an often cited mechanism for drug-induced liver injury [31]. ROS causes DNA damage, protein oxidation and lipid peroxidation which impairs normal cellular function and cell death. Given the fact that control of oxidative stress is of paramount importance for overall cell maintenance and cell survival, the KEAP1/Nrf2-mediated antioxidant adaptive stress program has evolved as a universal mechanism across cell types and species to control ROS-levels. It is for this reason that we established several BAC-GFP reporter cell lines to quantitatively monitor the Nrf2-mediated oxidative stress response at the single cell level using automated live cell imaging (see Chapter 2 and 3 for further details).

3.3. The unfolded protein response in the endoplasmic reticulum.

In the context of molecular toxicology the endoplasmic reticulum (ER) is most well-known for the xenobiotic enzymes located in the membrane folds such as P450 enzymes, UDP-glucuronosyltransferases (UGTs) and glutathione S-transferases and for its ability to store large amounts of intracellular calcium. Therefore the ER is a source of xenobiotic transformation and resultant reactive intermediates. Moreover, upon impairment of the ER membranes large amounts of calcium are released into the cytosol resulting in activation of pro-apoptotic signalling via the previously mentioned mitochondrial perturbations [32].

The main cell physiological role of the ER is protein folding and post translational protein modification. Compounds that interfere with these processes cause the activation of the so-called Unfolded Protein Response (UPR). The UPR consists of three major branches regulated via three transmembrane transducer proteins: activating transcription factor 6 (ATF6), protein kinase R-like ER kinase (PERK) and inositol-requiring enzyme 1-alpha (IRE-1α.) [33]. The activation level of these three distinct routes determines the balance of the resultant unfolded protein response. Thus, ATF6 and IRE-1 α activation lead to a more pro-survival physiologic response such as increased chaperone production; in contrast PERK activation leads to translation inhibition and with prolonged activation of downstream target Chop (DDIT3) and can sensitize cells to pro-apoptotic signalling [34]. Further details regarding on the ER-stress/ unfolded protein response are presented in chapter 2.

3.4. Inflammatory signalling and DILI.

The involvement of the immune system in DILI has been suggested for several drugs such as diclofenac, carbamazepine and methimazole [35] and are often known as hypersensitivity reactions. Several hypothesis have been suggested such as the hapten hypothesis, genetic disposition of HLA alleles or direct binding of the (parent) compound to T-cell receptors [36]. Recent work has demonstrated the importance of drug reactive metabolite-mediated adaptive

stress response signalling and cytokine-induced pro-apoptotic signalling in DILI [37, 38]. Following liver damage immune cells infiltrate the liver and secrete pro-inflammatory cytokines which can further exacerbate inflammation and liver damage.

Tumor necrosis factor-α (TNFα) is the main pro-inflammatory cytokine excreted by liver resident macrophages known as Kupffer cells. TNFα has been shown to increase liver damage caused by various drugs [39]. TNFα activates the tumor necrosis factor receptor (TNF receptor) which in turn activates pro-apoptotic signalling via its death domain and also activates NF-κB signalling. Upon activation NF-κB transiently translocates to the nucleus to activate downstream target genes mainly involved in cytoprotective (anti-apoptotic proteins) and inflammatory (cytokines) mechanisms (Liu et al. 1996). Reactive metabolites from DILI-related drugs typically provoke a cellular oxidative stress environment thereby initiating the stabilization and activation of the transcription factor Nrf2 (Li et al. 2005). Downstream antioxidant genes contribute to adaptation and protection of cells against the reactive metabolite induced oxidative stress. Several studies indicate that Nrf2 activation can act to suppress NF-κB-based immune signalling responses (Chen et al. 2006) which would indicate Nrf2 could be involved in NF-κB dysregulation in DILI. This is also part of our own investigations described in Chapter 5.

