High throughput microscopy of mechanism-based reporters in druginduced liver injury

Hiemstra, S.W.

Chapter 7

Summary and discussion

GFP stress response reporter system for DILI prediction

The establishment and application of a HepG2 GFP reporter cell line platform is central in this thesis. The past decade there has been a strong incentive to test adaptive stress response activation in toxicity screening, initiated by the report of the National Research Council in 2007 ('*Toxicity testing in the 21st century: a vision and strategy*'). We applied this strategy to better predict drugs with high iDILI hazard pre-clinically. Therefore, we build a large panel of GFP reporters cell lines to be able to follow stress response activation over time using confocal microscopy. In **chapter 2** we generated and characterized eleven GFP reporter cell lines. For three stress response pathways we chose an upstream sensor, the key transcription factor and a specific target of the key transcription factor. This approach resulted in Keap1-GFP (sensor), Nrf2-GFP (transcription factor) and Srxn1-GFP (target) for the anti-oxidant response pathway, 53bp1-GFP (sensor), p53-GFP (transcription factor) and p21-GFP (target) for the DNA damage repair pathway and Atf4-GFP and Xbp1-GFP (transcription factors) and BiP-GFP and Chop-GFP (targets) for the unfolded protein response. When we induced these reporter lines with known compounds activating the pathways we were able to observe perfect clustering by unsupervised hierarchical clustering. In addition, when we selected compounds based on transcriptomic analysis in primary human hepatocytes, we were able to find 100% sensitivity for Srxn1-GFP, 80% sensitivity for p21-GFP and 50% sensitivity for Chop-GFP and BiP-GFP. All together, these results ensured a stable foundation to test reporter activity after stimulation with a large batch of iDILI drugs. In addition, these results enable generation of additional reporter lines, also concerning other stress response pathways e.g. the heat shock response and hypoxia[223,224].

Therefore, in **chapter 3** we tested 176 drugs including drugs known to induce DILI in the clinic (most-DILI-concern drugs) and drugs known to be save for DILI in the clinic (no-DILI-concern drugs). These drugs were tested at 1, 5, 10, 50 and 100 Cmax for Srxn1-GFP, Chop-GFP and p21-GFP. We followed GFP activation for 24 hours after induction, generating time dynamic profiles of three GFP reporter lines for each compound concentration combination. GFP intensity was determined on single cell level. This allowed us to summarize not only the averaged GFP intensity of all cells per conditions, but also the fraction of cells exceeding the GFP background levels. When all data was summarized with unsupervised hierarchical clustering an overrepresentation of most-DILI-concern drugs is present in the activated clusters. This implies that the hypothesis, that adaptive stress response activation can be indicative for liver injury hazard in the clinic, shows good promise.

To be able to use this data in preclinical safety screening the prediction rate needs to be increased. This can be achieved by testing more stress response pathways, for example the innate immune signaling (NFκB) or the heat shock response. Also, more time points can be added, for example 48 and 72 hours. Furthermore, thus far, we only used fraction positive cells, GFP intensity and hierarchical clustering to rank the responses and discriminate between drugs.

However, this data can also be used to build a predictive classifier. Kraus and Frey reviewed the possibilities to classify fluorescent microscopy data using machine learning algorithms [225]. The classifier can then be used to predict whether preclinical stress response activation would lead to high DILI hazard in the clinic. Another approach is to determine the tipping point or point of no return[181]. By identifying the exact concentration where the system is not able to adapt to the stress anymore a more elaborate understanding of the switch between adaptation and cell death can be achieved. This knowledge can also be fed into the classifier algorithm, leading to enhanced predictive power. All together, the predictive rate of such a classifier will increase when we add more endpoints. That is why also other measurements could be of value. For example the size and shape of the lysosomes or total mass of mitochondria could play a role. In the end, when the classifier is optimized, the key features can be chosen to screen for in preclinical screening to reduce costs and workload. Taken together, chapter 2 and 3 form the basis for determining the usability of HepG2 GFP adaptive stress response reporters in preclinical pharmaceutical screening.

