Time- and Concentration-dependent response of a liver cell line to Benzo(a)pyrene

Author(s):
Jannuzzi Madureira, Danielle

Publication Date:
2012

Permanent Link:
https://doi.org/10.3929/ethz-a-007620572

Rights / License:
In Copyright - Non-Commercial Use Permitted
Time- and Concentration-dependent response of a liver cell line to Benzo(a)pyrene

Danielle Jannuzzi Madureira

2012
Time- and Concentration-dependent response of a liver cell line to Benzo(a)pyrene exposure

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Sciences

Presented by

Danielle Jannuzzi Madureira

Master in Genetics and Molecular Biology, State University of Campinas
born on November 12, 1981
Citizen of Brazil / Portugal

Accepted on the recommendation of
Prof. Dr. Kristin Schirmer, examiner
Prof. Dr. Martin Ackermann, co-examiner
Dr. Juliette Legler, co-examiner
Prof. Dr. Rik I. Eggen, co-examiner

2012
# Table of Content

**Abstract** ................................................................................................................................... V

**Zusammenfassung** .................................................................................................................. VIII

1. General Introduction .................................................................................................................. 1

   1.1 Organic Chemicals in the Environment ........................................................................ 1

       1.1.1 Polycyclic Aromatic Hydrocarbons ................................................................... 2

       1.1.2 Benzo(a)pyrene (BaP) ...................................................................................... 3

   1.2 The Aryl Hydrocarbon Receptor Signal Transduction Pathway .................................. 5

       1.2.1 The Aryl Hydrocarbon receptor (AhR) as transcription factor ..................... 5

       1.2.2 BaP as a ligand of AhR ...................................................................................... 8

   1.3 Toxicogenomics .............................................................................................................. 11

       1.3.1 Transcriptomics ................................................................................................ 12

       1.3.2 Phenotypic anchoring (Adverse outcome pathway) ....................................... 13

       1.3.3 Cellular models ................................................................................................. 14

   1.4 Scope of the Thesis ......................................................................................................... 15

   1.5 References ..................................................................................................................... 18

2. Differential time-resolved molecular response of a mouse liver cell line on low and high benzo(a)pyrene exposure .......................................................................................... 23

   **ABSTRACT** ....................................................................................................................... 24

   2.1 Introduction .................................................................................................................... 25

   2.2 Material and Methods ................................................................................................. 28

   2.3 Results .......................................................................................................................... 32

   2.4 Discussion ..................................................................................................................... 51

   2.5 References ..................................................................................................................... 56

   2.6 Support Information ..................................................................................................... 63

3. Characterization of tiPARP in BaP-exposed Hepa1c1c7 wild-type and tiPARP silenced cells ....................................................................................................................... 71

   **ABSTRACT** ....................................................................................................................... 72

   3.1 Introduction .................................................................................................................... 73

   3.2 Material and Methods ................................................................................................. 78

   3.3 Results .......................................................................................................................... 83
Benzo(a)pyrene (BaP) is a ubiquitous pollutant derived by incomplete combustion of organic matter. BaP is classified by the International Agency of Cancer (IARC) as carcinogenic to humans and is controlled in drinking water in the United States and Europe. BaP needs to be metabolized to become carcinogenic. Its metabolism is induced by the interaction between BaP and the aryl hydrocarbon receptor (AhR) and subsequent gene regulation. Despite this general knowledge, details about BaP entry into cells, distribution and internal concentrations able to elicit a cellular response are largely unexplored and mechanisms of actions incompletely understood. This PhD thesis therefore aims to investigate in a time and concentration dependent manner the molecular response of a murine hepatoma cell line, Hepa1c1c7, exposed to BaP, combining high throughput techniques, bioinformatics, phenotype characterization and chemical analysis.

To elucidate the temporal response of the murine liver cell line, Hepa1c1c7, a transcriptome data set was first established. Despite the 100-fold difference in BaP exposure concentration, i.e. 50 nM of BaP, where no cytotoxic effect was observed, and 5 µM of BaP, being a common concentration used in toxicology studies, the pattern of transcript regulation was shown to be comparable in cells for up to 4 h of exposure. A total of 20 genes was regulated by the low BaP concentration with a large fraction of them also being regulated by the high BaP exposure condition. However, while the transcriptome returned to control levels for the low concentration of BaP between 4 and 12 h of exposure, the transcriptome response diverged for the high concentration of BaP, with more than 1000 genes identified as significantly regulated. Selected descriptors of (sub)cellular phenotype were related to the regulation of selected groups of genes. Cells exposed to the low (50 nM) BaP showed a slight reduction in cell doubling time but no impact on cell viability or signs of formation of ROS. In contrast, cells exposed to the high (5 µM) BaP concentration greatly reduced proliferation and were prone to cell death. Both cyp1a1 mRNA transcripts and Cyp1a catalytic activity were induced by BaP, implying BaP metabolism and production of toxic and/or carcinogenic metabolites. Therefore, the slight decrease in proliferation observed for the low concentration of BaP is an indication that cells allow for damage repair in order to return to normal cell cycling. However, on exposure to the high BaP concentration, this defence is overruled and cells undergo cell death. Based on the regulation of the transcriptome, the 4 h time point is highlighted as an important window in the cells’ fate.
Abstract

Based on the transcriptome analysis an early regulated gene for both BaP concentrations used, namely TCDD-inducible poly(ADP-ribose) polymerase (tiPARP), was chosen to be further investigated. The tiPARP gene was silenced in Hepa1c1c7 cells and different end points were tested in wild type and tiPARP silenced cells, exposed and non-exposed to BaP. The hypothesis underlying this study was that tiPARP’s main function was to protect cells against damage induced by dioxin-like compounds, such as BaP. However, no change in cell viability upon exposure to BaP was seen when tiPARP was silenced by siRNA. One potential explanation was that tiPARP protein was not down-regulated sufficiently. However, attempts to detect the endogenous tiPARP protein levels using western blot and MudPIT were unsuccessful.

Transcriptome analysis indicated rapid uptake of BaP into cells along with transcription of genes related to BaP metabolism, cell cycle arrest and/or repair. At the same time, no impact on cell viability occurred and cell doubling time was only slightly increased for the low BaP concentration. In contrast, severe impact on cell viability and an almost complete block of cell proliferation was noted at the high BaP concentration. In order to investigate if these differences could be explained by cell-internal BaP amounts, BaP cellular concentrations were quantified using liquid scintillation counting (LSC) and high performance liquid chromatography (HPLC). BaP was indeed taken up by cells fast (within the first 4 h), although most of the BaP stayed in the culture medium, apparently associated to the culture medium components. The number of BaP molecules within one cell reached a maximum after 2 and 4 h and ranged between 1.2 to 1.9 $10^6$ and 1.2 to 1.7 $10^9$ for 50 nM and 5 µM, respectively. Biotransformation became apparent between 4 and 8 h of BaP exposure and most of the added BaP (> 90 %) was bio-transformed within 24 h for both concentrations. Rates of BaP uptake and biotransformation were calculated, showing that uptake is faster than biotransformation (0.17 and 0.15 h⁻¹ for 50 nM and 0.12 and 0.1 h⁻¹ for 5 µM BaP, respectively) for both concentrations. DNA adduct quantification showed the presence of adducts as early as 8 h of BaP exposure. Both DNA adduct isomers, t(+)-Benzo(a)pyrene diolepoxide (BPDE) dGuo and t(-)-BPDEdGuo, were detected in Hepa1c1c7 cells exposed to 50 nM and 5 µM BaP, with the first being more abundant. DNA repair was seen at the 24 h time point in the cells exposed to 50 nM BaP while repair mechanisms seemed to be overwhelmed at the 5 µM BaP exposure and DNA adducts continued to accumulate. These results are in agreement with the transcriptome analysis, where genes related to nucleotide excision repair (NER) were regulated at the high (5 µM) BaP concentration and later (12 and
24 h) time points. NER pathway is responsible to remove bulky DNA adducts as caused by BPDE.

In conclusion, this work demonstrates that studying the cell response at different levels in a time and concentration dependent manner helps to better understand the mechanisms allowing cells to survive or having cells succumb to chemical exposure. Such knowledge and the rate constants calculated in this thesis can contribute to building computational models with the goal to predict cellular responses to BaP in the Hepa1c1c7 cells or other cells, as well as to other stressors with similar mechanisms of toxic action.
Zusammenfassung


Um die zeitliche Antwort der Leberzellen auf BaP zu eruieren, wurde zunächst ein Transkriptomdatensatz etabliert. Es zeigte sich, dass das Muster der Transkriptregulation nach 2 und 4 Stunden BaP-Exposition für beide eingesetzten Konzentrationen, d. h. 50 nM (kein zytotoxischer Effekt) und 5 µM (häufig eingesetzte Konzentration in vorhergehenden Studien), vergleichbar waren. Bis zu 20 Gene wurden durch die niedrige BaP-Konzentration reguliert, wobei eine Reihe dieser Gene auch durch die hohe BaP-Konzentration reguliert wurde. Während jedoch das Transkriptom für die niedrige BaP-Konzentration zwischen 4 und 12 Stunden Exposition auf Kontrollniveau zurückreguliert wurde, löste die hohe BaP-Konzentration über die Zeit mit 1165 signifikant regulierten Genen eine umfangreiche Antwort des Transkriptoms aus. Ausgewählte Deskriptoren des (sub)zellulären Phänotyps wurden mit der Regulation von Gruppen von Genen in Verbindung gebracht. Zellen, die gegenüber der niedrigen (50 nM) BaP-Konzentration exponiert wurden zeigten einen leichten Anstieg der Verdopplungszeit aber keinen Einfluss auf die Zellvitalität oder die Bildung reaktiver Sauerstoffspezies. Dagegen war die Zellproliferation in Zellen, die gegenüber der hohen (5 µM) BaP-Konzentration exponiert wurden, sehr stark reduziert und die Zellen waren anfällig für Zelltod. Sowohl cyp1a1-mRNA als auch katalytische Aktivität von Cyp1a wurden durch BaP induziert, was anzeigt, dass BaP metabolisiert wurde und dabei toxische und/oder kanzerogene Metabolite gebildet wurden. Der leichte Anstieg der Verdopplungszeit, welche für die niedrige BaP Konzentration beobachtet wurde, ist somit ein Anzeichen dafür, dass die
Zusammenfassung

Zellen DNA-Schäden reparieren um danach wieder in den normalen Zellzyklus einzutreten. Bei Exposition gegenüber der hohen BaP-Konzentration reichen dagegen die Reparaturmechanismen nicht aus und die Zellen sterben ab. Aufgrund der Regulation des Transkriptoms wurde deutlich, dass 4 Stunden Exposition gegenüber dem Stressor ein sehr wichtiger Zeitpunkt für das Schicksal der Zellen ist.

Eines der Gene, welches bei den Transkriptomstudien bereits nach 2 Stunden durch beide eingesetzten BaP-Konzentrationen reguliert wurde, ist die TCDD-inducible poly(ADP-ribose)polymerase (tiPARP). Um dessen Funktion näher zu charakterisieren, wurden Hepa1c1c7 Zellen mit tiPARP silencing (si)RNA transfiziert und verschiedene Effekte auf die Zellvitalität in Wildtyp- und behandelten Zellen in der An- oder Abwesenheit von BaP untersucht. Die Hypothese war, dass die primäre Funktion von tiPARP der Schutz gegen Dioxin-ähnliche Verbindungen, wie BaP, ist. Allerdings konnten keine Unterschiede in der Zellvitalität festgestellt werden wenn tiparp-mRNA durch die siRNA stark reduziert vorlag. Eine mögliche Erklärung dafür war, dass das tiPARP Protein nur ungenügend reduziert werden konnte. Um dieser Erklärung nachzugehen wurde versucht, das Protein in den Zellen mittels Western Blot-Analyse und MudPIT (Multidimensional Protein Identification Technology) zu detektieren, was allerdings erfolglos blieb.


Abschliessend zeigt diese Studie, dass die Untersuchung der Zellantwort auf verschiedenen Ebenen in Abhängigkeit von Zeit und Konzentration sehr hilfreich für ein besseres Verständnis der Mechanismen ist, die einer Zelle erlauben, die Exposition gegenüber BaP unschädlich zu machen oder ihr zu unterliegen Dieses Wissen kann auch dazu beitragen,
Zusammenfassung

Computermodelle zu erstellen, die benutzt werden können, um zelluläre Antworten von Hepa1c1c7 oder anderen Zellen auf eine BaP-Exposition, oder auf andere Stressoren, die ähnliche molekulare Wirkmechanismen haben, vorherzusagen.
1. General Introduction

1.1 Organic chemicals in the environment

Presence of pollutants in the environment has always been part of human civilization (SPENGLER and SEXTON, 1983; HONG et al., 1994; HONG, 1996). Air pollution was first officially acknowledged by King Edward I of England through banning the burning of sea-coal in 1272 after the smoke became a problem in London (http://www.epa.gov/aboutepa/history/topics/perspect/london.html). However, it was the industrialization in the 18th century that started what we know today as environmental pollution. The development of factories and consumption of large quantities of coal and other fossil fuels gave rise to record air pollution and the large volume of industrial chemical discharges added to the growing load of untreated human waste.

“The Clean Air Act” of 1956 was the first major modern environmental legislation after “The Great Smog” of 1952 that killed approximately 4000 persons in London. In the middle of the 1950s and early 1970s the U.S Congress also approved several legislations (“Noise Control Act”, “The Clean Air Act”, “The Clean Water Act” and “The National Environmental Policy Act”), however, those did not stop severe incidents to happen, like the dumping of polychlorinated biphenyl (PCB) in Hudson River in 1974. International catastrophes also took place, such as the wreck of the Amoco Cadiz oil tanker off the coast of Brittany in 1978 and the Bhopal disaster in 1984 in India, demonstrating the universality of such events and the scale on which efforts to address them needed to engage (CONAN, 1982; BROUGHTON, 2005) (http://www.epa.gov/aboutepa/history/topics/perspect/london.html).

In 2001 an international environmental treaty, the Stockholm Convention on Persistent Organic Pollutants, was signed and was effective from May 2004, aiming to eliminate or restrict the production and use of persistent organic pollutants (POPs). POPs are organic chemical substances that possess a particular combination of physical and chemical properties such that, once released into the environment, they remain intact for exceptionally long periods of time, become widely distributed throughout the environment, accumulate in the fatty tissue of living organisms, thereby increasing the concentrations at higher levels in the food chain, and are toxic to both humans and wildlife (http://chm.pops.int/Convention/ThePOPs/tabid/673/Default.aspx). Efforts to ban POPs have thus far focused largely on chlorinated organic chemicals, i.e. certain pesticides, industrial chemicals as well as by-products of industrial processes. However, obligations to reduce
emissions of other POPs, such as polycyclic aromatic hydrocarbons (PAHs), are agreed upon (http://chm.pops.int/Convention/tabid/54/language/en-US/Default.aspx).

The first record of linking exposure to organic chemicals in the environment and human disease dates back to 1775 and the English surgeon Sir Percival Pott. He suggested that chimney sweeper’s cancer (scoot wart), a squamous cell carcinoma of the skin of the scrotum observed with high incidence in chimney sweeps at the time, originated from components of soot, marking the first occupational link to cancer (WALDRON, 1983). Yamagiwa and Ichikawa in 1914 were the first to report induction of squamous cell papilloma on rabbit ears after repeated application of coal tar. Kennaway in 1930 found that tumours in mouse skin could be produced by pure 1,2,5,6 dibenzanthracene (KENNAWAY and HIEGER, 1930). Later in 1932 Cook and collaborators isolated a highly carcinogenic crystalline compound from coal tar and showed that this chemical was the PAH benzo(a)pyrene (BaP) (COOK et al., 1932; WALDRON, 1983).

1.1.1 Polycyclic Aromatic Hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) are a large group of environmental organic pollutants that are of particular concern for humans and other vertebrates, as many of them are known or suspected carcinogens. PAHs are ubiquitous environmental contaminants derived from incomplete combustion of carbon and they belong to a family of chemicals that can act like dioxins, based on a similar compound structure and a common mechanism of toxic action (BOLS et al., 1999; DONG, 2008). Dioxin-like compounds include chemicals that are persistent, bio-accumulative and have biologic responses mediated via binding to the aryl hydrocarbon receptor, AhR, which is a specific high-affinity cellular protein (MANDAL, 2005).

Due to their potential risk to humans and animals, the United States Environmental Protection Agency (US-EPA) has designated 126 PAHs as priorities for environmental concern. Among those, the following 16, un-substituted PAHs are frequently monitored in environmental samples: naphthalene, acenaphthylene, acenaphthene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, benzo(a)anthracene, chrysene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(a)pyrene, dibenz(ah)anthracene, benzo(ghi)perylene, and indeno(1,2,3-cd)pyrene (http://water.epa.gov/scitech/methods/cwa/pollutants.cfm). Four of them, namely benzo(a)pyrene, benzo(k)fluoranthene, benzo(b)fluoranthene and indeno(1,2,3 – c,d)pyrene were chosen as indicator chemicals for monitoring obligations to reduce PAH emissions (http://water.epa.gov/scitech/methods/cwa/pollutants.cfm).
1.1.2 Benzo(a)pyrene (BaP)

BaP is a well characterized PAH not manufactured or used commercially. It is primarily a by-product of incomplete combustion but also occurs naturally in petroleum-based tars. BaP was identified as being the major tumour-producing agent in coal tar in 1932 (COOK et al., 1932; KENNAWAY, 1955). Because of its physicochemical properties (Table 1.1), BaP is found largely associated with particulate matter and has the tendency to accumulate in the environment through sorption to organic matter, animals and plants (JUHASZ and NAIDU, 2000; DONG, 2008). Since 1998 the US-EPA has assigned BaP as a priority level 1 due to its bioaccumulative, persistence and toxic capabilities. The BaP molecule is auto-fluorescent upon excitation with ultraviolet light, which enables its detection for quantitative analysis but also for visualization of chemical uptake (BARTLEY et al., 1982; SUREAU et al., 1990; BARHOUMI, 2002).

Table 1.1. Overview of physico-chemical properties of Benzo(a)pyrene (BaP).