3.5. Toxicogenomics legacy data: TG-GATES

The above adaptive stress response pathways are not the only cellular responses to toxic insults. In the past decade major progress within the toxicology field has been made to unravel a multitude of cellular responses that are initiated by xenobiotic exposure using omics technologies with the focus on transcriptomics and to lesser extent metabolomics and proteomics. More recently the attention has shifted to epigenomics and regulation by the microbiome [40] as well as the more recently developed transcriptomic technology RNA-sequencing [41]. Many gene-set based biomarkers have been reported in literature originating from omics-research [42, 43]; here optimized gene-, metabolic- or protein fingerprints/profiles were developed based on training compound sets. Several examples indicate that such profiles can add to hazard identification early in the drug development process [44]. The major benefit that emerged from these omics efforts is more likely to be the increased understanding of the biological pathways involved in the cellular, organ and organism adversity [45]. Omics-based technologies provide novel insights into the type of cellular processes initiated following chemical exposure. These insights are then followed up by more detailed mechanistic investigations. Thanks to the increased understanding of xenobiotic-induced toxicity it has become possible to create evidence-based frameworks (e.g. the AOP framework, see above) based on mechanistic understanding that can be implemented in safety testing strategies.

Throughout this thesis we used the legacy toxicogenomics dataset TG-GATES [46]. The TG-GATES dataset is available in the public domain in the form of Affymetrix .CEL files (http://toxico.nibio.go.jp/). The TG-GATES dataset originates from the Japanese Toxicogenomics Project which ran from 2004 to 2014 as a joint government-private sector project organized by the National Institute of Biomedical Innovation (NIBIO), the National Institute of Health Sciences (NIHS) and 18 pharmaceutical companies [46]. The main bulk of the data consists of Affymetrix microarray obtained transcript profiles from liver rat *in vivo* (acute and repeated dosing), primary rat hepatocytes (PRH) and primary human hepatocytes (PHH) at several time points and

concentrations. In total over 170 compounds were tested of which the majority are DILI-related drugs but also some kidney injury-related drugs and negative controls were included. In addition to the liver models a subset of the data consists of *in vivo* rat kidney repeat and acute dosing samples. The TG-GATES dataset is one of two examples of DILI-relevant and rich (encompassing large compound sets with various mechanisms of toxicity) datasets that are publicly available; DrugMatrix [47] being the second, and most large pharmaceutical companies maintain their own private rat (microarray) databases. Thus, the whole-genome transcriptomic profile of the majority of DILI related drugs is already known. Since time and dose-response relationship transcriptomic profiles are generally unavailable, due to the cost of performing microarray experiments, the TG-GATES data has paved the way to understanding general and specific responses of cells to DILI-related drugs. In addition, these transcriptomic datasets can be utilized for biomarker discovery as well as be the starting point for more functional and mechanistic understanding of DILI. In this thesis the TG-GATES dataset has extensively been used for several purposes: 1) the comparison of the HCI derived HepG2 reporter-protein levels with the transcript level in primary human hepatocytes; 2) data mining the TG-GATES data for suitable candidate BAC-reporter genes; 3) mapping of the transcriptomic profiles to more biological interpretable pathways to better understand the role of adaptive stress response pathways and inflammatory signalling to subsets of DILI-related compounds. In essence, the utilization of a previously established dataset and integration with HCI-derived data is a systems toxicology approach to efficiently combine diverse sets of models and datatypes to better understand cellular biology underlying DILI.

4. From mechanisms of toxicity to fluorescent reporters for high content imaging.

4.1. The BAC-GFP reporter platform.

To better assess the underlying mechanisms of chemical-induced liver toxicity we established a large set of reporter cell lines that quantitatively define the mode-of-action of chemicals. For this we have used the Bacterial Artificial Chromosome (BAC) cloning technology [48]. A BAC is a large plasmid containing double-strands genomic DNA and therefor includes regulatory sequences such as endogenous promoters and endogenous splicing mechanisms due to the inclusion of intron sequences. A BAC is selected based on the target gene of interest, and an additional requirement that on both sides of the outer exons of the gene of interest at least 10,000 base pairs must be situated to include regulatory elements. With homologous recombination a fluorescent tag, typically GFP, and resistance marker are cloned in the BAC plasmid, usually in frame at c-terminal side of the last intron of the gene of interest, resulting in the generation of a fusion of the gene-product of interest and GFP. The modified BAC can be introduced in cell lines. For this purpose we use the HepG2 cells, a hepatocellular carcinoma cell line with some human hepatocyte characteristics that is often used for early pre-clinical assessment of cytotoxicity liabilities. The introduction of the BAC-GFP constructs into HepG2 by transfection is followed by selection of monoclonal modified HepG2 cells which can be further selected based on suitability for imaging and homogeneity of the fluorescence level of the individual cells. HepG2 cells were selected based on several considerations. Firstly the HepG2 cell has several properties similar to that of hepatocytes. Secondly HepG2 cells are easy to culture and remain viable up to 30 passages and proliferate fast enough (doubling time of around 1 day) to allow selection after genetic modification. The application of this BAC-GFP technology in mechanistic toxicology is further