Application and advantages of 3D cultured HepG2 GFP reporters

A major gap in the GFP reporter system is the lack of metabolizing capacity of HepG2 cells. This would mean we will not be able to pick up drugs of which a toxic metabolite will cause injury in the liver. An alternative would be to use a more advanced liver model as primary hepatocytes, however this would encompass a restriction to cell death and ATP end points. Previously, we showed differentiation of HepG2 wild type cells when grown in 3D hydrogel[27]. In **Chapter 4** we combined the GFP reporter system with 3D spheroid culturing. To this end, differentiation profiles of six GFP reporters (Srxn1-GFP, Nqo1-GFP, BiP-GFP, Chop-GFP, p21-GFP and Btg2-GFP) were validated. Next, a screening method was set up to automatically screen reporter activation in 3D spheroids using confocal microscopy. First, a screen was performed with drugs and compounds activating the different reporter lines based on primary human transcriptomic data. This led to an overall sensitivity of 70%, with Srxn1 even at 100%. Then, we applied 33 drugs with known DILI hazard in six concentrations (1, 5, 10, 25, 50 and 100x Cmax). Hierarchical clustering revealed overrepresentation of drugs with most-DILI-concern in activated clusters compared to drugs with no-DILI-concern. Since we selected most-DILI-concern drugs which were not activated in 2D screening and were shown in literature to be metabolized in the liver, we also observed activated stress responses not present in 2D monolayer cultures. Furthermore, 3D spheroid culturing allows repeated exposures, which encompasses drug application each day for six days. This is a better representation of the human situation, where each day or multiple times a day a drug needs to be administrated. Six days repeated dose exposures demonstrated more significantly upregulated adaptive stress responses indicating this way of measuring delivers valuable information. Therefore, the strength of 3D spheroid culturing is not only the expression of enhanced liver like properties,

but also the possibility to screen drugs in a way that mimics the human situation better. Altogether, chapter 4 emphasizes the possibility to use the GFP reporter system grown in 3D spheroids allowing metabolism of drugs and repeated dose exposures.

As mentioned before, primary human hepatocytes are not suitable for the generation of stable fluorescent lines as they cannot be passaged. An alternative would be the use of pluripotent stem cells. These cells can be passaged before differentiation process is started, allowing genome editing. Recent advancements in the field of genome editing allows the use of CRISPR/Cas9 systems to specifically insert parts of DNA in the host genome[226]. If it would be possible to incorporate GFP DNA downstream of the gene of interest an endogenous fusion GFP protein will be formed. The pluripotent stem cells can then be differentiated to hepatocytes and be used for safety screening. As pluripotent stem cells do not have a restriction in passages, it would also be possible to generate one master cell line containing multiple fluorescent markers for different stress response pathways. In this way, one cell line would be sufficient to look at multiple stress responses, which would decrease costs dramatically. The deficit of this system is the costs it would take to differentiate pluripotent stem cells into hepatocytes. Also, pluripotent stem cells are not able to equal the advantage of 3D HepG2 cells to perform repeated exposure experiments. Taken together, genome editing in pluripotent stem cells show great promise in the application for future toxicity screening, but, as every system, also has its downsides.

Application of GFP reporters in unraveling mechanisms of DILI

As discussed in the introduction, DILI is a multifaceted problem with an drug intrinsic and a genetic factor involved. To be able to understand how these drugs can affect the liver, full understanding of stress response activation upon drug exposure is pivotal. Furthermore, mechanisms of crosstalk are key as simultaneous activation of multiple stress response pathways increases the change at cell death. The proposed crosstalk between the Nrf2 and p53 signaling pathways are contradictory and not wel studied in literature. Therefore, in **chapter 5** we set out to test whether model compounds activating a pathway were able to induce crosstalk mechanism. This resulted in Nrf2 activation upon stimulation with high concentrations of etoposide and p53 activation with high concentrations of diethyl maleate (DEM). To study the mechanism of crosstalk in our GFP reporter system, we conducted RNAi experiments. We identified induction of the p53 pathway after knock down of Keap1. This supported the hypothesis that Nqo1 could induce p53 signaling, however, more research is needed to define if this is the case in the HepG2 GFP system, but this is beyond the scope of this chapter.

The Nrf2 pathway is activated upon stimulation of a large battery of DILI compounds. To elucidate all components important in Nrf2 activation we conducted a large RNAi based screen to identify novel regulators of the Nrf2 pathway (**chapter 6**). Primary and deconvolution screening led to the identification of 58 suppressors and 19 inducers of the Nrf2 pathway.

Exposures with DEM, acetaminophen, azathioprine and diclofenac showed the identified hits were generic hits, which could be induced by several compounds, also DILI relevant ones. In addition, the genetic component was shown to be key in activating the stress response as most compounds induce the same pattern in Srxn1 activation. Knock down of the suppressors led to better protection against DEM-mediated cell death and knock down of the inducers led to worse protection against DEM mediated cell death. This implicated that these hits are important factors in the switch between adaptation and cell death. In addition, we evaluated other GFP reporters: Nrf2-GFP, Keap1-GFP, Nqo1-GFP and Hmox1-GFP to study the mechanisms of Nrf2 activation in more detail. Still, more in depth studies to specific hits can be performed. Firstly, we might require conditional shRNA approaches for validation of our siRNA based knock down studies. Recent, progress in CRISPR/Cas9 mediated knock out cell lines could provide stable knock out cell lines[227]. This could include the effect of specific knock downs on gene expression, ChIPseq of the identified co-transcriptional regulators of Nrf2, immunoprecipitation of candidate genes followed by proteomic analysis to unravel specific signaling complexes that modulate Nrf2 activity. Finally, generation of BAC-GFP reporters of candidate genes could unravel the understanding of the dynamics of these regulators during oxidative stress conditions. Susceptible individuals could have impaired Nrf2 signaling making them vulnerable to drugs inducing oxidative injury. Investigations in large human whole genome sequencing datasets may establish the full landscape of SNPs in our candidate Nrf2 regulators. Understanding the function of these SNPs in Nrf2 regulation could then at a next level contribute to the application in personalized medicine, where particular SNPs could then also predispose to DILI.