<table>
<thead>
<tr>
<th>BaP Physico-Chemical Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>C_{20}H_{12}</td>
</tr>
<tr>
<td>Molecular weight</td>
<td>252.31 g mol(^{-1})</td>
</tr>
<tr>
<td>Aggregate state</td>
<td>Solid - 1013 mbar at 20°C</td>
</tr>
<tr>
<td>Density</td>
<td>1.28 - 1.35 g m(^{-3})</td>
</tr>
<tr>
<td>Log Kow</td>
<td>6.06</td>
</tr>
<tr>
<td>Water solubility</td>
<td>1.5 \times 10^{-8} mol L(^{-1})</td>
</tr>
<tr>
<td>Henry's law constant</td>
<td>4.9 \times 10^{-7} atm m(^{3}) mol(^{-1})</td>
</tr>
<tr>
<td>Boiling point</td>
<td>310 - 312°C at 13 mbar</td>
</tr>
<tr>
<td>Vapour pressure</td>
<td>5.6 \times 10^{-9} mm Hg</td>
</tr>
<tr>
<td>Excitation/Emission</td>
<td>264/408 nm</td>
</tr>
</tbody>
</table>

The highest environmental concentrations of BaP are found in soil samples and usually range from 0.8 ng kg\(^{-1}\) to 100 mg kg\(^{-1}\). Food and plants have BaP concentrations between 0.1 and 150 μg kg\(^{-1}\), the air has a concentration of around 1.3 to 500 ng m\(^{-3}\) and drinking water generally contains 2.5 to 9 ng L\(^{-1}\) (ANGERER et al., 1997). Due to its potential harm to human health, concentrations of BaP are regulated in U.S. and European public drinking water supplies by the maximum contaminant level (MCL) of 200 ng L\(^{-1}\) (www.epa.gov) and 10 ng L\(^{-1}\) (http://ec.europa.eu/environment/water/water-
drink/index_en.html, www.epa.gov), respectively. In fact, the International Agency for Research on Cancer (IARC) listed BaP as carcinogenic to humans (group 1) as a result of its high carcinogenic and cell toxicity potential (www.iarc.fr). Hattemer-Frey and Travis (1991) reported that the general U.S. population has a long-term average daily intake from ingestion and inhaling of BaP of around 31 ng kg$^{-1}$ bodyweight day$^{-1}$.

Uptake into organisms proceeds via uptake by cells which comprise environment-organism or organism-internal barriers. Yet, while several investigations have been conducted to detect BaP in environmental samples by means of high performance liquid chromatography (HPLC), few previous studies have investigated the uptake and distribution of BaP within an organism or a cell population. For example, Marie et al (2010) assessed the distribution of BaP intravenously injected in rats. They found a bi-exponential elimination of BaP in blood, liver, skin and kidney and a mono-exponential elimination in the adipose tissue and lung. In this study, most of the BaP was eliminated in faeces in the metabolized form. Few studies showed that BaP is taken up by cells relatively fast (within minutes). Sureau et al (1990), using microspectrofluorimetric technique, studied the uptake kinetics of BaP in a mammary tumour cell. They observed that a saturation in the cytoplasm was reached after 10 min. Barhoumi et al (2000) showed that in rat cells exposed to 10 µM of BaP, a fluorescence saturation was reached only at 4 h, although BaP was rapidly partitioned into cells within several min. In this study they also used specific probes for organelles and found indications of extensive BaP localization after 16 h of 10 µM BaP in Golgi, and in cytoplasmic membranes but not within the nucleus, the mitochondrial matrix, or within the lysosomes (BARHOUMI et al., 2000). In contrast, Allison et al (1966) had reported on the identification of BaP, based on its auto-fluorescence, in the lysosomes. It appears that the time and concentration of exposure may play a decisive role in the amount and localization of BaP within cells aside from differing metabolic capabilities of different types of cells. For example, Schaefer and Selkirk (1985), studying parental and variants of Hepal1c1c7 cells, demonstrated that the different cell populations can exhibit significant differences in the rate and extent of metabolism of BaP.
1.2 The Aryl Hydrocarbon Receptor Signal Transduction Pathway

1.2.1 The Aryl Hydrocarbon receptor (AhR) as transcription factor

The first xenobiotic-activated receptor (XAR) identified was the aryl hydrocarbon receptor (AhR) (MA, 2007). The receptor is a key transcriptional regulatory protein involved in the altered gene expression and toxicity that result from exposure to dioxins or dioxin-like compounds. Furthermore, the AhR possesses physiological functions that are independent of exogenous chemical exposure (MANDAL, 2005). For example, AhR null mice present an abnormal liver phenotype suggesting that hepatocyte growth was altered by the receptor deficiency (GONZALEZ and FERNANDEZ-SALGUERO, 1998). Studies at the cellular level showed that AhR deficiency leads to reduced cell proliferation compared to wild type cells (MA and WHITLOCK, 1996).

Biochemical investigations demonstrated that in the absence of an agonist, the AhR exists in the cytosol as an inactive complex with two molecules of the chaperon Hsp90, as well as one molecule of each of the chaperones, ARA9 and p23 (PERDEW, 1988; PETRULIS, 2003). Molecular cloning studies showed that the AhR contains a PER-ARNT-SIM (PAS) domain similar to that found in other regulators of cellular responses to chemicals in the environment (NGUYEN and BRADFIELD, 2008).

The AhR pathway is broadly well understood. The agonist binds at the PAS domain of the AhR and leads to a conformational change in the receptor. This interaction alters the associations with the chaperons and exposes a nuclear localization signal on the AhR (Figure 1.1, step II). As a result, the receptor complex migrates into the nucleus where it dissociates from the chaperones and forms a heterodimer with another bHLH-PAS protein denominated aryl hydrocarbon receptor nuclear translocator (ARNT) (Figure 1.1, step IV) (HANKINSON, 1994). This interaction increases the heterodimer capacity to bind with high affinity to specific enhancer sequences adjacent to target promoters termed dioxin responsive elements (DREs) (Figure 1.1, step V). Binding to DREs results in an increased transcription of several genes, including the cytochrome P450 (CYP) genes, such as cyp1a1, cyp1a2, cyp1b1, as well as a number of phase II metabolizing enzymes (Figure 1.1, step VI) (NEBERT et al., 2004; NGUYEN and BRADFIELD, 2008).

The AhR is also subjected to negative regulation via the 26S proteosome pathway and the aryl hydrocarbon receptor repressor (AhRR) (Figure 1.1, steps VII and VIII, respectively). The activated AhR is degraded through the ubiquitin-26S proteosome pathway that is controlled by the AhR degradation promoting factor (ADPF) in the nucleus. Besides that, the
activity of the AhR-ARNT complex is attenuated by induced expression of the AhRR. This protein belongs to the same family as the AhR and ARNT and has a high similarity with them. The repressor competes with the AhR in two different ways: (I) by competing for the ARNT and (II) if bound to ARNT, by binding to the DREs, thus blocking further attachment of the AhR/ARNT complex. The reduction of AhR activity is due to a negative feedback loop and the receptor degradation may work as a protective mechanism from the transcriptional hyperstimulation by some potent agonist and to provide precise temporal control of this pathway (NGUYEN and BRADFIELD, 2008).

**Figure 1.1:** AhR Pathway: (I) BaP entrance and distribution within the cell; (II) interaction with AhR (aryl hydrocarbon receptor) signalling pathway; (III) or organelles like mitochondria and lysosomes; (IV) BaP-AhR complex is translocated into the nucleus; (V) AhR dimerizes with ARNT (aryl hydrocarbon receptor nuclear translocator) and binds to the DRE (dioxin responsive elements) and subsequently (VI) induces transcription of several genes important for detoxification and many other processes; (VII) The complex AhR-ARNT suffers degradation via the 26S proteosome pathway; (VIII) - AHRR (Aryl hydrocarbon receptor repressor) can bind to the ARNT and prevent that this molecule binds to the AhR; besides that, the AHRR-ARNT complex can bind to the DRE and physically block the access of the AhR-ARNT complex to the DNA; (IX) metabolism of BaP by induced enzymes, e.g. Cyp1a1; (X) metabolites excreted from the cells; (XI) Some metabolites can enter the nucleus and form DNA adducts, e.g. BaP 7,8-diol9,10-epoxide.
Some evidence, like the similarity of amino acid sequence of the ligand-binding PAS domain protein, indicates that the AhR is quite conserved among the vertebrates, from marine to terrestrial and avian environments. Besides that, developmental abnormalities and pathologies in AhR null mice and ontogeny could be interpreted as a probable sign of an existence of an evolutionary endogenous regulator (Nguyen and Bradfield, 2008; Oesch-Bartlomowicz and Oesch, 2009). Several reports also suggest that up-regulation of AhR-target genes is occurring during embryonic development, participating in physiological processes involving the immune, cardiovascular and liver system. Support for this suggestion comes from an AhR-negative mouse liver Hepa1c1c7 cell line and mouse AhR-/- embryonic fibroblasts which grow more slowly than their AhR-positive counterparts (Ma and Whitlock, 1996). Several theories on the identity of endogenous ligands have been reported; nevertheless, all endogenous compounds tested until today were not proven to be an agonist with real relevance to AhR-associated physiology (Puga et al., 2002; Elferink, 2003; Fong et al., 2005; Andrysík, 2007; Nguyen and Bradfield, 2008).

The most well studied genes controlled by the AhR are those coding for enzymes involved in biotransformation, i.e. chemical metabolism, e.g. cyp1a1. However, recent studies using microarray have identified a substantial number of additional genes regulated by this signalling pathway. As part of the research programme of which this thesis was a part, Michaelson et al (2011) used a Random Forest classifier to predict AhR direct targets (Table 1.2). They also compared the results from their classifier to a study from Dere et al (2011), who explored a murine liver tissue after TCDD exposure for gene expression, and found that over 70 % of genes predicted by the classifier to be AhR responders had also been identified to be located in regions of AhR enriched binding. Among the genes predicted to be AhR direct targets, 12 were transcription factors (TFs, Table 1.2). Due to their predicted interactions, the regulation of those TFs indicates a cascade of induction of other TFs, initiated by the AhR. Along these lines, van Delft et al (2010) exposed HepG2 cell to 3 µM BaP for time scales varying from 3 to 60 h and stated that a complex network of a few transcription factors (TF) is responsible for the majority of the transcriptional changes induced by BaP.
Table 1.2. Primary targets of AhR predicted based on a time and concentration dependent gene expression data set (MICHAELSON et al., 2011).

<table>
<thead>
<tr>
<th>Hspa4l</th>
<th>Jun</th>
<th>Gm10122</th>
</tr>
</thead>
<tbody>
<tr>
<td>2410066E13Rik</td>
<td>Cdkn1a</td>
<td>Snx30</td>
</tr>
<tr>
<td>Arl6ip5</td>
<td>Abce4</td>
<td>Cdkn1b</td>
</tr>
<tr>
<td>Plsr2</td>
<td>Slec6a9</td>
<td>Slec26a2</td>
</tr>
<tr>
<td>Mpp2</td>
<td>Adh7</td>
<td>Plk2</td>
</tr>
<tr>
<td>Tiparp</td>
<td>Usp18</td>
<td>Zscan29</td>
</tr>
<tr>
<td>Sdpr</td>
<td>Npc1</td>
<td></td>
</tr>
<tr>
<td>Ndr1</td>
<td>Casp3</td>
<td>Nrg1</td>
</tr>
<tr>
<td>Nrn1</td>
<td>Aldh3a1</td>
<td>Abcd2</td>
</tr>
<tr>
<td>Cyp2s1</td>
<td>Slec35d1</td>
<td>Klf9</td>
</tr>
<tr>
<td>Tnfaip2</td>
<td>Cyp1b1</td>
<td></td>
</tr>
<tr>
<td>Cpox</td>
<td>Intu</td>
<td>Tnfaip8</td>
</tr>
<tr>
<td>Ospbl2</td>
<td>Pitpnc1</td>
<td>9330175E14Rik</td>
</tr>
<tr>
<td>Rbks</td>
<td>Sesn2</td>
<td>Lrc30</td>
</tr>
<tr>
<td>Ppard</td>
<td>Ceng2</td>
<td></td>
</tr>
<tr>
<td>Tbc1d16</td>
<td>Fam198b</td>
<td>Bmf</td>
</tr>
<tr>
<td>Arrdc3</td>
<td>Ddit4</td>
<td>Rnf39</td>
</tr>
<tr>
<td>Lpin1</td>
<td>Ubl3</td>
<td>St6gal1</td>
</tr>
<tr>
<td>Id2</td>
<td>Nqo1</td>
<td>Zfp36i1</td>
</tr>
<tr>
<td>Xdh</td>
<td>Trp53inp1</td>
<td>Nrd1d1</td>
</tr>
<tr>
<td>Gramd3</td>
<td>Cyp1a1</td>
<td>Irs2</td>
</tr>
<tr>
<td>Serpine1</td>
<td>Abca6</td>
<td>Ets2</td>
</tr>
<tr>
<td>Pfkb3</td>
<td>Hmox1</td>
<td>Nfe212</td>
</tr>
<tr>
<td>Jub</td>
<td>Aldh4a1</td>
<td>Irf1</td>
</tr>
<tr>
<td>Ddx58</td>
<td>Npffr1</td>
<td>Cib2</td>
</tr>
<tr>
<td>Zfp418</td>
<td>Btg2</td>
<td>S1pr1</td>
</tr>
<tr>
<td>Sgk1</td>
<td>Nr3e1</td>
<td>Traf5</td>
</tr>
</tbody>
</table>

Known transcriptional regulators (TFs) (i.e. annotated with relevant GO terms) are bolded.

1.2.2 BaP as a ligand of AhR

BaP is a well-known ligand of the AhR. According to the processes depicted in Figure 1.1, it binds to the AhR and triggers the translocation of the receptor to the nucleus. In the nucleus, the AhR-BaP binds to the ARNT and this new complex binds to the DREs. This leads to the regulation of diverse genes, such as genes enabling biotransformation and biological activation, as well as genes involved in immune response, cell development and female reproduction (BURCHIEL and LUSTER, 2001; PUGA et al., 2002).
Despite this general knowledge, details about the BaP entry into cells, distribution and internal concentrations able to elicit responses are largely unexplored and mechanisms of action incompletely understood. For example, Holmes and Pollenz (1997), using quantitative western blot, calculated the number of AhRs per different mouse, rat and human cells, but information such as the number of molecules actually present in the cells required to activate the AhR is lacking. Several analyses of regulation of selected genes were performed in vivo and in vitro to provide better understanding of the molecular response patterns on exposure to BaP but generally high concentrations (in the micro-molar range) and comparatively long exposure times (i.e. generally around 24 h) were employed (BARTOSIEWICZ et al., 2001). Those conditions were expected to cause a cellular response, however, they blend the action of BaP and potential BaP metabolites as well as gene expression under conditions of severe cell damage. Recently, efforts have been undertaken to derive more detailed time- and concentration-dependent transcriptome response profiles and link these to a phenotype response (HOCKLEY et al., 2006; HOCKLEY et al., 2007; HOCKLEY et al., 2008; VAN DELFT et al., 2010; MICHAELSON et al., 2011).

It is widely accepted that BaP requires biological activation through oxidative metabolism to be carcinogenic. BaP induces its own metabolism by activating the AhR pathway, leading to the induction of biotransformation enzymes (COSMAN et al., 1992; HANKINSON, 1995; MILLER and RAMOS, 2001). The metabolism of BaP is divided into three main phases. During phase I, BaP is functionalized by an enzymatic reaction or adoption of electrophile or nucleophile groups by oxidation, reduction or hydrolysis. Therefore, cytochrome P450s, e.g. Cyp1a1 and Cyp1b1, epoxide reductases and epoxide hydrolases are the most important phase I enzymes (GELBOIN, 1980). The monooxygenation of BaP leads to formation of oxidized metabolites and those are further metabolized by phase II enzymes. During phase II, the BaP metabolites are conjugated to hydrophilic molecules such as glucuronic acid or glutathione, making them more hydrophilic and for that reason easier to be excreted. The glutathione transferases, UDP-glucuronyl transferases, and sulfotransferases are enzymes involved in these processes (GELBOIN, 1980). Finally in phase III, metabolites are eliminated from cells through transport proteins, like the ATP binding cassette transporters (ABC-transporters) (MILLER and RAMOS, 2001).

Taken together, BaP metabolism leads to the production of a number of metabolites (Figure 1.2). The main groups of those metabolites are epoxides, dihydrodiols, phenols and quinones. The main epoxides are the 2,3-, 4,5-, 7,8- and 9,10-isomers, which are generated by monooxygenases (GELBOIN, 1980). Those epoxides are transformed by epoxide hydrolase to
General Introduction

4,5-, 7,8-, and 9-10-trans-dihydrodiol isomers. The epoxides can also be transformed into BaP-phenols (1-,3-,6-,7-, and 9-BaP-phenol-isomers) via non-enzymatic reactions.

Figure 1.2: Different primary and secondary metabolites of BaP which are formed through initial oxidation of the parent compound by cytochrome P450 enzymes (modified from Miller and Ramos (2001)).
BaP metabolites with a high carcinogenic potential can be formed during the transformation processes. One of the best characterized carcinogenic BaP intermediate is BaP-7,8-dihydrodiol-9,10-epoxide (BPDE) (Figure 1.3) (Frank et al., 1998; Miller and Ramos, 2001; Dreij et al., 2005). BPDE is able to bind to DNA at the N²-position of deoxyguanosine (BPDE-N²-dG) and to proteins, forming both DNA and protein adducts (Flowers et al., 1996). It is assumed that the formation of BPDE is responsible for the generation of the majority of DNA adducts in cells (Li et al., 1995; Geacintov et al., 1997; Shukla et al., 1997; Kozack and Loechler, 1999; Penning et al., 1999; Boysen and Hecht, 2003). DNA adducts may lead to cancer by causing mutations in genes essential for key functions, including apoptosis, proliferation and differentiation (Skipper et al., 1994; Chi et al., 2009). DNA repair is known to play a role in adduct removal. Adduct number has been shown to slightly decrease after cells were kept 72 h in BaP free medium (Bartley et al., 1982).

Figure 1.3: Chemical structure of the model compound BaP, investigated in this thesis, and some of the enzymes involved in its phase I metabolism, leading to the formation of BPDE.

1.3 Toxicogenomics

Toxicogenomics deals with the collection, interpretation and storage of information about gene and protein activity within particular cells or tissues of an organism in response to toxic substances. It combines toxicology with genomics or other high throughput molecular profiling technologies such as transcriptomics, proteomics and metabolomics (Lovett, 2000; Meyer et al., 2008). These high throughput profiling technologies allow scientists to address
mechanistic questions concerning the effects of toxicants. Recent advances, like gene expression arrays and bioinformatics, allow to assess the effects of xenobiotics on thousands of genes simultaneously, where before only isolated genes and/or pathways were addressed (LOVETT, 2000).