introduced in chapter 2, while throughout the thesis BAC-GFP reporter cell lines have been applied for interrogation of pathway of toxicity activation.

4.2. High content imaging of key adaptive stress networks.

As discussed briefly above the primary repair and defense mechanism of cells following chemical exposure are the adaptive stress response programs (Fig. 3). The key rationale to monitor in particular these adaptive stress response pathways follows from the paradigm that the molecular initiating events result in an altered (intracellular) biochemical environment that, as a first step, provokes activation of the various adaptive stress response pathways. Which pathways are activated depends on the mode-of-action as well as the concentration of individual compounds.

The cell senses these environmental changes via upstream biochemical sensors (e.g. Keap1 for the Nrf2-activated oxidative stress response) and responds via signalling networks called adaptive stress response pathways. These pathways are physiological pathways and as such activation does by itself not correspond to or imply toxicity. We anticipate that it is essential to monitor the magnitude of the response of all individual adaptive stress response pathways and to understand at which point a threshold, or breaking point, is reached when the cell can no longer cope with the chemically-imposed stress. At this stage cells would die by either apoptosis, necrosis, necroptosis or in milder cases where only organelles are damaged autophagy. By monitoring the concentration and time relationships of chemical-induced stress responses, it is possible to separate the initial (causative) adaptive stress type from secondary adaptive stress-types. To illustrate this, imagine the following: some compound affect proper folding of proteins in the endoplasmic reticulum resulting in the activation of the unfolded protein response (UPR), translation initiation inhibition and selective protein translation of chaperones. At a certain concentration the UPR is unable to re-establish homeostasis in the endoplasmic reticulum and proper protein folding and transport is impaired as well as release of Ca^{2+} from the lumen of the ER. Due to this impairment, secondary effects, such as oxidative stress, are expected to occur. Only by obtaining detailed maps of the concentration and time course relationships of the activation of the various adaptive stress response networks can such time resolved primary and secondary mechanistic insights be uncovered.

During the live cell imaging-based assessment of pathway activation one can simultaneously monitor the fate of the cells, the onset of necrosis or apoptosis, cell proliferation, and cell migratory behaviour. This allows the evaluation the relationship between mode and magnitude of toxicity pathway activation and cell fate caused by xenobiotic exposure. These relationships inform us on the mechanism of toxicity. Such combinations of biological readouts can be interpreted as a dynamic biological ‘fingerprint’. Importantly, such fingerprints allow unsupervised clustering methods to define sets of compounds with similar mode of action and cell fate. Such strategies can aid in prediction of toxic liabilities.

5. High content imaging.

High content imaging (HCI) has evolved greatly in the past two decades due to the technological advancement of microscopy technology. Automated microscopes can be fitted with incubation systems enabling live cell imaging. Robotic arms for multi-well imaging plate loading in combination with robotic liquid handlers lead to high throughput imaging-based experimentation.