In conclusion, we demonstrated the GFP reporter system can also be used to identify mechanisms of adaptive stress response activation. By combining RNAi based screening with the GFP reporters we are able to establish a network of regulators important in Nrf2 signaling. This approach can now also be pursuit for ER stress activation and p53 signaling, all together leading to a panel of key factors in cell death and DILI onset. This is pivotal knowledge in the battle against DILI.

Future perspectives and overall conclusions

The adaptive stress reporter platform is a powerful tool for high throughput toxicity screening. By obtaining important information about stress response activation early on in preclinical testing, we could provide danger signals early on in the process of lead identification and lead optimization. At this moment, the BAC-GFP reporter platform is ready to be implemented by pharmaceutical companies in drug safety testing regimes. To increase such a take up, the development of a reliable and robust classifier is a first next step. In addition, we would need to identify other critical markers that may help to identify other drugs with potential DILI liabilities, these markers could then also be introduced in HepG2 cells to generate additional

BAC-GFP reporters. The current library of GFP reporters were all chosen based on their function in stress response pathway activation. Another approach would be to select genes based on transcriptomic data after exposure with DILI drugs. In this way, selected genes would function as biomarkers for DILI. To identify these biomarker GFP reporters different approaches can be used. Large transcriptomic datasets including the TG-GATES dataset can be used. Also, the recent approach of Sutherland *et al*. can be used[228]. This group performed weighted gene co-expression network analysis (WGCNA) on known transcriptomics datasets like TG-GATES. Here they defined networks of co-regulated genes based on a large set of compounds. When these networks are known, important genes in the most prominently activated networks by DILI drugs can be selected to generate GFP reporters for. By selecting novel candidate genes in this way, the predictivity of a GFP reporter-based classifier will likely increase.

One can argue whether it is essential to validate stress response activation *in vivo*. In principal, when stress response activation would provide a classifier which is indicative for risk at DILI, do we need to know whether stress response activation is also occurring in another model system? Recent advancements in intravital imaging could aid in validating stress response activation *in vivo*[229]. However, ideally, stress response activation should be validated in humans during clinical trials. This would strengthen the hypothesis of stress response activation as a marker for increased risk at DILI. Studying extensive time and dose responses in humans will be impossible, however novel developments in the field of liver exosomes could provide more insight in stress response activation in humans. It has been shown in rat with liver injury that HSP70 and HSP90 protein levels are increased compared to the healthy control rats[230]. This indicates liver exosomes could potentially be used to validate stress response activation in human after perturbation with various drugs inducing DILI.

Another important issue is the question how different time and dose responses *in vitro* are translatable to what happens in humans. In humans, route of exposure including everyday circumstances, mainly determines how much and when the drug will end up in the hepatocytes. Therefore, the intracellular drug concentration and time dynamics will differ between each individual and also each time an individual takes a drug. This does not mean time dynamics and dose responses are not useful to measure *in vitro*. Time dynamics can give information on how strongly the system needs to respond to avoid cytotoxicity. For example, for p53 signaling it has been shown that p53 onset in cancer cells is the critical factor in cisplatin induced apoptosis[231]. Therefore, time dynamics can be of pivotal importance when building a predictive classifier.

As a concluding thought we can ask ourselves the question: what if we would generate a classifier predicting 100% sensitivity and specificity for all DILI drugs screened? This would not mean DILI will never occur again. As DILI drugs cause DILI in several (unknown) ways it will always be possible a newly designed drug will slip through the barriers and be indicated as a false positive. The reason would then be the new mechanism of action is not incorporated in

the GFP reporter set. How valuable this GFP reporter system is, can only be determined after years of preclinical testing. However, this thesis ensures a stable foundation to already start collaborations with pharmaceutical companies to test the GFP reporter system in their preclinical screening.

Concluding remark

Altogether, this thesis contributes on one hand to a better preclinical prediction of DILI and on the other hand to a more elaborate understanding of the mechanisms underlying DILI. This was achieved by the development of an adaptive stress response GFP reporter system functioning as a two-edged-sword. By quantitative high throughput assessment of GFP tagged adaptive stress response reporters we brought the solution to the multifaceted DILI problem one step closer.