Assembling this high throughput data together can be done by using a systems biology approach. The goal of systems biology is to facilitate the comprehension of the changes in an integrated manner, through mechanism-oriented experimentation and modelling and by understanding the behaviour of entire cell populations or an organism by characterizing intrinsic interactions between its constituent parts (LOVETT, 2000; HECKER et al., 2009; KOIDE et al., 2009). This systems approach can be applied in toxicology in order to predict the cell response to a chemical, or to a mixture of stressors. A computational model of cell/organ/body system function would be desirable for several reasons. For example, it would enable a more targeted investigation and understanding of complex scenarios, such as the effect of multiple stressors. This model could then be used to predict chemical-cell interactions, thereby identifying toxic chemicals but also chemical structures or doses that are safe (MATERI and WISHART, 2007; HANDY, 2008).

Another method recently used in toxicology is the network biology approach. This approach also integrates different “omics” data within a systems biology framework but uses statistical inferences to predict connectivity among genes or proteins. Williams et al (2011) showed that it is possible to predict the presence of a chemical pollutant in complex mixtures and health outcomes for the exposed organism through this approach. Toxicogenomics is still in its infancy but is advancing rapidly due to an increasing pressure to reduce the number of animal tests in pharmacology and toxicology and also due to a rapid progress in the available technologies (LOVETT, 2000).

1.3.1 Transcriptomics

Transcriptomics builds on the knowledge of genomes. Although the genome is basically the same in every cell of an organism, different cells show different patterns of gene expression for different purposes but also to respond to external stimuli, such as chemicals. The transcriptome represents all RNA molecules, including mRNA, which are further translated into proteins. The entirety of mRNA is a mirror of the genes that are actively expressed in a cell or an organism at a certain time, thus the transcriptome constitutes a snapshot of all actively expressed genes at any given time (SCHIRMER et al., 2010).
Since the development of the microarray technique in 1995, there has been an enormous increase in gene expression data from several organisms. DNA microarrays allow examination of the function of thousands of genes at once in parallel. This method exploits the unique feature of single stranded DNA to hybridize to complementary DNA sequences (SELLHEYER and BELBIN, 2004). For example, several studies using mouse arrays had been done in the past years for purposes such as gene discovery, gene function and pathway research, toxicological studies and molecular characterization of disease (DERE et al., 2006; TIJET et al., 2006; BEYER et al., 2007; BOUTROS et al., 2009). Different mouse arrays are available commercially with one example being the Mouse Exon 1.0 ST array by Affymetrix that enables two complementary levels of analysis: gene expression and alternative splicing (www.affymetrix.com). Another example is the whole Mouse Genome array from Agilent Technologies, which offers arrays with all known genes and transcripts of the entire mouse genome (www.agilent.com). Another option is to design arrays with genes of interest for more specific uses. These customized arrays enable to focus on pre-selected genes, while the whole genome arrays provide an overall response without dependence on prior knowledge of (potentially) regulated genes.

DNA microarrays require previous knowledge of the DNA sequence and are thus applied frequently in model organisms. Other high throughput methods for gene expression analysis not requiring any previous knowledge about transcripts are increasingly available and of particular interest for non-model organisms. An example of one such method is whole transcriptome shotgun sequencing, designated as RNA-Seq (SCHIRMER et al., 2010).

1.3.2 Phenotypic anchoring (Adverse outcome pathway)

Phenotype anchoring is defined as the linking of molecular responses, including those at the transcript, protein, and metabolite levels, to changes of the phenotype observed at the cell, organismal or population level (CUNNINGHAM et al., 2003; PAULES, 2003; SCHMIDT, 2003; SCHIRMER et al., 2010). Until now phenotype anchoring studies have mostly built on the transcriptome to directly connect it to the organism or population phenotype (SCHIRMER et al., 2010). As an extension to phenotype anchoring, the so-called adverse outcome pathway (AOP) approach has recently been proposed as a concept to link sub-cellular mechanisms to chemical response to outcomes of relevance for environmental risk assessment, e.g. altered growth and reproduction as population level effects (ANKLEY et al., 2010). Each AOP is a set of chemical, biochemical, cellular, physiological, behavioural, among others, responses which
characterize the biological effect cascade resulting from a particular molecular initiating event. An AOP is a description of plausible causal linkages which illustrate how molecular initiating events result in other critical biological effects quantified at the cellular, tissue, organ and whole animal levels of observation. A particular chemical might have a variety of interactions with a biological system. The molecular initiating event is normally the first critical step of the sequence of events that are essential to the induction of the outcome (Figure 1.4) (ANKLEY et al., 2010; SCHULTZ, 2010). Risk assessment will better predict effects once the integration of mechanistic information and relevant phenotypes are achieved (SCHIRMER et al., 2010).

![Figure 1.4. Conceptual diagram of key features of an adverse outcome pathway (ANKLEY et al., 2010).](image)

### 1.3.3 Cellular models

Cell lines provide an almost unlimited supply of cells with similar genotypes and phenotypes. Their use avoids variation between individuals and bypasses moral and ethical issues associated with animal and human experiments. Unless specialized culture conditions are used, within a few passages of initiation of a cell culture, a relatively uniform population of proliferative cells is selected. If these cells can be propagated, a cell line arises. Cell lines are generally easier to control and more reproducible than complex systems, e.g. animals and tissues models (HANKINSON, 1995; MASTERS, 2000; BOLS et al., 2005). An external stress can be applied in a cell-based system under controlled conditions and sampling done more easily at various conditions and time points. Moreover, the use of cell cultures enables comparison of different species at the cellular level under similar circumstances (BOLS et al., 2005).
One such cellular model, which was selected for this thesis work, is the well characterized mouse hepatoma cell line Hepa1c1c7 (Figure 5). This cell line was derived from the BW7756 hepatoma, which arose in a C57L mouse and propagated in C57L/J mice (HANKINSON, 1979). It conserves several hepatic functions, e.g. transferrin secretion and haptoglobin synthesis, as well as phase I and phase II biotransformation reactions (FONG et al., 2005). This cell line specifically has been studied with regard to the AhR pathway and AhR quantification, (MA and WHITLOCK, 1996; HOLMES and POLLENZ, 1997; BARHOUMI et al., 2000) and several mutants of this cell line are available for AhR-related gene expression and functional studies (HANKINSON, 1979; FONG et al., 2005).

![Figure 5](image.png)

**Figure 5:** Hepa1c1c7 cell line after 48 h (left) and after 72 h (right) in culture, plated at a density of $10^4$ cells cm$^{-2}$ and viewed by confocal microscopy transmitted light.

### 1.4 Scope of the thesis

This thesis research was performed as part of a Systems Biology project entitled: “From contaminant molecules to cellular response: system quantification and predictive model development” funded by the Helmholtz Association, Alliance on Systems Biology (Germany). This project aimed at understanding cellular responses to chemical stressors with a systems perspective, which involved intra-cellular transport of the chemical, reaction with sub-cellular target sites and the cell’s response at the transcriptional and posttranscriptional level. The PAH Benzo(a)pyrene (BaP) was selected as the model chemical because it represents an important class of environmental contaminants and because it affects the aryl
hydrocarbon receptor (AhR) signalling pathway, which comprises a central route for toxic effects of many organic chemicals. The research project overall was divided among several research groups into 5 modules as follows: 1) uptake and intracellular distribution; 2) interaction with the AhR pathway; 3) gene expression; 4) protein expression; 5) in silico signalling pathway.

The goal of this thesis specifically was to investigate the alteration of gene expression in the murine liver cell line, Hepa1c1c7, in a time- and concentration-dependent manner on exposure to BaP, combining high throughput technique, bioinformatics, phenotype characterization and chemical analysis. Time of exposure varied from 2 to 48 h and the BaP concentration ranged from 0.5 nM to 5 µM for most of the experiments.

The first objective (Chapter 2) was to determine the phenotype and alteration of the transcriptome of Hepa1c1c7 cells exposed to BaP. Thus, phenotype end points such as cell proliferation, lysosome integrity, metabolic activity, formation of reactive oxygen species (ROS) and Cyp1a catalytic activity were measured. Based on the effects on the cell line, two BaP concentrations (50 nM and 5 µM BaP) and four time points (2, 4, 12 and 24 h) were then selected for transcriptome analysis by whole mouse genome microarray and the data integrated by various bioinformatics techniques.

Based on the transcriptome response acquired in Chapter 2, the second objective (Chapter 3) of this thesis was to understand the functional role of TCDD-inducible poly(ADP)ribose protein (tiPARP). tiPARP was characterized as an AhR dependent regulated gene and frequently found in microarray studies of cells exposed to dioxin like compounds. tiPARP regulation occurred very rapidly (within the first hours) and transiently (for less than 12 h) in Hepa 1c1c7 cells upon exposure to both BaP concentrations (Chapter 2). This response pattern led to the hypothesis that tiPARP was involved in early damage control induced by dioxin like compounds. In order to study its function, small interference RNA (siRNA) was used to silence Hepa1c1c7 cells and physiological endpoints were tested to determine the differences between the wild type from the silenced cells. Protein determination by western blot and mass spectrometry was attempted to define the endogenous presence of this protein. Unfortunately, the efforts here used were not sufficient to detect the protein in a regular state or to reveal its function.

The third objective of this thesis (Chapter 4) was to quantify the BaP existing in the cells. Due to the results obtained in Chapter 2, it was theorized that for the low (50 nM) BaP concentration, the cells were able to metabolize and excrete most of the BaP, and repair any potential DNA damage, while for the high concentration of BaP (5 µM), damage by BaP and
its metabolites overwhelmed cellular defence and mechanisms of repair. To prove this hypothesis, cells were exposed to radiolabelled BaP under the same conditions as for the transcriptome study (Chapter 2). Liquid scintillation (LS) and fluorescence and radio high pressure liquid chromatography (HPLC) were used to determine the total amount of radioactivity and BaP concentration, respectively, in each compartment of the exposure setup (culture medium, plastic and cells). Cell number was determined for each time point of each exposure condition in order to calculate internal BaP concentrations per cell. DNA adducts were determined to correlate internal BaP concentration to DNA damage state.

Due to the collaborative nature of this project, a portion of the data obtained in this thesis supported other project parts of the Systems Biology initiative. Specifically, the transcriptome analysis was instrumental in the development of a Random Forest machine learning approach to differentiate specific, receptor-driven transcriptional responses from the more general toxic response. This work was published in BMC genomics (MICHAELSON et al., 2011) and is presented in the Appendix section. Moreover, contributions were made to a review on “Transcriptomics in ecotoxicology” (SCHIRMER et al., 2010). Finally throughout this thesis, reference is given to the study of Dautel et al (2011). This study also resulted from the Systems Biology project whereby the same cell material was analysed by proteomics (DAUTEL et al., 2011) and transcriptomics (this thesis).
1.5 References


General Introduction


Hankinson, O., The role of the aryl hydrocarbon receptor nuclear translocator protein in aryl hydrocarbon receptor action. *TEM* 1994, 5, 240-244.


Holmes, J. L.; Pollenz, R. S., Determination of aryl hydrocarbon receptor nuclear translocator protein concentration and subcellular localization in hepatic and nonhepatic cell culture lines: Development of quantitative western blotting protocols for calculation of aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein in total cell lysates. *Molecular Pharmacology* 1997, 52 (2), 202-211.


Hong, S., Candelone, J.-P., Patterson, C.C. and Boutron, C.F., History of Ancient Copper Smelting Pollution During Roman and Medieval Times Recorded in Greenland Ice. *Science* 1996, 272 (5259), 246-249.


Penning, T. M.; Burczynski, M. E.; Hung, C. F.; McCoull, K. D.; Palackal, N. T.; Tsuruda, L. S., Dihydrodiol dehydrogenases and polycyclic aromatic hydrocarbon activation:
General Introduction


Schaefer, E. L.; Selkirk, J. K., Metabolism of Benzo(a)Pyrene by Variant Mouse Hepatoma-Cells. Cancer Research 1985, 45 (8), 3487-3492.

Schirmer, K.; Fischer, B. B.; Madureira, D. J.; Pillai, S., Transcriptomics in ecotoxicology. Analytical and Bioanalytical Chemistry 2010, 397 (3), 917-923.

Schmidt, C. W., Toxicogenomics. An emerging discipline. EHP toxicogenomics : journal of the National Institute of Environmental Health Sciences 2003, 111 (1T), A20-5.


Chapter 2

2. Differential time-resolved molecular response of a mouse liver cell line on low and high benzo(a)pyrene exposure

Submitted to BMC Genomics.
Chapter 2

Abstract

Benzo(a)pyrene (BaP) is a well-known and widely distributed environmental carcinogen produced by incomplete combustion of organic matter. BaP induces its own metabolism through the activation of the aryl hydrocarbon receptor (AhR) pathway. During its metabolism, several by-products and reactive metabolites are produced which contribute to toxicity and carcinogenesis. Although these processes are generally well understood and time and concentration are two well accepted variables of cellular response to chemical exposure, detailed knowledge of the molecular networks leading to different phenotypic responses are thus far little understood. The research presented here therefore explores time-resolved transcriptional signatures of a murine hepatoma cell line on low (50 nM) and high (5 µM) BaP exposure, leading either to recovery from initial insult or to sever damage of cells. Despite the 100-fold difference in BaP exposure concentrations, the pattern of transcript regulation was comparable in cells for up to 4 h of exposure. A total of 20 genes were regulated by the low BaP concentration with about half of them also being regulated by the high BaP exposure condition. However, while the transcriptome returned to control levels for the low concentration of BaP between 4 and 12 h of exposure, the transcriptome response diverged for the high concentration of BaP, with more than 1000 genes identified as significantly regulated. Selected descriptors of (sub)cellular phenotype were related to the regulation of selected groups of genes. Metabolism of BaP was indicated by regulation, among others, of cyp1a1 mRNA and Cyp1a catalytic activity for both exposure concentrations. Cells exposed to the low BaP concentration showed a slight increase in cell doubling time but no impact on cell viability or signs of formation of ROS. In contrast, cells exposed to the high BaP concentration greatly reduced proliferation, showed accumulation of ROS and were prone to cell death. Based on the regulation of the transcriptome, the 4 h time point appears to be decisive for the cells’ fate. The slight decrease in proliferation on low exposure to BaP is an indication that cells allow for damage repair in order to return to normal cycling. However, on exposure to the high BaP concentration, this defence is overruled and cells undergo cell death.
Benzo(a)pyrene (BaP) is a widely studied and well-known environmental carcinogenic pollutant belonging to the polycyclic aromatic hydrocarbons (PAHs). BaP is formed by incomplete combustion of organic materials, like cigarette smoke and car exhaust, and is found in air pollution, diet and in some occupational atmospheres, being identified as a major tumor producing agent in coal tar already in 1932 (KENNAWAY, 1955). It is generally accepted that BaP requires biological activation through oxidative metabolism to be toxic and carcinogenic. BaP induces its own metabolism by activating the aryl hydrocarbon receptor (AhR) pathway. Indeed, cells deficient in AhR-dependent BaP activation have been shown to be refractory to BaP exposure (SCHIRMER et al., 2000) and animal knockout studies demonstrated that the loss of AhR results in reduced toxicity and carcinogenicity of BaP and TCDD (GONZALEZ and FERNANDEZ-SALGUERO, 1998; SHIMIZU et al., 2000). AhR is a cytosolic transcription factor that, upon activation, translocates into the nucleus and binds to responsive DNA elements, thereby regulating the expression of many genes, including several encoding for enzymes that enable biotransformation. Examples of such enzymes are cytochrome P450s, e.g. Cyp1a1, epoxide hydrolases, dehydrogenases and some peroxidases.

Enzymatic transformation of BaP leads to a number of (in part chemically active) metabolites and by-products which may contribute in different ways to toxicity and carcinogenesis. Among the by-products are reactive oxygen species (ROS), which can cause oxidative stress, apoptosis and cancer through induction of DNA damage (KLAUNIG et al., 2010). Chemically active metabolites are formed, like BaP quinones, which are capable of electron redox-cycling leading to the formation of ROS (BURDICK et al., 2003). Most attention has been given to the 7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo(a)pyrene (BPDE) metabolite, which is known to undergo covalent binding to DNA forming stable DNA adducts (SIMS et al., 1974). These adducts may lead to cancer by causing mutations in genes essential for key functions, including apoptosis, proliferation and differentiation (BAIRD et al., 2005; XUE and WARSHAWSKY, 2005). Despite the general knowledge on BaP activation, the dynamics of molecular responses by cells to BaP exposure are not yet well understood. Evidently, BaP elicits a multitude of outcomes which depend on the cellular environment and the time- and concentration of exposure.

Characterizing transcriptional signatures, initiated either directly by the AhR or by secondary responses not regulated by AhR, are key to fundamental knowledge of the interaction of BaP with cells. High throughput technologies, such as microarrays, have
become instrumental in the discovery of such signatures (BARTOSIEWICZ et al., 2001; JOHNSON et al., 2003; KESHAVA et al., 2005). However, only recently have these approaches been applied to derive more detailed time- and concentration-dependent transcriptome response profiles (HOCKLEY et al., 2006; HOCKLEY et al., 2007; HOCKLEY et al., 2008; VAN DELFT et al., 2010; MICHAELSON et al., 2011). Hockley et al (2006; 2007) exposed the breast cancer cell line, MCF-7, and the human hepatoma cell line, HepG2, for 6, 24 and 48 h to 0.25, 1, 2.5 and 5 µM BaP in order to identify gene expression changes related to carcinogenesis using a 6 and 15 K cDNA microarrays. Indeed, several tens of genes were regulated in a time- and concentration dependent fashion, belonging to processes of nucleosome assembly, chromatin structure organisation, oncogenesis and cell cycle progression. However, marked differences between the two cell lines in the time- and concentration-dependent regulation of genes and presence of DNA adduct formation were noted. One such difference was that MCF-7 cells, but not HepG2, responded with a clear pattern of continuous gene induction/repression over time to the two high (2.5 and 5 µM) BaP concentrations while for the two low (0.25 and 1 µM) BaP concentrations, gene expression overall appeared to return to control levels toward the end, i.e. 48 h, of exposure. Using a more comprehensive 22 K oligonucleotide microarray, van Delft et al (2010) explored the gene expression of HepG2 cells exposed to 3 µM BaP for time scales varying from 3 to 60 h in order to identify relevant transcription factor networks. Independent of whether the transcriptional response was early (up to ~12h), intermediate (~12 to 30h) or late (~30 to 60h), the majority of the regulated genes was found to be connected to five transcription factors (NF-κB, c-MYC, SRF, AP1 and E2F1), indicating that a complex network of a few TFs is responsible for a large part of the transcriptional response. However, because this network hardly changed over time while the transcriptome response to BaP clearly did, the authors concluded that also other regulatory mechanisms must be involved. Michaelson et al (2011), using the data set established in this thesis, explored a matrix of time-resolved gene expression values, derived from BaP exposed murine hepatoma cells, Hepa1c1c7, to train a Random Forest classifier that allows to predict direct targets of the AhR and distinguish them from genes responding to AhR-independent side-effects. Twelve TFs were found among the AhR direct targets, suggesting that these regulators could trigger a transcriptional cascade that starts with the activation of AhR.