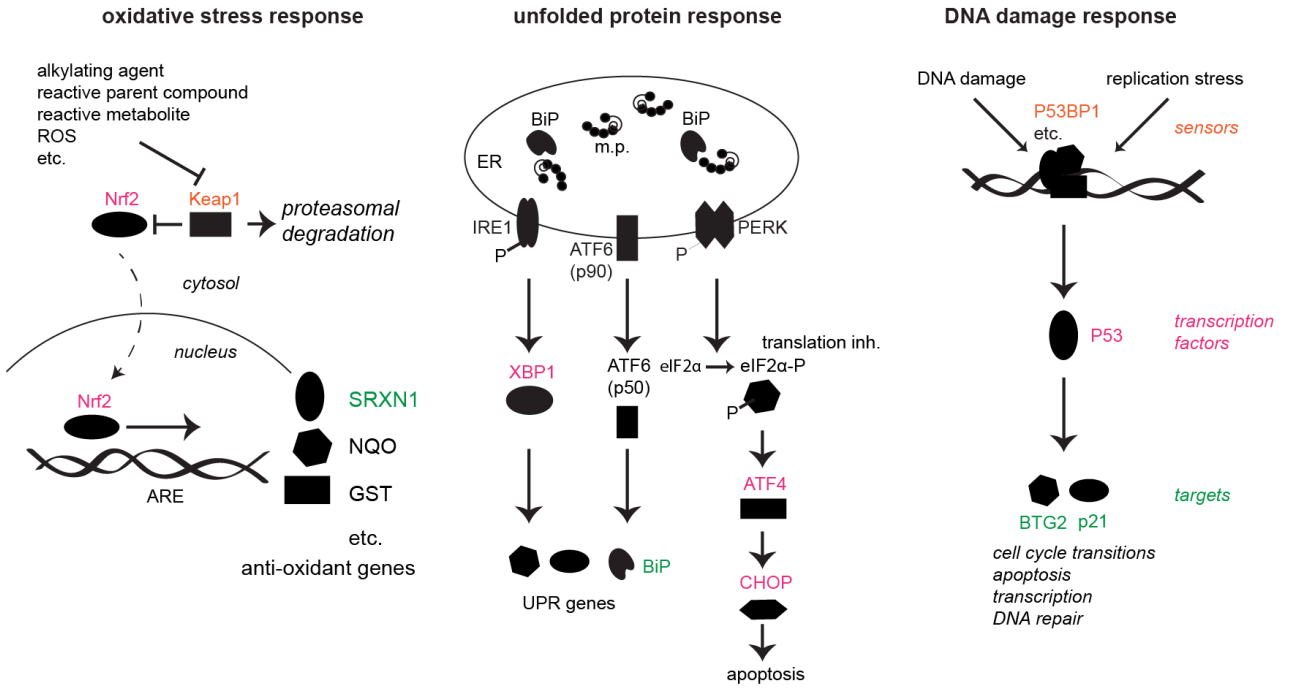


Figure 3: **Primary adaptive stress response programs.** Left) The oxidative stress response program which is activated by the translocation of Nrf2 to the nucleus where it binds to the antioxidant response element (ARE). Middle) The three arms of the unfolded protein response activated by misfolding (m.p.) or overload of proteins in the ER. Right) The DNA damage response which involves a sensory and repair machinery and main regulator p53. Each adaptive stress response pathway consists of sensor proteins, transcription factors and down stream targets.

Confocal microscopes allow a high level of detail of in practice up to four different fluorescence emission wavelengths. In addition, most confocal microscopes allow simultaneous image acquisition of transmitted light, phase contrast or other non-fluorescence based image modalities. Because of the high throughput applications and enormous amount of imaging data, proper and efficient file storage systems are essential. Moreover, all images contain massive amount of information regarding the level and localization of fluorescence at the single cell level. The analysis of these data in an efficient manner requires specific informatics-based automated strategies. The development of an image analysis pipeline to enable current efforts in high throughput HCI is described in chapter 2 and chapter 7.

6. Tool development for datamining in HCI.

As indicated, HCI efforts results in enormous amounts of quantitative data. Due to technological development and automation of HCI hardware data output can be staggering which requires software in the form of data base management systems (DBMS) and application interfaces and computational hardware in the form of a storage file system and sufficient computational power. To enable datamining of large datasets one must consider two important physical aspects of a "dataset"; the size and the computational load the desired operations on the data require. A database engine based on structure query language (SQL) such as MySQL allows unlimited storage of data in relational tables. The organization of HCI data in a relational database is quite intuitive as each object will have a primary index as well as secondary indexes linking the parent-child relationships which allows the join-operations to be performed on multiple table to, for example,