In this study, a transcriptome data set was established (the same being subsequently used by Michaelson et al (2011)) in order to compare and contrast the temporal transcriptional response of Hepa1c1c7 cells to a low (50 nM) BaP concentration, where no phenotype response was observed, and a high (5 µM) BaP concentration, where cellular damage was
present. The cell line was selected because of knowledge available from previous studies with the cell line regarding to AhR pathway signalling and quantification of the AhR receptor (HOLMES and POLLENZ, 1997b; MA, 2007). The overall response to the two explored BaP concentrations was found to greatly diverge over time with a tendency to return to control levels, being noticeable for the low concentration of BaP. It was demonstrated that the greatest overlap in the transcriptional response occur as early as 4 h of exposure and the detailed response patterns of groups of genes and their relation to selected descriptors of (sub)cellular phenotype while cells succumb to BaP toxicity or recover from the initial insult was explored.
2.2 Material and Methods

**Chemicals**

Benzo(a)pyrene (≥ 96 % purity, Sigma Aldrich, Switzerland) was pre-dissolved in Dimethyl sulfoxide (DMSO, ≥ 99.9 % purity, Sigma Aldrich, Switzerland) in a stock solution with a concentration of 10 mM. 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) was a kind gift from Dr. Peter Schmid from EMPA, Material and Science Technology, Switzerland and was prepared as 0.5 mM stock solution. TCDD was only used as a positive control for measuring the Cyp1a catalytic activity in Hepa1c1c7 cells.

**Cell cultivation**

The mouse hepatoma Hepa1c1c7 cell line was grown in Dulbecco’s modified medium without phenol red (Biochrom AG, Berlin) and complemented with 7 % of fetal calf serum (Biochrom AG), 1 % of penicillin/streptomycin (Sigma Aldrich) and 2 mM of L-alanine and L-glutamine (Biochrom AG, Berlin).

**Experimental design**

The cell density of $10^4$ cells cm$^{-2}$ was chosen based on previous works (FONG et al., 2005) and because it allowed the cells to proliferate throughout the chosen time course. The cells were seeded on the first day, one day after the medium was exchanged, and two days later the cells were exposed to the chemical, BaP in DMSO or DMSO as solvent control. The final concentration of the vehicle in the culture medium was 0.05 % (v/v). BaP was used in a range varying from 0.5 nM to 5 µM and the time of exposure was from 2 to 48 h. The concentration range was selected based on environmental relevance rarely studied thus far (nanomolar range) and concentrations frequently used in previous studies focusing on BaP cellular toxicity (micromolar range).

**Phenotype characterization**

*Cell proliferation.* Cell number was assessed to determine the population growth rate and the generation doubling time. Cells were grown at the same density as described above in 75 cm$^2$ culture flasks. Cell number was investigated for a period of 96 h, having the culture medium exchanged one day after plating. After another 24, 48, 50, 52, 60, 72, 96 h, the cells were detached from the flasks using trypsin and counted using a cell counter (Casy®, Schärfe System GmbH). The growth rate ($\mu$) was determined by the slope of the linear regression and
the doubling time (Td) was calculated using the formula: $T_d = \frac{\ln(2)}{\mu}$. Cell number was also determined by counting the cells with the Neubauer chamber for each time point and BaP concentration from the exposure set up. All experiments were done in biological triplicates.

**Cell viability assay.** For the determination of the sensitivity of Hepa1c1c7 cells to BaP, the cells were exposed to the chemical in 12 well plates (Greiner, Switzerland). After exposure, cell viability assays using different fluorescent dyes were performed as described by Schirmer et al (1998), altering only the incubation temperature to 37° C. AlamarBlue (Invitrogen, Basel, Switzerland) was used as a measure of metabolic activity, 5-carboxyfluorescein diacetate acetoxyethyl ester (CFDA-AM, Invitrogen, Basel, Switzerland) as an indicator of membrane permeability, and neutral red (Invitrogen, Basel, Switzerland) as a measure of lysosome membrane integrity. All the measurements were made with technical and biological triplicates. The results were expressed as percentage of the appropriate control.

**EC$_{50}$ determination.** Data collected from cell proliferation and viability assays were plotted in order to determine which concentration of BaP is needed to inhibit the tested endpoint by 50 % (EC$_{50}$). The formula used to derive the EC$_{50}$ value was $Y=$Bottom+$\frac{(Top-Bottom)}{(1+10^{((LogEC50-Log[BaP])*HillSlope)})}$, where $Y$ is the response, ranging from 0 % (bottom) to 100 % (top) with a sigmoidal shape. The software GraphPad Prism 4 was used for the calculations.

**ROS assay.** Production of reaction oxygen species (ROS) was accessed using H$_2$DFFDA (Molecular Probes™, Invitrogen, Switzerland) (10 µM as final concentration) in a time (2, 4, 8, 12 and 24 h) and concentration (0.5 nM to 5 µM BaP) dependent manner. Fluorescence was measured by plate reader and excitation and emission wavelengths were set to 485 and 530 nm, respectively. Total protein was also measured in order to normalise the data. Excitation and emission wavelengths were set to 360 and 460 nm and bovine serum albumin (BSA) curves was used as protein standard. The experiment was done in technical and biological triplicates.

**Cyp1a catalytic activity.** 7-Ethoxyresorufin-O-deethylase (EROD) assays were performed using the fluorometric assay described by Kennedy and Jones (1994) using a plate reader. Total protein was determined using bovine serum albumin as a protein standard. Excitation and emission wavelengths were set to 530/25 nm and 590/35 nm, respectively, for the EROD
assay. Total protein was also measured using excitation and emission wavelengths set at 360 and 460 nm and bovine serum albumin (BSA) curves were used as standard. The experiment was performed in triplicates using both BaP and TCDD.

**RNA isolation, Microarray & Bioinformatics Analysis**

For the microarray study the cells were grown in 75cm² flasks and the experiment was done in triplicates. The time points chosen were 2, 4, 12 and 24 h and two concentrations, namely 50 nM and 5 µM BaP was used. After the exposure time the cells were harvested using Trizol (Invitrogen) and RNA was extracted using RNA easy kit (Qiagen, Valencia, CA, USA). RNA was quantified and integrity verified on a Bioanalyzer (Agilent Technologies, Palo Alto, CA). Sample preparation for Affymetrix GeneChip Mouse Exon 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA) was performed following the manufacturer's recommendations.

Microarrays were normalized using Robust Multichip Average (RMA) and a one way ANOVA and LIMMA analysis were performed. One way ANOVA was performed using as independent variable the different treatments. The data was then cut-off for a false discovery rate (fdr) < 0.1. This generated a list of genes per time point. Then, this data was used to do a principal component analysis using the MATLAB software (version 2010a).

LIMMA analysis (SMYTH, 2004), which uses an empirical Bayes method, was used to identify genes differentially expressed in one group compared to the control group. In this case this method was used to differentiate the gene expression over time using the cells exposed to DMSO as control. The cut-off of the gene expression was done using a fdr < 0.1, a p-value < 0.05 and a log2 ratio of 0.5 for both up and down regulated genes.

Annotation was performed by means of the Database for Annotation, Visualisation and Integrated Discovery (DAVID) (http://david.abcc.ncifcrf.gov/) using the list of differentially expressed genes identified by LIMMA analysis (HUANG et al., 2009b; HUANG et al., 2009a).

Pathway, network and gene-set enrichment analyses were applied system-wide, using the MetaCore™ software (GeneGo, San Diego, CA) (EKINS et al., 2005; NIKOLSKY et al., 2005). The significance of changes in expression in pathways or networks is based on the degree of overlap between the dataset described in this study and a set of genes corresponding to a specific network or pathway queried. MetaCore™ was used to examine connectivity between the genes, using their network building algorithms (EKINS et al., 2007; DEZSO et al., 2009). Networks were analysed using the ‟Transcription regulation” algorithm, where the
created networks are centered on transcriptional factors. The network is built dynamically, connecting all genes via direct mechanistic interactions based on manually curated literature evidence.

*Real time quantitative polymerase chain reaction (qPCR) gene confirmation.* Among of the early (2 and 4 h) response genes, a subset of 11 was used to confirm the microarray expression, as shown in Table 2.1. The RNA was extracted according to description above. cDNA synthesis was performed according to manufacture descriptions of TaqMan®, Reverse Transcription Reagents (Applied Biosystems, Switzerland), using a cycle of 25°C for 5min, 48°C for 30 minutes, 95°C for 10 min and 4°C for 1min. For real time amplification a SyBR Green® PCR master mix (Applied Biosystems, Switzerland) was used and the cycle used was 95°C for 10 min and 94°C for 15 seconds for 40 cycles, and 60°C for 1 minute for elongation. qPCR expression levels relative to untreated controls were calculated using the relative expression method described by Pfaffl et al (2002). Thus, statistically significant differences between control samples and treated samples were evaluated in group means by randomization test using REST software. Differences were considered to be significant at $p < 0.05$. 
2.3 Results

(Sub-)cellular phenotype

Cell proliferation and viability: BaP impaired the proliferation of cells in a time- and concentration-dependent manner as judged by cell numbers (Figure 2.1A, Figure S1) and doubling time (Figure 2.1B). Exposure to 50 and 500 nM BaP led to a slight increase in doubling time, while for 5 µM BaP concentration proliferation was completely inhibited. Viability of adherent cells was significantly impaired only at the highest exposure concentration and only after 24 h or more (Figure 2.1C). Metabolic activity (as judged by alamarblue staining) and lysosomal membrane integrity (neutral red staining) were more strongly affected than cytoplasm membrane integrity (CFDA-AM staining) (Figure 2.1C, D). Alteration of cell volumes was not observed (data not shown). Because of its severe impact on cell viability, BaP exposures for 48 h were omitted from subsequent investigations.
Figure 2.1: Figure 1: Impact of BaP on Hepa1c1c7 cell proliferation and viability. All experiments were performed in biological triplicates and average and standard deviations are shown. A- Cell number determination after DMSO and 50 nM/5 µM BaP exposure over time (see Fig. S1 for results obtained for additional BaP concentrations). Asterisks indicate significant difference as compared to the 2 h time point as determined by one-way ANOVA followed by Dunnett’s Multiple Comparison Test (*p-value < 0.05; **p-value < 0.01). B- Cell doubling time in the absence (DMSO control) and presence of BaP ranging in concentration from 0.5 nM to 5 µM. C- Dose response relationships obtained for three different indicators of viability for cells exposed over time and the 0.5 nM to 5 µM concentration range of BaP). D- EC50 values obtained for the three indicators of viability after 24 and 48 h exposure to BaP.
Chapter 2

**Cyp1a catalytic activity:** Cells possessed an easily measurable constitutive Cyp1a catalytic activity, represented in the graphic with the solid line, which tended to decline with culture time [as seen both for exposure to BaP (Figure 2.2) and TCDD (Figure S2), with the latter being used as positive control]. These results indicate the ability of the cells to oxidize and thus metabolize low concentrations of BaP without up-regulating Cyp1a enzyme activity at least up to 12 h of static exposure. A concentration-dependent induction of Cyp1a activity was seen at this time point and further increased over the next hours up to the 24 h observation point. The apparent decline of Cyp1a activity for the two highest BaP concentrations can at least in part be explained by a reduced cell viability (see Fig. 2.1C).

![Figure 2.2](image)

**Figure 2.2.** Cyp1a enzymatic activity, measured as 7-ethoxyresorufin-O-deethylase (EROD) activity in Hepa1c1c7 cells exposed to BaP ranging in concentration from 1 nM to 5 µM. Experiments were performed in biological triplicates and average and standard deviations are shown. Asterisks indicate significant difference as compared to the DMSO control as determined by one-way ANOVA followed by Dunnett’s Multiple Comparison Test * p-value < 0.05 and ** p-value<0.01. **A**- 2 h, **B**- 4 h, **C**- 12 h and **D**- 24 h of exposure (see Fig. S2 for results obtained with TCDD, which served as positive control). Solid line represents basal Cyp1a activity while dashed lines reflects the confidence interval.
**Reactive oxygen species (ROS) formation:** Significant ROS accumulation was seen only for 12 and 24 h of exposure to the highest exposure concentration. This was demonstrated both quantitatively as well as by microscopic observation (Figure 2.3 and Figure S3).

![Figure 2.3](image)

Figure 2.3. Reactive oxygen species (ROS) quantification in Hepa1c1c7 cells exposed to BaP, ranging in concentration from 0.5 nM to 5 µM, using H2DFFDA as indicator: A- quantification of fluorescence normalized to total protein. Experiments were performed in biological triplicates and average and standard deviations are shown. Asterisks indicate significant difference as compared to the DMSO control as determined by one-way ANOVA followed by Dunnett's Multiple Comparison Test **p-value<0.01; B- visualisation by confocal microscopy. Representative images are shown for DMSO controls (top panel), and 12 h (middle panel) and 24 h (bottom panel) 5 µM BaP exposed cells. Small (inset) pictures (25 µm) show details of few cells while the bigger pictures are representative of the area (50 µm).
Analyses of Microarray

Microarray data were analysed using two different statistical approaches: (I) one way ANOVA was used to differentiate the response among the treatments (DMSO, 50 nM and 5 µM BaP) and only the significant ones where subsequently used in the principle component analysis (PCA); and (II) LIMMA (Linear models of microarray analysis) (SMYTH, 2004) was used to identify the time-dependence of the transcriptional response, using DMSO treated cells at each respective time point as a control (Supplement information_excel file).

**PCA:** A total of 1004 genes were found to be regulated in a concentration dependent manner. PCA revealed significant differences in the expression of genes in cells exposed to the high (5 µM) BaP concentration for 12 and 24 h compared to all other conditions, namely DMSO and the low (50 nM) BaP concentration (Figure 2.4). This indicates that cells exposed to the two BaP concentrations start to diverge in responses between 4 and 12 h of exposure.

![Principal Component Scatter plot](image)

**Figure 2.4.** Principal component analysis based on statistically significantly regulated genes among the treatments (DMSO, 50 nM and 5 µM BaP) as determined by one way ANOVA. Blue dots represent DMSO control samples, orange dots 50nM BaP and red dots 5µM BaP samples for each biological replicate (n=3).
**Time dependence of transcriptional response:** Different genes were regulated at the different time points; however, when taking all time points together, a total of 20 genes was differentially expressed on exposure to 50 nM BaP and 2251 genes were differentially regulated by exposure to 5 µM BaP (Figure 2.5). The 20 genes identified for the low BaP exposure were also differentially expressed on high BaP exposure with the majority of them being direct targets of the AhR (Table 2.1). A subset of genes regulated at the early time points was selected for confirmation of microarray results by real time PCR (Table S1, S2). The greatest correspondence between genes regulated at the low and the high BaP concentration was at 2 and 4h (Table 2.1, Table S2 and Figure 2.6). After this time point, the initially regulated genes returned to control levels for the low BaP concentration with one exception, *cyp1a1* (Table 2.1). In contrast, most of the genes regulated at 4 h were also differentially expressed at 12 and 24 h for the high BaP concentration (Table 2.1). Moreover, the number of regulated genes drastically increased after the 4 h time point (Figure 2.5). The time-dependent differences for the two exposure concentrations are explained in more detail below.

**Figure 2.5:** Venn diagram showing the distribution of genes regulated over time on exposure to **A**- 50 nM (for the 24 h time point no gene was significantly regulated) and **B**- 5 µM BaP. The purple circle represents the number of genes regulated at 2 h, the yellow at 4 h, green at 12 h and blue at 24 h of exposure to BaP.
Table 2.1. Group classification, identity and expression pattern over time of genes significantly regulated within the first 4 h of exposure of Hepa1c1c7 cells to BaP. Red colour = up-regulation; blue colour = down-regulation.

<table>
<thead>
<tr>
<th>Group</th>
<th>Genes</th>
<th>AhR target</th>
<th>qPCR 2 h</th>
<th>4 h</th>
<th>12 h</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td>Biotransformation</td>
<td>cyp1a1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nqo1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>aldh7</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>aldh3a1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cyp2s1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cpox</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>myot1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ndf1tc</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mmp2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nedd4d</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nm1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>slt1a4</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>abcc4</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tbc1d16</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stdpl</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>adipor2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>serpine</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>jub</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tpr53/3np1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>asg1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hspa4l</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ic2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>foxq1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ddx58</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>arrdc3</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>zfp608</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ndrg1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>nif39</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tiparp</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>spr-psi1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>tmem71*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>trav8d1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9030624G23Rik*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* no described function * data also at Michaelson et al (2011).
The 2 and 4 h time points: The genes regulated at both concentrations after 2 and 4 h exposure can be assigned into five groups: biotransformation, cytoskeleton and membrane repair and transport, cell cycle and survival, transcription factors (TF) and TF activators, and miscellaneous (Table 2.2 and Table S3). Identical genes were regulated at 2 h: mpp2, tiparp, rnf39 and cyp1a1, all direct AhR targets (Table 2.1). Among those, two genes, cyp1a1 and mpp2, remained regulated for both BaP concentrations at 4 h of exposure, while rnf39 and tiparp, remained up-regulated in the high BaP concentration only. Additional genes started to be regulated at this time point for the high BaP concentration (Figure 2.5 and 2.6). No down-regulated genes appeared in the high concentration, while for the lower concentration, spr-ps1 (sepiapterin reductase pseudogene 1), trav8d-1 (T cell receptor alpha variable 8D-1) and 9030624G23RIK (RIKEN cDNA 9030624G23 gene) were down regulated. Unfortunately, very little information exists regarding the function of those genes (Table 2.1).
### Table 2.2. Genes differentially expressed at 2 and 4 h for both 50 nM and 5μM BaP exposed cells.