query all measurements for a certain cell containing multiple objects such as the nucleus, organelles, cytoplasm over several frames in time. CellProfiler utility allows the output data derived from the image analysis to be stored in such databases [49]. This is often the recommended approach for handling large and complex datasets resulting from image analysis pipelines of single cell data and screens. Yet, this also requires informatics expertise to maintain a server-side database, or keep track of lots of small local databases. CellProfiler also allows run-time data dumps to the HDF5 file storage system [50]. Several applications (e.g. HDFView) allow easy browsing through this file system making it more intuitive. Memory wise it is possible to load subsets of data from the HDF5 file to analyze determined by the amount of memory available on the local system. After the user selected data is loaded into R-memory, speed is a non-issue as no more read/writing operations to hard drives have to be performed. By using a binary- ordered indexed table structure provided by the R-package 'data.table' all further data analysis operations are near-instant. Calculations on millions of data entries involving very small subsets of large datasets can still require a lot of time on a single CPU. R provides parallel computation with for example 'doSNOW' and 'parallel' packages without data duplication resulting in trivial additional memory usage. As many modern desktops prove 8-24 cpu-threads, this speeds up computations considerably.

The work presented in this thesis involves HCI data representing high throughput experiments for various reporters exposed to >170 compounds at the different concentrations with up to 50 time points per condition. Each condition would capture information of 100-200 individual cells. This type of data is no longer suited for detailed analysis using spreadsheet software tools as a single overnight experiment can easily lead to several GB of numeric data. The memory of spreadsheet tools is insufficient and the computational load too severe for even a single over-night HCI dataset. Often such data types will be analyzed by computational/informatics-oriented staff which will perform the analysis based on the requests of the biologist. The output from the analysis is then interpreted by the biologist who will then usually return to the (bio)-informatician who will optimize the analysis. Meanwhile more data will be produced. This leads to a cycle which is very time consuming and leads to communication problems as biologists and informaticians have to understand the jargon and science from two different worlds.

For the above reason a user-friendly R-package was developed (H5CellProfiler) that allows biologists to analyze their own large datasets such as perform aggregations, statistics, plate-normalizations and basic table manipulations such as selection, filtering, division and counting. In addition, live-cell tracking specific operations can be performed on the single cell data such as making the data suitable for single cell tracking by creating unique track labels for each time-lapse movie and also reconnecting "broke" tracks based on user defined thresholds. After appropriate data summarization H5CellProfiler allows interactive plotting of selected treatments to create time course or dose response graphs or the user can plot all data in a single graphic using the faceting utility provided by the R-package 'ggplot2'. The described utilities are provided in the form of a browser based menu; a graphical user interface which sends commands to the (local) R-server. A detailed description of H5CellProfiler architecture is provided in chapter 7.

7. Overall aim and objectives of this thesis.

Our long term vision is to establish an imaging-based platform that can quantitatively assess the activation of individual key events relevant to AOPs. The initial focus in this thesis is on adaptive stress response pathways, that are typically part of AOPs and related to adverse drug reactions. In this thesis, we have established and characterized various reporter cell lines. We have established the infrastructure for automated imaging and image analysis of these reporters cell lines. We have integrated these reporter cell lines in mechanistic understanding of DILI. We have used these reporter cell lines to improve the prediction of DILI. These topics are described in the following chapters:

Chapter 2 provides an overview of the background of adaptive stress responses, the development of reporter cell lines to monitor these pathways, and examples of the application of these reporters.

Chapter 3 details the generation and characterization of pathway reporters that reflect the KEAP1/Nrf2 pathways, the unfolded protein response and the DNA damage response.

Chapter 4 describes the interaction between the Nrf2 pathway and the NF-κB pathway in relation to DILI. I present how DILI compounds with strong Nrf2 activation affects the NF-κB signalling pathway and susceptibility to DILI.

Chapter 5 deals with the role of oxidative stress and the unfolded protein response in the control of DILI compound-mediated cytotoxicity. I present how the activation of these pathways by DILI compounds affect the susceptibility to TNFα-induced cell death. Here we applied different BAC-GFP reporters to monitor the UPR.

Chapter 6 involves the application of adaptive stress response reporters in the prediction of DILI. Here we used >170 different compounds, that largely cover most DILI types, and demonstrate the applicability of the reporter systems in DILI prediction.

Chapter 7 describes the HDF5CellProfiler pipeline for the datamining of high content imaging data.

Chapter 8 provides an overall summary of the work and discussion on future perspectives and further challenges on the application of the reporter cell lines in DILI safety assessment.