<table>
<thead>
<tr>
<th>Function thus far reported</th>
<th>Gene name</th>
<th>Comments</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>cyp1al</em></td>
<td>AhR direct target; PAHs metabolism</td>
<td>(GUENGERICH and SHIMADA, 1991)</td>
</tr>
<tr>
<td></td>
<td><em>np1</em> (NAD(P)H dehydrogenase, quinone 1)</td>
<td>AhR direct target; well-known gene involved in the metabolism of PAH; over expression inhibits p53 degradation (due to its property to stabilize p53).</td>
<td>(DOUGHERTY et al., 2008; DANKOVA-KOSTOVA and TALALAY, 2010; TSYETKOV et al., 2010)</td>
</tr>
<tr>
<td></td>
<td><em>Cyp2a1</em> (cytochrome P450, family 2, subfamily s, polypeptide 1)</td>
<td>AhR direct target; regulated both by AhR receptor and by hypoxia, however no biological function was yet determined, some studies suggest that this enzyme may play a role in the detoxification and bioactivation of xenobiotics and has been also proposed to be involved in embryonic development.</td>
<td>(DUB and BANDIERA, 2010; BEBENEK et al., 2011)</td>
</tr>
<tr>
<td></td>
<td><em>adhr</em> (alcohol dehydrogenase 7 class IV, mu or sigma polypeptide)</td>
<td>AhR direct target; represents a metabolic barrier against ethanol by having a predominant role in the first-pass metabolism, however this enzyme can have other substrates like lipid peroxidation products like 4-hydroxynonenal.</td>
<td>(GELSKI and SZMIEKOWSKI, 2008)</td>
</tr>
<tr>
<td></td>
<td><em>Aldh3a1</em> (aldehyde dehydrogenase family 3, subfamily A1)</td>
<td>AhR direct target; known to protect against HNE (4 hydroxy-2-nonenal, aldehyde derived from lipid peroxidation) induced protein adduct formation and growth inhibition and is also known to enhance cell survival, <em>aldh3a1</em> expression is activated by the AhR pathway, where it might play a role in cell survival by oxidizing peroxides generated by oxygen radicals.</td>
<td>(REISDORPH and LINDAHL, 2007)</td>
</tr>
<tr>
<td></td>
<td><em>Cpox</em> (coproporphyrinogen oxidase)</td>
<td>AhR direct targets; codifies for an enzyme that is involved in heme synthesis; might be responsible for the regulation of heme synthesis.</td>
<td>(TAKETANI et al., 2001)</td>
</tr>
<tr>
<td></td>
<td><em>Myof</em> (myoferlin)</td>
<td>Belongs to the ferlin family of membrane proteins, which has been associated to a variety of processes as cell mobility, growth factor receptor stability (VEGF receptor 2), endocytosis, and membrane repair.</td>
<td>(BERNATCHEZ et al., 2007; EISENBERG et al., 2011; YU et al., 2011)</td>
</tr>
<tr>
<td></td>
<td><em>Nsl1</em> (NSFL1 (p97) cofactor (p47))</td>
<td>ATPase associated with various cellular activities involved in membrane trafficking events.</td>
<td>(DREVENY et al., 2004; BEURON et al., 2006; UCHIYAMA et al., 2006)</td>
</tr>
<tr>
<td>Gene</td>
<td>Function and Regulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>npp2</strong></td>
<td>AhR direct target; tumour suppressor proteins in vertebrates (structural and regulatory role in tissue organization and maintenance and in cell signalling and membrane trafficking); negatively regulates c-Src kinase activity preventing disorganization of the actin cytoskeleton in lamellipodia of cells. (Baumgartner et al., 2009)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nedd41</strong> (neural precursor cell expressed, developmentally down-regulated gene 4-like)</td>
<td>Regulates several cellular processes such as excretion, plasma membrane channel regulation (Na⁺), protein catabolism and TGF-β signalling, most expressed in liver, brain, heart, kidney and lung. (Hellwinkel et al., 2011; Takeuchi et al., 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nrm1</strong> (neurexin 1)</td>
<td>AhR direct targets; correlated with maturation of hepatocytes, it is regulated by hypoxia and was implicated in tumorigenesis by promoting changes in cell morphology, anchorage-independent growth and tumour formation. (Le Jan et al., 2006; Han et al., 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Scl1a4</strong> (solute carrier family 1 (glutamate/neural amino acid transporter), member 4)</td>
<td>Encodes a Na⁺-dependent neutral amino acid transporter, in brain is preferentially expressed in glial cells where L-serine is synthesised and only liberated by SLC1A4 in exchange of other substrates. (Kanai and Hediger, 2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Abcc4</strong> (ATP-binding cassette, subfamily C (CFTR/MRP), member 4)</td>
<td>AhR direct target; related to transporting BaP metabolites. (Xu et al., 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tbc1d16</strong> (TBC1 domain family, member 16)</td>
<td>AhR direct target; predicted to be involved in vesicular trafficking, however it was demonstrated that this gene might play a role in cell survival and proliferation. (Akaavia et al., 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oshd2</strong> (oxysterol binding protein-like 2)</td>
<td>Protein localized in lipid droplets (inhibiting triglyceride, phospholipid and cholesterol ester metabolism) but is also known to disrupt the trafficking ER-Golgi when is over expressed. (Ridgway, 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adipor2</strong> (adiponectin receptor 2)</td>
<td>Codifies for an integral membrane protein that binds to the hormone adiponectin, that is reported to promote cell survival and growth but might also stimulate apoptosis; increases migration of human chondrosarcoma cells, mainly expressed in liver cells. (Chiu et al., 2009; Burchler et al., 2010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Serpine1</strong> (serine (or cysteine) peptidase inhibitor, clade E, member 1)</td>
<td>AhR direct target; also known as pas-1, is involved in suppressing fibrinolysis, leading to the pathological fibrin deposition and tissue damage, is also involved in cell replication, negative regulation of cell cycle, and angiogenesis and is considered to be an acute phase reactant, being influenced by inflammatory cytokines. (McMahon and Kwaan, 2008; Cesari et al., 2010; Higgins et al., 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Name</td>
<td>Description</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td><em>Jub</em> (jubha)</td>
<td>Ahr direct target; protein that shuttles in and out of the nucleus and is involved in different functions like, cell-cell junction, cell migration, cell cycle regulation and modulator of IL-1 induced NF-κB activation.</td>
<td>(Hirota et al., 2003; Marie et al., 2003; Feng and Longmore, 2005; Amri et al., 2006; Montoya-Duran et al., 2008).</td>
<td></td>
</tr>
<tr>
<td>Trp53inp1 (transformation related protein 53 inducible nuclear protein 1)</td>
<td>Ahr direct target; key stress induced p53 target gene, that modulates apoptosis and cell cycle arrest at G1-phase.</td>
<td>(Tomasini et al., 2005; Nguessan et al., 2011; Seux et al., 2011).</td>
<td></td>
</tr>
<tr>
<td>Oxigin1 (Oxidative stress induced growth inhibitor 1)</td>
<td>Negative regulator of cell growth, proapoptosis function, may also protect cells against oxidative stress.</td>
<td>(Li et al., 2007; Yao et al., 2008).</td>
<td></td>
</tr>
<tr>
<td>Hspa4l (heat shock protein 4-like)</td>
<td>Ahr direct target; was suggested by to be involved in anti-apoptosis.</td>
<td>(Takahashi et al., 2007).</td>
<td></td>
</tr>
<tr>
<td>Id2 (inhibitor of DNA binding 2)</td>
<td>Ahr direct target; transcription factor described to play an important role in defining the differentiation fate of peripheral lymphocytes and their response to infection; necessary for the differentiation of several immune cell lineages and also implicated in the development of red blood cells, mammary glands, spermatogenesis and normal lung development; induce cell cycle arrest and prevent differentiation of enteroendocrine precursor during embryogenesis.</td>
<td>(Belz and Rankin, 2011).</td>
<td></td>
</tr>
<tr>
<td>Foxp1 (forkhead box Q1)</td>
<td>Belongs to the forkhead-factor family of transcription factor and has been showed to directly control the expression of p21 and to enhance tumor growth and tumorigenicity in colorectal carcinomas, cell shape and size, actin cytoskeleton, cell migration, cell plasticity, cell-cell contact and cell proliferation in epithelial cells; foxp1 expression changed TGFβ1 induced cell dissociation and TGFβ1 dependent gene expression.</td>
<td>(Feuerborn et al., 2011).</td>
<td></td>
</tr>
<tr>
<td>Ddx58 (DEAH (Asp-Glu-Ala-His) box polypeptide 8)</td>
<td>Ahr direct targets; also known as rig-1 (retinoic acid-inducible gene 1 protein) is a gene involved in virus recognition, although recent studies indicates that can also have other function, e.g., in inflammation, cancer and apoptosis, it also initiates a signalling cascade that converge to three families of transcription factors, NFκB, IRF3/7 and ATF-2/Jun.</td>
<td>(Vidalain and Tangy, 2010; Salminen et al., 2011).</td>
<td></td>
</tr>
<tr>
<td>Arrdc3 (arrestin domain containing 3)</td>
<td>Ahr direct target; component of the adhrβ2 (adrenergic receptor, beta 2) down regulation process.</td>
<td>(Nabian et al., 2010).</td>
<td></td>
</tr>
<tr>
<td>Zfp608 (zinc finger protein 608)</td>
<td>Ahr direct targets; controlling the thymocyte development.</td>
<td>(Zhang et al., 2006).</td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Description</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>Ndrgl</td>
<td>AhR direct target; highly conserved among multicellular organisms and it is expressed ubiquitously in response to cellular stress signals.</td>
<td>(Ellen et al., 2008)</td>
<td></td>
</tr>
<tr>
<td>mlf59</td>
<td>AhR direct target; expressed in rat, late after birth, in the cortex, cerebellum and hippocampus and in adults in the lung, liver, stomach, small intestine and in the brain</td>
<td>(Matsuo et al., 2001)</td>
<td></td>
</tr>
<tr>
<td>Tiparp</td>
<td>AhR direct target; Hepatic gluconeogenesis</td>
<td>(Diani-Moore et al., 2010)</td>
<td></td>
</tr>
<tr>
<td>Spr-psl</td>
<td>82% homology to the spr gene involved in BH4 pathway.</td>
<td>(Auerbach et al., 1997, Lee et al., 1999)</td>
<td></td>
</tr>
<tr>
<td>Tmeq7l</td>
<td>No described function.</td>
<td>MGI</td>
<td></td>
</tr>
<tr>
<td>Traw8d-1</td>
<td>No described function.</td>
<td>MGI</td>
<td></td>
</tr>
<tr>
<td>9030624G23Rik</td>
<td>No described function.</td>
<td>MGI</td>
<td></td>
</tr>
</tbody>
</table>

MGI: Mouse Genomics Informatics (www.mgi.org).
Chapter 2

Figure 2.6: Venn diagram showing the differentially expressed genes at 4 h on 50nM and 5µM of BaP exposure. Red indicates up-regulated, while blue indicates down-regulated, genes.

The 12 and 24 h time points: At the lower BaP concentration all the genes, with the exception of the cyp1a1 at 12 h, returned to control levels, while for the higher concentration the number of genes up and down regulated increases dramatically. A total of 656 genes are regulated for both time points, while 473 are only regulated at the 12 h and 1036 only at the 24 h time point.

According to gene ontology (GO) analysis from the DAVID annotation, regulated groups of genes were related to cell cycle, organelle, regulation of transcription and oxidation reduction. Identified through KEGG pathway analysis were the cell cycle and Huntington’s disease at 12 h and p53 signalling pathway at 24 h of exposure (Table S3).

Pathway analysis: Time-dependent pathway analysis was carried out for the 5µM BaP case including all but the 2 h time point (the latter having too few genes). A total of 11 pathways were identified, in particular, due to the number of genes being regulated, AP-1 dependent regulation of cell metabolism, cytoskeleton and membrane remodelling and TGF-dependent regulation (Table 2.3).
Table 2.3: GeneGo pathway-analysis listing the significance of pathway regulation at 4, 12 and 24 h for Hepa1c1c7 cells exposed to 5µM BaP.

<table>
<thead>
<tr>
<th>GeneGo Pathway-Maps</th>
<th>Time point</th>
<th>pValue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. AP-1 in regulation of cellular metabolism</td>
<td>4</td>
<td>3.6 e-2</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>4.5 e-8</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>3.1 e-7</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.4 e-2</td>
</tr>
<tr>
<td>2. Cytoskeleton remodelling</td>
<td>12</td>
<td>8.1 e-8</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>6.9 e-5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.7 e-2</td>
</tr>
<tr>
<td>3. TGF-beta receptor signalling</td>
<td>12</td>
<td>1.6 e-7</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>1.0 e-5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.0 e-1</td>
</tr>
<tr>
<td>4. TGF, WNT, and cytoskeletal remodelling</td>
<td>12</td>
<td>3.3 e-7</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>3.8 e-7</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.7 e-4</td>
</tr>
<tr>
<td>5. TGF-beta-dependent induction of EMT via SMADS</td>
<td>12</td>
<td>1.7 e-6</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>2.2 e-2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.2 e-2</td>
</tr>
<tr>
<td>6. IL-1 signalling pathway</td>
<td>12</td>
<td>1.1 e-4</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>2.1 e-6</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.4 e-2</td>
</tr>
<tr>
<td>7. TGF-beta-dependent induction of EMT via MAPK</td>
<td>12</td>
<td>4.5 e-6</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>6.2 e-4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.2 e-2</td>
</tr>
<tr>
<td>8. Chemokines and adhesion</td>
<td>12</td>
<td>7.0 e-6</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>1.8 e-4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2.3 e-2</td>
</tr>
<tr>
<td>9. Glucocorticoid receptor signalling</td>
<td>12</td>
<td>5.7 e-5</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>5.5 e-4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6.0 e-2</td>
</tr>
<tr>
<td>10. Regulation of epithelial-to-mesenchymal transition (EMT)</td>
<td>12</td>
<td>6.9 e-3</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>1.6 e-4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3.7 e-2</td>
</tr>
<tr>
<td>11. PLAU signalling</td>
<td>12</td>
<td>6.9 e-3</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>3.6 e-2</td>
</tr>
</tbody>
</table>
Transcription factors (TF): The number and identity of regulated TFs was investigated in order to highlight their role in the time-dependent increase in gene regulation for the exposure of cells to the high concentration of BaP. A total of 38 TFs were regulated at 12 and 24 h, with 14 being up regulated and 24 down regulated (Figure 2.7). Of the TFs being regulated at 12 h of exposure, four (EGR1, c-Jun, c-Fos, ATF-3) were up-regulated while all others were down-regulated. At the 24 hour exposure point, the number of up-regulated TFs was much higher and additional four TFs were down-regulated (SREBF2, Maf1, XBP1, ARNT).

The network-building algorithm on transcriptional regulation from MetaCore® (GeneGo) was used to examine whether the modulated genes are connected to TF. Analysis of the 2251 genes being regulated in total upon exposure to the high BaP concentration revealed significant networks centred around 20 TFs. The number of genes in each network varied from 137 to 487 (Table 2.4).

Figure 2.7. Venn diagram showing the distribution of significantly regulated transcription factors over time. Gene names in red indicate up-regulation while blue indicates down-regulation.
### Table 2.4: Key transcription factors targeted in the 5 μM BaP data set.

<table>
<thead>
<tr>
<th>Network</th>
<th>GO processes</th>
<th>Total nodes</th>
<th>Seed nodes</th>
<th>p-Value</th>
<th>zScore</th>
<th>gScore</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP1</td>
<td>cellular process (94.2%; 1.391e-43), negative regulation of biological process (41.3%; 8.829e-43), cellular metabolic process (71.3%; 1.291e-42), negative regulation of cellular process (39.0%; 2.131e-42), response to organic substance (33.7%; 4.719e-42)</td>
<td>487</td>
<td>486</td>
<td>0.000E+00</td>
<td>115.94</td>
<td>115.94</td>
</tr>
<tr>
<td>HNF4-alpha</td>
<td>cellular metabolic process (65.5%; 2.193e-26), metabolic process (70.2%; 4.703e-25), primary metabolic process (62.8%; 7.746e-21), cellular macromolecule metabolic process (47.9%; 9.268e-16), nitrogen compound metabolic process (42.5%; 1.950e-15)</td>
<td>456</td>
<td>455</td>
<td>0.000E+00</td>
<td>112.02</td>
<td>112.02</td>
</tr>
<tr>
<td>c-Myc</td>
<td>cellular metabolic process (75.8%; 8.706e-44), metabolic process (78.2%; 2.581e-37), primary metabolic process (72.9%; 2.949e-36), cellular process (92.6%; 4.830e-29), cellular nitrogen compound metabolic process (51.3%; 1.641e-28)</td>
<td>381</td>
<td>381</td>
<td>0.000E+00</td>
<td>102.68</td>
<td>102.68</td>
</tr>
<tr>
<td>p53</td>
<td>positive regulation of cellular process (48.2%; 1.011e-32), negative regulation of biological process (47.0%; 1.286e-31), negative regulation of cellular process (44.6%; 2.601e-31), regulation of cell death (31.7%; 1.265e-30), regulation of cell cycle (24.5%; 2.986e-30)</td>
<td>250</td>
<td>249</td>
<td>0.000E+00</td>
<td>82.73</td>
<td>82.73</td>
</tr>
<tr>
<td>ESR1 (nuclear)</td>
<td>cellular process (92.7%; 7.243e-19), positive regulation of biological process (41.5%; 1.277e-17), positive regulation of cellular process (38.9%; 2.337e-17), regulation of transcription from RNA polymerase II promoter (22.6%; 1.458e-16), negative regulation of cellular process (35.5%; 2.464e-16)</td>
<td>236</td>
<td>235</td>
<td>0.000E+00</td>
<td>80.35</td>
<td>80.35</td>
</tr>
<tr>
<td>CREB1</td>
<td>response to organic substance (42.5%; 1.743e-34), response to stress (49.6%; 2.944e-34), negative regulation of biological process (50.0%; 1.050e-33), negative regulation of cellular process (47.3%; 4.586e-33), cellular response to</td>
<td>228</td>
<td>227</td>
<td>0.000E+00</td>
<td>78.96</td>
<td>78.96</td>
</tr>
<tr>
<td>Gene</td>
<td>Function and Regulation Details</td>
<td>Score</td>
<td>FDR</td>
<td>p-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------</td>
<td>-------</td>
<td>-----</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGR1</td>
<td>regulation of cell proliferation (34.6%; 6.664e-30), positive regulation of biological process (52.9%; 2.339e-28), negative regulation of cellular process (47.1%; 7.990e-28), negative regulation of biological process (49.2%; 1.399e-27), positive regulation of metabolic process (38.2%; 2.939e-27)</td>
<td>194</td>
<td>194</td>
<td>7.090E-287</td>
<td>73.15</td>
<td>73.15</td>
</tr>
<tr>
<td>AP-1</td>
<td>response to organic substance (45.3%; 1.200e-31), response to stress (52.0%; 1.014e-30), developmental process (60.9%; 1.230e-26), response to chemical stimulus (54.7%; 1.273e-26), regulation of cell proliferation (33.5%; 1.740e-26)</td>
<td>180</td>
<td>180</td>
<td>5.940E-266</td>
<td>70.45</td>
<td>70.45</td>
</tr>
<tr>
<td>NF-kB</td>
<td>positive regulation of biological process (61.5%; 5.214e-37), response to organic substance (49.1%; 1.105e-35), positive regulation of cellular process (57.4%; 9.344e-35), regulation of cell proliferation (38.5%; 1.145e-32), response to stress (54.4%; 1.693e-32)</td>
<td>171</td>
<td>171</td>
<td>1.590E-252</td>
<td>68.66</td>
<td>68.66</td>
</tr>
<tr>
<td>RelA (p65 NF-kB subunit)</td>
<td>positive regulation of biological process (64.3%; 1.895e-38), positive regulation of cellular process (61.1%; 1.225e-37), negative regulation of cellular process (54.8%; 7.782e-33), negative regulation of biological process (56.7%; 3.154e-32), positive regulation of metabolic process (43.9%; 3.021e-30)</td>
<td>160</td>
<td>160</td>
<td>3.880E-236</td>
<td>66.41</td>
<td>66.41</td>
</tr>
<tr>
<td>GATA-1</td>
<td>positive regulation of biological process (47.2%; 1.089e-17), response to stress (43.4%; 1.155e-17), positive regulation of cellular process (44.7%; 1.258e-17), protein modification process (34.6%; 4.039e-16), macromolecule modification (35.2%; 9.980e-16)</td>
<td>160</td>
<td>159</td>
<td>1.850E-232</td>
<td>65.98</td>
<td>65.98</td>
</tr>
<tr>
<td>E2F1</td>
<td>positive regulation of cellular process (54.8%; 6.083e-29), positive regulation of biological process (56.7%; 4.234e-28), regulation of cell cycle (29.9%; 9.847e-28), cellular metabolic process (81.5%; 1.233e-25), cellular macromolecule metabolic process (70.1%; 4.405e-25)</td>
<td>159</td>
<td>158</td>
<td>5.650E-231</td>
<td>65.77</td>
<td>65.77</td>
</tr>
<tr>
<td>Protein</td>
<td>Description</td>
<td>p-value</td>
<td>FC</td>
<td>Score</td>
<td>p-value</td>
<td>Score</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>--------</td>
<td>-----</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>GCR-alpha</td>
<td>negative regulation of cellular process (48.1%; 8.920e-24), system development (53.9%; 4.333e-23), negative regulation of biological process (49.4%; 1.169e-22), positive regulation of macromolecule metabolic process (37.0%; 1.370e-22), positive regulation of metabolic process (38.3%; 2.377e-22)</td>
<td>155</td>
<td>154</td>
<td>4.920E-225</td>
<td>64.93</td>
<td>64.93</td>
</tr>
<tr>
<td>c-Jun</td>
<td>regulation of cell proliferation (42.7%; 2.142e-35), response to organic substance (50.7%; 5.716e-34), positive regulation of biological process (60.0%; 5.382e-31), positive regulation of cellular process (57.3%; 6.475e-31), cellular response to chemical stimulus (42.7%; 1.886e-30)</td>
<td>152</td>
<td>152</td>
<td>3.090E-224</td>
<td>64.73</td>
<td>64.73</td>
</tr>
<tr>
<td>Androgen receptor</td>
<td>positive regulation of metabolic process (38.2%; 6.819e-22), positive regulation of biological process (52.0%; 7.342e-22), positive regulation of cellular process (49.3%; 9.665e-22), regulation of macromolecule metabolic process (57.2%; 6.510e-21), regulation of gene expression (51.3%; 1.030e-20)</td>
<td>154</td>
<td>153</td>
<td>1.500E-223</td>
<td>64.71</td>
<td>64.71</td>
</tr>
<tr>
<td>ETS1</td>
<td>positive regulation of biological process (53.1%; 2.437e-22), positive regulation of cellular process (49.7%; 2.075e-21), regulation of cell death (34.0%; 2.367e-21), negative regulation of biological process (49.0%; 2.692e-21), regulation of programmed cell death (32.7%; 2.988e-20)</td>
<td>151</td>
<td>151</td>
<td>9.470E-223</td>
<td>64.51</td>
<td>64.51</td>
</tr>
<tr>
<td>AP-2A</td>
<td>response to stress (51.0%; 2.763e-25), negative regulation of biological process (51.0%; 2.232e-24), negative regulation of cellular process (47.1%; 1.715e-22), regulation of cell death (34.0%; 3.913e-22), cellular response to stress (29.4%; 3.267e-21)</td>
<td>153</td>
<td>152</td>
<td>4.570E-222</td>
<td>64.5</td>
<td>64.5</td>
</tr>
<tr>
<td>C/EBPbeta</td>
<td>positive regulation of cellular process (55.4%; 4.019e-28), positive regulation of biological process (56.1%; 7.345e-26), developmental process (62.2%; 1.269e-23), response to organic substance (42.6%; 3.856e-23), multicellular organismal development (58.1%; 4.698e-22)</td>
<td>149</td>
<td>148</td>
<td>3.920E-216</td>
<td>63.64</td>
<td>63.64</td>
</tr>
<tr>
<td>Gene</td>
<td>Function and Pathways</td>
<td>Fold Change</td>
<td>p-value</td>
<td>q-value</td>
<td>FDR</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Elk-1</td>
<td>toll-like receptor 2 signaling pathway (13.4%; 1.767e-22), toll-like receptor 1</td>
<td>135</td>
<td>5.400E-199</td>
<td>60.99</td>
<td>60.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>signaling pathway (13.4%; 1.767e-22), MyD88-dependent toll-like receptor signaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pathway (13.4%; 3.994e-22), toll-like receptor 4 signaling pathway (13.4%; 7.185e-22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Toll signaling pathway (13.4%; 8.699e-22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YY1</td>
<td>cellular metabolic process (83.7%; 1.070e-24), metabolic process (87.4%; 1.221e-23),</td>
<td>137</td>
<td>2.350E-198</td>
<td>60.98</td>
<td>60.98</td>
<td></td>
</tr>
<tr>
<td></td>
<td>response to inorganic substance (20.0%; 3.488e-16), primary metabolic process (75.6%;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.812e-16), respiratory electron transport chain (11.1%; 2.775e-15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.4 Discussion

While time and concentration are two well accepted variables of cellular response to chemical exposure, detailed knowledge of the molecular networks leading to different phenotypic responses are thus far little understood. This study demonstrates that a low (50 nM) and a high (5 µM) BaP concentration lead to a similar initial transcriptional response (up to 4 h of exposure). This response is characterized by direct targets of the transcription factor AhR, of which BaP is a ligand. After this initial response, the transcriptional patterns largely diverge and are associated with very different phenotypes. While the transcriptome overall returns to control levels for the low BaP exposure concentration, the number of regulated genes and involved pathways continues to increases for the high BaP concentration over the observed 24 h of exposure. Exposure for both concentrations was associated with a sustained induction of cyp1a1 mRNA although resulting catalytic enzyme activity, measured as EROD activity, was significantly induced only at concentrations > 100 nM BaP. Induction of ROS was significant after 12 h of exposure to 5 µM BaP, indicating the formation of reactive by-products in the Cyp1a catalysed transformation of BaP. Cytotoxicity commenced within 24 h for exposure concentrations beyond 100 nM (BaP). The only phenotypic impact for the 50 nM BaP exposure concentration was a slight increase in cell doubling time while for the 5 µM BaP concentration, proliferation was completely abolished.

Despite a 100-fold difference in BaP exposure concentrations, the transcriptional response was small but identical after 2 h, and similar in terms of extent and groups of regulated genes after 4 h, of exposure. This observation indicates fast uptake of BaP by the cells and a low response threshold for the involved genes. According to Barhoumi et al (2000), cell internal steady-state concentrations of BaP (10 µM) were reached within 4 h of exposure of a rat liver cell line. As demonstrated in Chapter 4 of this thesis, investigations with the Hepa1c1c7 cell line confirmed these observations. It was determined that the 4 h exposure to 50 nM BaP resulted in $1.9 \times 10^6$ molecules of BaP per cell while exposure to 5 µM BaP resulted in $1.7 \times 10^9$ molecules at this time point (Chapter 4). Holmes and Pollenz (1997a) determined the number of AhRs and ARNT per cell of the Hepa1c1c7 cell line to be $3.2 \times 10^5$ and $3.3 \times 10^4$, respectively. Thus, given the ratio of BaP molecules vs. the number of AhR per cell, maximum occupation and activation of the AhR/ARNT at both BaP concentrations may explain the similarity in the transcriptional response to both 50 nM and 5 µM BaP for the first hours of exposure.
Four genes, all known or predicted direct targets of AhR (BARTOSIEWICZ et al., 2001; HOCKLEY et al., 2006; MICHAELSON et al., 2011), were regulated for both BaP concentrations at the 2 hour time point: \textit{cyp1a1} (cytochrome P450 1A1), \textit{mpp2} (membrane protein palmitoylated 2), \textit{tiparp} (TCDD-inducible poly(ADP-ribose) polymerase, PARP7) and \textit{rnf39} (ring finger protein 39). \textit{Cyp1a1} is a well-known gene involved in phase I of PAH metabolism. The other genes are much less well defined. \textit{Mpp2} was co-localized with the cytoskeleton in human breast endothelial cells and found to control activity of c-Src, which is known to regulate many cell functions, including cell proliferation (BAUMGARTNER et al., 2009). \textit{tiPARP} was described as a mediator of TCDD-induced suppression of hepatic gluconeogenesis (DIANI-MOORE et al., 2010), and \textit{rnf39}, one of those human synonyms is Hematopoietic Zinc Finger (Hzf), appears to be involved in p53-mediated cell cycle arrest in response to DNA damage (SUGIMOTO et al., 2006). The time- and concentration-dependent induction of those genes beyond the 2 h period differed. In general, induction was more sustained for the 5 µM BaP exposure. This is in support of the prolonged nuclear translocation of the AhR observed for the 5 µM vs. 50 nM BaP concentration after 24 h of exposure (MICHAELSON et al., 2011). Despite this, maximum induction of \textit{tiparp} and \textit{rnf39} occurred only at 2 h for both BaP concentrations and only \textit{cyp1a1} and \textit{mpp2} were induced throughout the 24 h for the high concentration of BaP. The reason for this differential expression can only be speculated upon at this point but a fast, transient response may indicate important regulatory roles and tight regulation of \textit{tiparp} and \textit{rnf39} (Hzf). Das et al. (2007) have proposed a model in which \textit{hzf}, induced upon mild genotoxic stress, binds to p53 in an auto-regulatory feedback loop, directing p53 preferentially toward cell cycle arrest growth promoters, while severe or irreparable stress causes ubiquination of Hzf, thereby preventing this regulation and allowing p53 to activate pro-apoptotic targets. Moreover, MacPherson and Matthews have experimental support to suggest that the principal role of tiPARP is to negatively regulate AhR-mediated signalling (personal communication, August 2012). Indeed, establishing the functional role of \textit{mpp2}, \textit{tiparp} and \textit{rnf39} in response to BaP would be important next steps.

The highest number of genes being regulated by the low BaP concentration was found at the 4 hour time point. About half of those genes were also regulated upon exposure to the high BaP concentration and the majority are known or predicted targets of the AhR. Considering the functional groups to which the regulated genes can be tentatively assigned supports the importance of the 4 hour time point in cell fate. For example, in addition to \textit{cyp1a1}, \textit{nqo1} is induced as a second major phase I PAH metabolising enzyme and induction
of *cpox* is likely to support sufficient synthesis of heme protein, which is required as a cofactor for *cyp1a1*. These results, together with the confirmed Cyp1a catalytic activity, indicate extensive metabolism of BaP by the liver cells. Interestingly, *adh7* and *adh3a1*, which are being induced by the high BaP concentration only, have been described to play a protective role by metabolizing peroxides generated by ROS (Reisdorph and Lindahl, 2007; Jelski and Szmitkowski, 2008). Apparently, such protective mechanisms are overwhelmed by the high concentration of BaP, leading to accumulation of ROS and cytotoxicity as time progresses. Another interesting functional group concerns genes involved in cytoskeleton organization and membrane trafficking and repair. Tekpli et al (2010) revealed that 50 nM of BaP lead to membrane remodelling which in turn impacts on transmembrane proteins, such as the Na+/H+ exchanger isoform 1 (NHE1). In the current study, several genes involved in the regulation of plasma membrane proteins were induced. For example, *nedd4l*, which plays a role in Na+ channel regulation, was induced by 50 nM BaP at the 4 hour time point and *Slc1a4*, a gene encoding for a Na+-dependent neutral amino acid transporter was regulated by the high BaP concentration at 4 h of exposure and beyond (Kanai and Hediger, 2004; Hellwinkel et al., 2011; Takeuchi et al., 2011). Regarding to genes involved in membrane trafficking, *nslflc* was regulated by 50 nM BaP at the 4 h time point. The protein related to this genes has also been shown to be regulated by BaP (Dautel et al., 2011). Nsfl1c has an ATPase activity and is present in Golgi stacks (Dreveny et al., 2004; Beuron et al., 2006; Uchiyama et al., 2006). It thus appears that cells respond rapidly on the gene expression level to changes brought about to the plasma or cell internal membranes by BaP diffusing through or into these structures.

While gene expression returned to constitutive levels for 50 nM BaP beyond 4 h of exposure, a diverse and complex transcriptional response was seen for the high BaP concentration. Regulated genes could broadly be assigned to xenobiotic biotransformation, cytoskeleton/membrane repair/transformation, cell cycle/survival, TF/TF activator and miscellaneous. Pathway analysis showed AP-1 dependent regulation as one of the most important pathways, which is in accordance with van Delft et al (2010) who identified the TF AP-1 to modulate a high percentage of genes on exposure of HepG2 cells to 3 µM BaP. Cytoskeleton remodelling was also among the most regulated pathways by MetaCore™, confirming the findings of Dautel et al (2011) and Tekpli et al (2010). TGFβ and glucocorticoid are also relevant pathways due to their broad control of the cell machinery. TGFβ has an anti-proliferative effect in normal cells and in early cancer stages. The glucocorticoid receptor control genes that control immune response. Both pathways also
cross-talk to other pathways, thus controlling other cellular functions. The type of function employed by those three pathways are in accordance with the phenotypes observed for the high BaP concentration, like losses of ability to proliferate and severe loss in cell viability. Identification of group of genes that function in the same pathway reduces complexity and identifying pathways that differ between two conditions can have more explanatory power than a simple list of different genes or proteins (KHATRI et al., 2012).

Aside from AP-1, nineteen other TFs were found to be at the centre of the large transcriptional response to 5 µM BaP. Among those, Ap-1, SP-1, c-Myc, p53, ESR1, CREB-1, EGR-1, NF-κβ, E2F1, GCRα, c-Jun and Androgen receptor, were also described by van Delft et al (2010) to be the most regulated. They revealed that the most influential TFs were NFκβ, c-Myc, SRF, AP-1, E2F1, P53, CREB1, NRF2 and ESR1, in a descending order (VAN DELFT et al., 2010). From those only Srf and Nrf2 (also known as nfe2l2) were not present in the analysis presented here. The number of genes being altered, and pathways trigged, indicate that the high BaP concentration generated a very complex gene response, possibly through a complex induction network. The binding of the AhR to the binding site leads to the induction of certain transcription factors that will further induce other genes and transcription factors creating a cascade of response, and thus increasing the number of genes being regulated. van Delft et al (2010) using HepG2 cells exposed to 3 µM of BaP for 8 and 24 h showed that 30 significant networks were centred around the AhR transcription factor. This cascade of regulation together with cross talks between AhR and other TFs makes the BaP signalling a very complex system.

The transcriptome response was analysed in this study under conditions where no measurable impact on cell viability was observed for the low BaP concentration (50 nM) while for the high BaP concentration (5 µM), cell viability was significantly impaired. Compared to other cell lines, the Hepa1c1c7 murine hepatocarcinoma cell line appears to be particularly responsive to BaP. For comparison, no impact on cell viability on exposure to 5 µM BaP for at least 48 h was observed for the human breast carcinoma cell line, MCF-7, and the human hepatocarcinoma cell line, HepG2 (HOCKLEY et al., 2006). Other differences can be seen in the sensitivity of the different cell based systems used. Landolph et al (1976) showed that when BaP metabolic enzymes were inhibited in NMuLi cells, the cytotoxic effect was reduced. Others have shown a positive correlation between cytotoxic effect and induction of enzymes involved in BaP metabolism (BARTHOLOMEW et al., 1975; BABICH et al., 1988; LEE et al., 1993; PANDEY et al., 2006). Also, Landolph et al (1976) showed that cells in exponential phase produce more metabolic enzymes, what makes them more sensitive to BaP
than stationary phase cells. They too argue that the presence of more metabolic enzymes alone could not be responsible for the differential toxicity among growing and resting cells and further suggest that cellular division may also play an important role in the toxicity. Holme et al (2007), comparing Hepa1c1c7 cells to a rat liver epithelial F258 cell, showed that the rate and maximal capacity of metabolic activation was higher in Hepa1c1c7 than in F258 cells, due to higher induction of \textit{cyp1a1}. Also, they claim that the ROS production in F258 cells is an initial event while in Hepa1c1c7 cells is a consequence of toxicity. Hamouchene et al (2011) linked cell cycle phase in MCF-7 cells (present at S or G2/M) with an increase in DNA adduct formation. Hockley et al (2006) showed that MCF-7 cell viability reduced by 60\%, compared to control, after 96 h exposed to 1 to 5 µM BaP while HepG2 reduced by less than 40\% and only to 2.5 and 5 µM.

In conclusion, this study demonstrates that the genes being regulated within the first hours of exposure of cells to BaP are similar irrespective of a 100-fold exposure concentration. Moreover, the study reveals that the 4 h time point is decisive for the cell fate. Depending on the chemical concentration, the cells can recover from exposure or enter a complex response, both being seen in the transcriptional response as well as in the cell phenotype. Previously, Michaelson et al (2011), by comparing gene expression of AhR targets in Hepa1c1c7 and tao BpRc1 cells, deficient in endogenous AhR, exposed to BaP, TCDD and BPDE, revealed that genes that are direct AhR targets were induced as soon as 1 h after the start of treatment in Hepa1c1c7 cells, while there was no significant induction compared to vehicle control samples detectable in tao BpRc1 cells up to 4 h after exposure. At this time point genes involved in xenobiotic biotransformation, \textit{cyp1a1}, \textit{nqo1} and \textit{cpox} are regulated. Cyp enzymes are fully active at 24 h after BaP addition and metabolism is taking place (Fig 3).

The reduction in proliferation (small increase in doubling time) at the low BaP concentration indicates an arrest in cell cycle progression, apparently due to repair of DNA, since Cyp1a is active and transcription is regulated early on. At the high BaP concentration, the cells also enter arrest, with a very strong increase in doubling time, to repair DNA damage (12 and 24 h show a regulation of genes related to the NER pathway), however, the defence is overruled and cells render into death.
2.5 References


Bebenek, I. G.; Solaimani, P.; Bui, P.; Hankinson, O., CYP2S1 is negatively regulated by corticosteroids in human cell lines. *Toxicology Letters* 2011, 209 (1), 30-34.


Deb, S.; Bandiera, S. M., Characterization of a new cytochrome P450 enzyme, CYP2S1, in rats: Its regulation by aryl hydrocarbon receptor agonists. Toxicology 2010, 267 (1-3), 91-98.


Diani-Moore, S.; Ram, P.; Li, X.; Mondal, P.; Youn, D. Y.; Sauve, A. A.; Rifkind, A. B., Identification of the aryl hydrocarbon receptor target gene TiPARP as a mediator of suppression of hepatic gluconeogenesis by 2,3,7,8-tetrachlorodibenzo-p-dioxin and of nicotinamide as a corrective agent for this effect. Journal of Biological Chemistry 2010, 285 (50), 38801-38810.


Holmes, J. L.; Pollenz, R. S., Determination of aryl hydrocarbon receptor nuclear translocator protein concentration and subcellular localization in hepatic and nonhepatic cell culture lines: development of quantitative Western blotting protocols for calculation of aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein in total cell lysates. *Mol Pharmacol* 1997a, 52 (2), 202-11.

Holmes, J. L.; Pollenz, R. S., Determination of aryl hydrocarbon receptor nuclear translocator protein concentration and subcellular localization in hepatic and nonhepatic cell culture lines: Development of quantitative western blotting protocols for calculation of aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein in total cell lysates. *Molecular Pharmacology* 1997b, 52 (2), 202-211.


Jelski, W.; Szmitkowski, M., Alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) in the cancer diseases. *Clinica Chimica Acta* **2008**, *395* (1-2), 1-5.


Matsuo, R.; Asada, A.; Fujitani, K.; Inokuchi, K., LIRF, a gene induced during hippocampal long-term potentiation as an immediate-early gene, encodes a novel RING finger...


Pfaffl, M. W.; Horgan, G. W.; Dempfle, L., Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* **2002**, 30 (9), e36.


2.6 Support information

Figure S1: Cell number over time and in different BaP concentrations. A- Control (no chemical), B- 0.5 nM, C- 5 nM and D- 500 nM; (n=3). Asterisks indicate significant difference as compared to the 2 h time point as determined by one-way ANOVA followed by Dunnett's Multiple Comparison Test (**p-value < 0.01).
Chapter 2: Support Information

Figure S2: Cyp1a enzymatic activity, measured as 7-ethoxyresorufin-O-deethylase (EROD) activity in Hepa1c1c7 cells exposed to BaP ranging in concentration from 0.1 nM to 12.5 nM. Experiments were performed in biological triplicates and average and standard deviations are shown. Asterisks indicate significant difference as compared to the DMSO control as determined by one-way ANOVA followed by Dunnett's Multiple Comparison Test ** p-value<0.01. A- 2 h, B- 4 h, C- 12 h and D- 24 h of exposure. Solid lines represent Cyp1a basal activity and dash lines reflect the confidence intervals.
Figure S3. Reactive oxygen species (ROS) quantification in Hepa1c1c7 cells exposed to BaP, ranging in concentration from 0.5 nM to 5 µM, using H2DFFDA as indicator. Quantification of fluorescence was normalized to total protein. Experiments were performed in biological triplicates and average and standard deviations are shown. Asterisks indicate significant difference as compared to the DMSO control as determined by one-way ANOVA followed by Dunnett's Multiple Comparison Test **p-value<0.01; A- 4h, B- 8 h and C- 2, 4, 8, 12, 24 h control cells and cells exposed to TBH for 2 h.
Table S2. Gene expression of selected genes over time by real time PCR.

<table>
<thead>
<tr>
<th>Gene</th>
<th>50μM</th>
<th>2h</th>
<th>4h</th>
<th>5μM</th>
<th>2h</th>
<th>4h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Microarray</td>
<td>qPCR</td>
<td>Microarray</td>
<td>qPCR</td>
<td>Microarray</td>
<td>qPCR</td>
</tr>
<tr>
<td></td>
<td>Log2 ratio</td>
<td>p-value</td>
<td>Fold change</td>
<td>p-value</td>
<td>Log2 ratio</td>
<td>p-value</td>
</tr>
<tr>
<td>Mpp2</td>
<td>1.09</td>
<td>3.51e-05</td>
<td>3.79‡</td>
<td>&lt;0.0001</td>
<td>1.57</td>
<td>3.83e-05</td>
</tr>
<tr>
<td>MyoI</td>
<td>-</td>
<td>-</td>
<td>2.05</td>
<td>0.369</td>
<td>0.89</td>
<td>0.000282</td>
</tr>
<tr>
<td>Nsf11c</td>
<td>-</td>
<td>-</td>
<td>0.98</td>
<td>0.836</td>
<td>0.83</td>
<td>0.000234</td>
</tr>
<tr>
<td>Nqo1</td>
<td>-</td>
<td>-</td>
<td>2.57</td>
<td>0.069</td>
<td>1.19</td>
<td>1.96e-05</td>
</tr>
<tr>
<td>Adh7</td>
<td>-</td>
<td>-</td>
<td>1.71</td>
<td>0.181</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Adipor2</td>
<td>-</td>
<td>-</td>
<td>1.29</td>
<td>0.472</td>
<td>0.74</td>
<td>0.0002</td>
</tr>
<tr>
<td>Cpox</td>
<td>-</td>
<td>-</td>
<td>2.18</td>
<td>0.096</td>
<td>1.3</td>
<td>0.00021</td>
</tr>
<tr>
<td>Cyp1a1</td>
<td>2.86</td>
<td>2.73e-06</td>
<td>22.24‡</td>
<td>&lt;0.0001</td>
<td>3.16</td>
<td>2.82e-07</td>
</tr>
<tr>
<td>Oaspl</td>
<td>-</td>
<td>-</td>
<td>2.43‡</td>
<td>&lt;0.0001</td>
<td>1.3</td>
<td>1.94e-05</td>
</tr>
<tr>
<td>Tiparp</td>
<td>2.76</td>
<td>2.01e-05</td>
<td>9.01‡</td>
<td>&lt;0.0001</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tmem71</td>
<td>-</td>
<td>-</td>
<td>1.01</td>
<td>0.984</td>
<td>0.83</td>
<td>0.00022</td>
</tr>
</tbody>
</table>

qPCR fold changes values and p-values (n=3; p<0.05) were calculated by REST software. ‡ confirmation of up regulation.
<table>
<thead>
<tr>
<th>Gene name</th>
<th>Forward primer</th>
<th>Reverse primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyp1a1</td>
<td>TACCTGCCTAACTCTTTCC</td>
<td>ATGTCCTGTAGTGCTTT</td>
</tr>
<tr>
<td>Nqo1</td>
<td>TTCTCTGCGCCGATTCAGAT</td>
<td>GGCTGCTTTGGAGCAAATAG</td>
</tr>
<tr>
<td>Adh7</td>
<td>GCTGTCCTATGGGGAGTGAA</td>
<td>CACTCTCCACAACCCAACT</td>
</tr>
<tr>
<td>Cpox</td>
<td>CCAATCGCTTTTACTTGC</td>
<td>GGCGAGACGTGTTCAGTAA</td>
</tr>
<tr>
<td>Tiparp</td>
<td>AACTACGAAGGCTGTCTA</td>
<td>GATACTCTCTCCATCCAAG</td>
</tr>
<tr>
<td>Mpp2</td>
<td>TGTCCTTTACGTGTGTCCA</td>
<td>TCGAAGTAATGCCCATAGCC</td>
</tr>
<tr>
<td>Osbpl</td>
<td>GAGGCCCCTCAGAACTCCACAG</td>
<td>CCACAGGACACTCCAGGAT</td>
</tr>
<tr>
<td>Myof</td>
<td>AGTGCTGGCTGAAGGTTTC</td>
<td>TTTACAGACCACGGACCCCTA</td>
</tr>
<tr>
<td>Nsfl1c</td>
<td>CCAAGCTAGCTCGTCCATC</td>
<td>CAGCCAGCTTTGTAGGG</td>
</tr>
<tr>
<td>Tmem71</td>
<td>AGTGGAGATCTCTGGCTGGA</td>
<td>TCTTTCGAAAGTGGGCTGAT</td>
</tr>
<tr>
<td>Adipor2</td>
<td>GGATGTGGAAGTCGTGTG</td>
<td>ACCTGGTCAAACGAGACACC</td>
</tr>
</tbody>
</table>
### Table S3. David annotation.

**12 h of 5 µM BaP**

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>PValue</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0031981--nuclear lumen</td>
<td>3.02E-19</td>
<td>4.29E-16</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0031974--membrane-enclosed lumen</td>
<td>3.20E-18</td>
<td>4.55E-15</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0070013--intracellular organelle lumen</td>
<td>1.79E-17</td>
<td>2.54E-14</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0043233--organelle lumen</td>
<td>2.22E-17</td>
<td>3.16E-14</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0005654--nucleoplasm</td>
<td>1.25E-12</td>
<td>1.78E-09</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0007049--cell cycle</td>
<td>8.15E-12</td>
<td>1.45E-08</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0044451--nucleoplasm part</td>
<td>5.58E-11</td>
<td>7.93E-08</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0006793--phosphorus metabolic process</td>
<td>4.60E-09</td>
<td>8.16E-06</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0006796--phosphate metabolic process</td>
<td>4.60E-09</td>
<td>8.16E-06</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0043228--non-membrane-bounded organelle</td>
<td>5.49E-09</td>
<td>7.80E-06</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0043232--intracellular non-membrane-bounded organelle</td>
<td>5.49E-09</td>
<td>7.80E-06</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0005667--transcription factor complex</td>
<td>5.34E-08</td>
<td>7.59E-05</td>
</tr>
<tr>
<td>KEGG_PATHWAY</td>
<td>mmu04110:Cell cycle</td>
<td>1.53E-07</td>
<td>1.86E-04</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0016310--phosphorylation</td>
<td>1.82E-07</td>
<td>3.24E-04</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0044427--chromosomal part</td>
<td>2.67E-07</td>
<td>3.79E-04</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0022402--cell cycle process</td>
<td>3.49E-07</td>
<td>6.20E-04</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>GO:0008134--transcription factor binding</td>
<td>4.16E-07</td>
<td>6.44E-04</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0031967--organelle envelope</td>
<td>5.79E-07</td>
<td>8.23E-04</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0031975--envelope</td>
<td>6.57E-07</td>
<td>9.34E-04</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0005694--chromosome</td>
<td>6.96E-07</td>
<td>9.89E-04</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0005730--nucleolus</td>
<td>9.82E-07</td>
<td>0.00139639</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0033554--cellular response to stress</td>
<td>4.17E-06</td>
<td>0.00740252</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0006357--regulation of transcription from RNA polymerase II promoter</td>
<td>5.39E-06</td>
<td>0.00956228</td>
</tr>
<tr>
<td>GOTERM_MF_FAT</td>
<td>GO:0004672--protein kinase activity</td>
<td>5.81E-06</td>
<td>0.00900367</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0005819--spindle</td>
<td>6.43E-06</td>
<td>0.0091335</td>
</tr>
<tr>
<td>GOTERM_CC_FAT</td>
<td>GO:0000775--chromosome, centromeric region</td>
<td>8.45E-06</td>
<td>0.01201548</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0045941--positive regulation of transcription</td>
<td>8.59E-06</td>
<td>0.0152434</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0010628--positive regulation of gene expression</td>
<td>9.00E-06</td>
<td>0.01596137</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0006468--protein amino acid phosphorylation</td>
<td>9.81E-06</td>
<td>0.01741465</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0045944--positive regulation of transcription from RNA polymerase II promoter</td>
<td>1.02E-05</td>
<td>0.01813898</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0051254--positive regulation of RNA metabolic process</td>
<td>1.14E-05</td>
<td>0.02031532</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0051301--cell division</td>
<td>1.31E-05</td>
<td>0.02327879</td>
</tr>
<tr>
<td>GOTERM_BP_FAT</td>
<td>GO:0045935--positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process</td>
<td>1.58E-05</td>
<td>0.02799733</td>
</tr>
</tbody>
</table>
### Chapter 2: Support Information

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Description</th>
<th>Enrichment Score</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GOTERM_CC_FAT</strong></td>
<td>GO:0000793</td>
<td>condensed chromosomal condensation</td>
<td>1.86E-05</td>
<td>0.02641504</td>
</tr>
<tr>
<td><strong>KEGG_PATHWAY</strong></td>
<td>mmu05016</td>
<td>Huntington's disease</td>
<td>1.88E-05</td>
<td>0.02293016</td>
</tr>
<tr>
<td><strong>GOTERM_BP_FAT</strong></td>
<td>GO:0051173</td>
<td>positive regulation of nitrogen compound metabolic process</td>
<td>1.93E-05</td>
<td>0.0342563</td>
</tr>
<tr>
<td></td>
<td>GO:0045893</td>
<td>positive regulation of transcription, DNA-dependent</td>
<td>2.02E-05</td>
<td>0.03589541</td>
</tr>
<tr>
<td></td>
<td>GO:0010557</td>
<td>positive regulation of macromolecule biosynthetic process</td>
<td>2.37E-05</td>
<td>0.04211239</td>
</tr>
</tbody>
</table>

**24 h of 5 µM BaP**

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Description</th>
<th>Enrichment Score</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GOTERM_CC_FAT</strong></td>
<td>GO:0031974</td>
<td>membrane-enclosed lumen</td>
<td>9.91E-15</td>
<td>1.42E-11</td>
</tr>
<tr>
<td></td>
<td>GO:0070013</td>
<td>intracellular organelle lumen</td>
<td>2.39E-14</td>
<td>3.46E-11</td>
</tr>
<tr>
<td></td>
<td>GO:0043233</td>
<td>organelle lumen</td>
<td>3.11E-14</td>
<td>4.46E-11</td>
</tr>
<tr>
<td></td>
<td>GO:0031967</td>
<td>organelle envelope</td>
<td>1.80E-10</td>
<td>2.58E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0031975</td>
<td>envelope</td>
<td>2.19E-10</td>
<td>3.14E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0005739</td>
<td>mitochondrial lumen</td>
<td>2.64E-10</td>
<td>3.79E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0031981</td>
<td>nuclear lumen</td>
<td>4.06E-10</td>
<td>5.82E-07</td>
</tr>
<tr>
<td></td>
<td>GO:0055114</td>
<td>oxidation reduction</td>
<td>1.34E-08</td>
<td>2.42E-05</td>
</tr>
<tr>
<td></td>
<td>GO:0005764</td>
<td>lysosome</td>
<td>9.73E-07</td>
<td>0.00139659</td>
</tr>
<tr>
<td></td>
<td>GO:0000323</td>
<td>lytic vacuole</td>
<td>1.12E-06</td>
<td>0.00161153</td>
</tr>
<tr>
<td></td>
<td>GO:0005654</td>
<td>nucleoplasm</td>
<td>1.60E-07</td>
<td>2.30E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0005783</td>
<td>endoplasmic reticulum</td>
<td>1.25E-05</td>
<td>0.02252508</td>
</tr>
<tr>
<td></td>
<td>GO:0044429</td>
<td>mitochondrial part</td>
<td>3.51E-07</td>
<td>5.03E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0044451</td>
<td>nucleoplasm part</td>
<td>5.66E-07</td>
<td>8.12E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0008134</td>
<td>transcription factor binding</td>
<td>7.28E-07</td>
<td>0.00116439</td>
</tr>
<tr>
<td></td>
<td>GO:0005764</td>
<td>lysosome</td>
<td>9.73E-07</td>
<td>0.00139659</td>
</tr>
<tr>
<td></td>
<td>GO:0000323</td>
<td>lytic vacuole</td>
<td>1.12E-06</td>
<td>0.00161153</td>
</tr>
<tr>
<td></td>
<td>GO:0005654</td>
<td>nucleoplasm</td>
<td>1.60E-07</td>
<td>2.30E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0005783</td>
<td>endoplasmic reticulum</td>
<td>1.25E-05</td>
<td>0.02252508</td>
</tr>
<tr>
<td></td>
<td>GO:0044429</td>
<td>mitochondrial part</td>
<td>3.51E-07</td>
<td>5.03E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0044451</td>
<td>nucleoplasm part</td>
<td>5.66E-07</td>
<td>8.12E-04</td>
</tr>
<tr>
<td></td>
<td>GO:0008134</td>
<td>transcription factor binding</td>
<td>7.28E-07</td>
<td>0.00116439</td>
</tr>
</tbody>
</table>

**GOTERM_MF_FAT**

<table>
<thead>
<tr>
<th>Category</th>
<th>Term</th>
<th>Description</th>
<th>Enrichment Score</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GO:0000166</td>
<td>nucleotide binding</td>
<td>2.02E-06</td>
<td>0.00323286</td>
</tr>
<tr>
<td></td>
<td>GO:0005829</td>
<td>cytosol</td>
<td>2.37E-06</td>
<td>0.00339694</td>
</tr>
<tr>
<td></td>
<td>GO:0007049</td>
<td>cell cycle</td>
<td>3.76E-06</td>
<td>0.00680239</td>
</tr>
<tr>
<td></td>
<td>GO:0007050</td>
<td>cell cycle arrest</td>
<td>6.87E-06</td>
<td>0.01242866</td>
</tr>
<tr>
<td></td>
<td>GO:0005502</td>
<td>cell cycle process</td>
<td>9.75E-06</td>
<td>0.01764164</td>
</tr>
<tr>
<td></td>
<td>mmu04115</td>
<td>p53 signaling pathway</td>
<td>1.24E-05</td>
<td>0.01537487</td>
</tr>
<tr>
<td></td>
<td>GO:0065003</td>
<td>macromolecular complex assembly</td>
<td>1.25E-05</td>
<td>0.02252508</td>
</tr>
<tr>
<td></td>
<td>GO:0016765</td>
<td>transferase activity, transferring alkyl or aryl (other than methyl) groups</td>
<td>3.04E-05</td>
<td>0.04860927</td>
</tr>
</tbody>
</table>
3. Characterization of tiPARP in BaP-exposed Hepa1c1c7 wild-type and tiPARP silenced cells
Chapter 3

Abstract

TCDD-inducible poly(ADP-ribose) polymerase (tiPARP) belongs to a family of proteins called poly (ADP-ribose) polymerase (PARP) that are thought to be involved in a number of cellular processes related to DNA repair and programmed cell death, with PARP-1 being the most well-known member of this family. tiPARP has first been described as a gene responsive to TCDD in Hepa1c1c7 cells, although studies, including the work presented in Chapter 2, have shown that this gene is responsive to several dioxin-like compounds under the regulation of the aryl hydrocarbon receptor (AhR). A number of activities have been proposed for tiPARP, however, none of those seems to be the main function of this protein, when taking in account the gene regulation and the poly(ADP-ribose) activity. In Chapter 2, the transcriptome of Hepa1c1c7 cells exposed to 50 nM and 5 μM BaP concentrations over 24 h was accessed and tiPARP was identified to be regulated rapidly and tightly, within the first 4 h while other PARP genes, PARP-1 and PARP-10, were regulated only at later time points. With regard to all available information on tiPARP and other PARP genes, it was hypothesized that tiPARP’s primary function was to protect the cells against damage induced by dioxin like compounds. To explore this hypothesis, phenotype end points were chosen to monitor the cellular viability of wild type and tiPARP silenced Hepa1c1c7 cells exposed and non-exposed to BaP. However, no alteration was observed for the time points tested. In order to confirm the reduction on tiPARP protein levels in silenced cells, endogenous tiPARP protein detection was attempted with western blot and global mass spectrometry based proteomics (MUDPIT). Nonetheless, these efforts remained unsuccessful, probably due to a tight regulation of tiPARP protein and subsequently insufficient levels of this protein to be detected by these methods. Other strategies need to be developed in order to detect this protein.
3.1 Introduction

TCDD-inducible poly(ADP-ribose) polymerase (tiPARP) was discovered a decade ago in a murine hepatoma cell (Hepa1c1c7) exposed to TCDD, using mRNA differential display (MA et al., 2001). tiPARP encodes for a protein of 657 amino acids, it consist of a zinc finger domain, a WWE domain (a putative protein-protein interaction motif that contains two conserved tryptophan residues and one glutamic acid residue) and a PARP catalytic domain, which shows homology to Parp-1 and tankyrase-1 (MEYER, 2006). Since its discovery, this gene has repeatedly been demonstrated to respond to dioxin exposure under the control of AhR. (ITO et al., 2006; DE WAARD et al., 2008; DIANI-MOORE et al., 2010; AUNG et al., 2011; BOUTROS et al., 2011; MICHAELSON et al., 2011; GUALTIERI et al., 2012). Several activities have been proposed for this protein. In accordance with its PARP catalytic domain, in vitro expressed murine tiPARP has been shown to have Parp-1 activity (MEYER, 2006) and recently tiPARP was proposed to mediate suppression of hepatic gluconeogenesis after TCDD treatment in chicken embryos (DIANI-MOORE et al., 2010). However, based on the knowledge of its transcriptional regulation and its poly (ADP-ribose) polymerase activity, the primary function of this protein has not yet been determined.

Poly (ADP-ribose) polymerase (PARP) is a family of proteins involved in a number of cellular processes that thus far have been described to relate to DNA repair and programmed cell death. The PARP family comprises 17 members (10 putative). The conservation in the PARP signature varies among the PARP-family members. The PARP domain is located at the C terminus of the protein, except in PARP-4, and is adjacent to various domains that are involved in DNA or RNA binding, protein-protein interactions or cell signalling (MEYER, 2006; SCHREIBER et al., 2006). A tentative classification of PARP-family members can be proposed according to their putative functional domains or established functions (Figure 3.1): DNA-dependent PARPs (PARP-1 and PARP-2); tankyrases (tankyrase-1 and tankyrase-2); CCCH-type PARPs (PARP-12, PARP-13 and tiPARP, also called PARP-7) and macroPARPs (PARP-9, PARP-14 and PARP-15) (MEYER, 2006; SCHREIBER et al., 2006).
Figure 3.1. The PARP superfamily (SCHREIBER et al., 2006).
As described in Chapter 2, the transcriptome of Hepa1c1c7 cells exposed to benzo(a)pyrene (BaP) in a time and concentration dependent manner was accessed by microarray. The transcriptome revealed a very distinct regulation among the regulated PARPs (tiPARP, PARP-1 and PARP-10, Figure 3.2).

**Figure 3.2.** Gene expression of tiPARP, Parp-1 and Parp-10 in Hepa1c1c7 cells exposed to BaP in the microarray A) 50 nM and B) 5 µM. The smaller graphs are the tiPARP expression over time by qPCR performed in independent samples from the ones used for the microarray (for primer sequences, see Table 2.1). It is worthwhile to notice that qPCR analysis is more sensitive than microarray analysis.
PARP-1 has been the most extensively studied member of the PARP-family. This protein is a molecular sensor of DNA strand breaks: it has a key role in the spatial and temporal organization of their repair. In response to DNA strand breaks, PARP-1 uses NAD$^+$ as a substrate to catalyse the covalent attachment of ADP-ribose units on the carboxyl group of acceptor proteins that are usually associated with DNA transactions or on PARP-1 itself (SCHREIBER et al., 2006; LIN and YANG, 2008). Michaelson et al (2011) indicate that PARP-1 is a secondary AhR target, responding to BPDE damage.

PARP-10 is a partner of the proto-oncoprotein c-Myc, a key transcriptional regulator that controls cell proliferation. PARP-10 contains an RNA recognition motif and a Gly-rich domain, both of which have also been found to mediate RNA binding in a c-Myc nucleolar partner protein, nucleolin. PARP-10 shuttles between the cytoplasm and the nucleus and accumulates within the nucleus where it acquires a CDK2-dependent phosphorylation during late-G1-S phase and during pro-metaphase to cytokinesis. PARP-10 is a potent inhibitor of the cell transformation that is mediated by c-Myc in the presence of Ha-Ras, but its PARP activity is not required for this function. PARP-10 was able to counteract some nuclear c-Myc functions that promote cellular transformation. Furthermore, PARP-10 poly(ADP-ribosyl)ates histone H2A, which indicates that it has a role in chromatin regulation (SCHREIBER et al., 2006).

Given the knowledge of the functions of PARP-1 and PARP-10, their regulation at high BaP concentrations at exposure times equal or more than 12 h underlines their likely role in DNA repair and cell cycle control. However, the question remains as to what the role of tiPARP is upon exposure to BaP. One way to understand the function of a gene is to observe a biological system that lacks that gene (KITTLER et al., 2004; LEUNG and WHITTAKER, 2005; KITTLER et al., 2007). Mammalian gene function has been determined traditionally by methods such as disruption of genes, the introduction of transgenes, the molecular characterization of human hereditary diseases, and targeting of genes by antisense or ribosomes techniques (ELBASHIR et al., 2002; FRIEDMAN and PERRIMON, 2004; GAMA SOSA et al., 2010). Working backward from the phenotype one can determine the gene responsible for that characteristic (LEUNG and WHITTAKER, 2005; GAO and ZHANG, 2007; STURZL et al., 2008).

Several techniques for gene knock-down were developed over the years with one of the recent ones being silencing RNA (ROZEMA and LEWIS, 2003; LEUNG and WHITTAKER, 2005; GAO and ZHANG, 2007; BOONANUNTANASARN, 2008; BOUTROS and AHRINGER, 2008; STURZL et al., 2008; KLEINHAMMER et al., 2011). This technology can be applied on wild-type
cells, thus circumventing potential difficulties of creating mutant strains. RNA interference aims to degrade mRNA from a specific gene (ROZEMA and LEWIS, 2003; LEUNG and WHITTAKER, 2005; GAO and ZHANG, 2007). This technology was developed by Andrew Fire and Craig Mello during their research in Caenorhabditis elegans (FIRE et al., 1998; SCHERER and ROSSI, 2003; KITTLER et al., 2007; BOUTROS and AHRINGER, 2008). It is based on the introduction of a double strand (ds)RNA with a complementary sequence to a gene of interest (ELBASHIR et al., 2002; ROZEMA and LEWIS, 2003; SCHERER and ROSSI, 2003; STURZL et al., 2008; KLEINHAMMER et al., 2011). This sequence is then recognized by the organism as being an exogenous genetic material and activates the siRNA pathway (ROZEMA and LEWIS, 2003; SCHERER and ROSSI, 2003; LEUNG and WHITTAKER, 2005; GAO and ZHANG, 2007; KITTLER et al., 2007; BOUTROS and AHRINGER, 2008; KLEINHAMMER et al., 2011). In this way, expression of a target gene is decreased, making possible the study of its physiologic role (ROZEMA and LEWIS, 2003; SCHERER and ROSSI, 2003; GAO and ZHANG, 2007; BOUTROS and AHRINGER, 2008).

Taking into account all available information about the PARP genes (Figure 3.3) and their expression under the here applied BaP exposure conditions, it was hypothesized that tiPARP function is to protect cells against damage induced by the presence of dioxin like compounds, e.g. BaP. To study the role of tiPARP, siRNA experiments were designed and phenotype end points were tested. To confirm if tiPARP levels were also reduced on the protein level, detection of tiPARP protein was attempted in different ways but remained unsuccessful.

**Figure 3.3.** Metabolism of poly(ADP-ribose) during DNA damage and repair (SCHREIBER et al., 2006). PARG, Poly(ADP-ribose)glycohydrolase is the major enzyme responsible for the catabolism of poly(ADP-ribose).
Chapter 3

3.2 Material and Methods

Chemical

Benzo(a)pyrene (≥ 96 % purity, Sigma Aldrich, Switzerland) was pre-dissolved in Dimethyl sulfoxide (DMSO, ≥ 99.9 % purity, Sigma Aldrich, Switzerland) in a stock solution with a concentration of 10 mM and 0.1 mM.

Cell cultivation

The mouse hepatoma Hepa1c1c7 cell line was grown in Dulbecco’s modified medium without phenol red (Biochrom AG, Berlin) and complemented with 7 % of fetal calf serum (Biochrom AG), 1 % of penicillin/streptomycin (Sigma Aldrich) and 2 mM of L-alanine and L-glutamine (Biochrom AG, Berlin).

Cell Viability assay

For the determination of the sensitivity of tiPARP silenced Hepa1c1c7 cells to BaP (vessels marked with sitiPARP 50nM or sitiPARP 5µM), a set of cells were used to measure cell viability, using different fluorescent dyes (SCHIRMER et al., 1998). AlamarBlue (Invitrogen, Basel, Switzerland) was used as a measure of metabolic activity, 5-carboxyfluorecein diacetate acetoxyethyl ester (CFDA-AM, Invitrogen, Basel, Switzerland) as an indicative of membrane permeability, and neutral red (Invitrogen, Basel, Switzerland) as a measure of lysosome membrane integrity. After the exposure time, the cells were washed with 1 mL of PBS (Phosphate buffered saline, 2.68 mM KCl, 1.47 mM KH2PO4, 136.89 mM NaCl and 8.10 mM NaHPO4) and incubated for 1 hour with a solution of PBS with AlamarBlue (5 % v/v) and CFDA-AM (4 µM). Thereafter the fluorescence was measured using a fluorescence plate reader (TECAN, Switzerland) with excitation/emission 530/595 nm and 493/541 nm. The solution was then discarded and the cells were again washed and incubated with PBS with neutral red (1.5 % v/v) for another hour. After this incubation period the cells were washed with a fixative solution (45 mM CaCl2 and 37 % formaldehyde) and incubated for 30 minutes with an extraction solution (1 % acetic acid and 50 % ethanol). Then the fluorescence was measured also using the fluorescence plate reader at excitation/emission 530/645 nm (SCHIRMER et al., 1998). All the measurements were made with technical and biological triplicates. The results were expressed as percentage of the appropriate control (unexposed cells).
Small interfering RNAs transfection

Small interfering RNAs (siRNA) of tiPARP (NM_178892), cyclophilin B (D-001820-02-05) and non-targeting control siRNA (D-001810-10-05) were obtained from ON-TARGETplus SMARTpool of Dharmacon Research (Thermo Scientific, United Kingdom). Cyclophilin B siRNA probes were used as positive control. Each siRNA set was designed to target four different regions of the specific gene. Cells were transfected with the siRNA at a concentration of 25 nM using the DharmaFECT siRNA transfection reagent in an antibiotic-free medium.

A cell density of 10⁴ cells cm⁻² was used in a 12 well plate (Greiner, Switzerland) containing 1.5 mL of cell culture medium per well without antibiotics throughout the whole experiment. The cells were seeded on the first day, one day after the medium was exchanged and the transfection reagents were added, and two day later the cells were exposed to the chemical, BaP in DMSO. The final concentration of the vehicle in the culture medium was 0.05 % (v/v). BaP was used in two different final concentration, 50 nM and 5 µM, and the time of exposure was 24 h. The cells were then used for RNA extraction or to determine cell viability.

qPCR

After the experimental procedure cells were harvested using Trizol (Invitrogen) and RNA was extracted using RNA easy kit (Qiagen, Valencia, CA, USA). RNA was quantified using nanodrop (ThermoScientific, Switzerland). cDNA synthesis was performed according to manufacture descriptions of TaqMan®, Reverse Transcription Reagents (Applied Biosystems, Switzerland), using a cycle of 25°C for 5min, 48°C for 30 minutes, 95°C for 10 min and 4°C for 1min. For the real time amplification a SyBR Green® PCR master mix (Applied Biosystems, Switzerland) was used and the cycle used was 95°C for 10 min and 94°C for 15 seconds for 40 cycles, and 60°C for 1 minute for elongation. The primers used and their efficiency are shown in Table 2.1. qPCR expression levels relative to untreated controls were calculated according to Muller et al (2002). The advantage of this calculation is that it takes varying PCR amplification efficiencies into account, according to: MNE = ((Eref)Ctref,mean) / ((Etarget)Cttarget,mean), where MNE stands for mean normalized expression; Eref is the amplification efficiency for the reference gene; Etarget is target gene amplification efficiency; Ctref, mean is mean Ct value for the reference gene; and Cttarget, mean is the mean Ct value for the target gene. Expression data for target genes were presented relative to the reference gene.
Table 2.1. qPCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward</th>
<th>Reverse</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>tiPARP</td>
<td>AACTACGAAGGCTGTCTA</td>
<td>GATACTCTCTCCATCCAAAG</td>
<td>99 %</td>
</tr>
<tr>
<td>cyclophilinB</td>
<td>TCGCCATTGCAAGGAGTGAGA</td>
<td>AACCTTGACTGGCTACCTTCGT</td>
<td>89 %</td>
</tr>
<tr>
<td>actinB</td>
<td>GATGGCACAGGAGGAAAGAG</td>
<td>CCAGGAGGTCTTGACTGTGG</td>
<td>100 %</td>
</tr>
</tbody>
</table>

**Protein analysis**

**Western blot:** Cells untreated and exposed for 24 h to 50 nM or 5 µM BaP were submitted to protein extraction with 150 mM NaCl, 10 mM Tris-HCl (pH 7.4), 5 mM EDTA (pH 8), 1 % tritonX100, 5 M DTT and 69 mM PMSF. The homogenates were centrifuged (15 min, 10000 rpm at 4°C) and the protein content of the supernatant was determined by Pierce® BCA Protein Assay Kit (Themofisher, Switzerland). Aliquots of supernatant (100 µg/µL total protein) were boiled for 5 min in sample buffer (0.2 M Tris–HCl buffer, 10 % glycerol, 2 % SDS, 0.02 % β-mercaptoethanol). Equal amounts of protein from each of the samples were resolved by sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS–PAGE) on 10 % gel and transferred electrophoretically overnight at 4°C to a nitrocellulose membrane (3MM Whatman, Switzerland). The membrane was blocked with TBSTM (20 mM Tris–HCl; 500 mM NaCl; 0.05 % Tween20; 5 % w/v skimmed milk powder) at room temperature for 1 h. The membrane was then incubated with respective primary antibodies (rabbit monoclonal against tiPARP, Abgent, United Kingdom) for 2 h at room temperature. The membrane was then washed two times with TBSTM and incubated with goat anti-rabbit antibody coupled with biotin and horseradish peroxidase-conjugated secondary antibody (Abgent, United Kingdom) for 1 h. The membrane was then washed again three times with TBSTM. Protein antibody complexes were detected by the addition of CN/DAB substrate kit (Thermofisher, Switzerland).

**Global protein analysis by Multidimensional Protein Identification Technology (MudPIT):**

**Protein extraction and preparation of tryptic digests:** Cells were exposed to BaP for 24 h, harvested as described above and frozen at -80°C until further processing, which was done by Dr. Ksenia Groh (Eawag-Department of Environmental Toxicology). The proteins were extracted by placing the cells in the lysis buffer (9 M urea, 2 M thiourea, 0.1 M Tris-HCl, 1 % CHAPS, Protease Inhibitor Cocktail 1x (Sigma), homogenizing the suspension by thorough pipetting and sonicating it on ice (three successive 15s bursts with 30s pauses; Labsonic M, Sartorius Stedim Biotech, Goettingen, Germany). The supernatant was recovered after centrifugation for 15 min at 4° C. To prepare the samples for MudPIT analysis, the
proteins were precipitated from the supernatant using a conventional methanol-chloroform method (WESSEL and FLUGGE, 1984), air-dried for 5 min and redissolved in resolubilization buffer (9 M urea, 2 M thiourea, 50 mM Tris–HCl, pH 8.5). To improve the solubility, the pellet was wetted with 0.2 M NaOH before addition of resolubilization buffer. The protein concentration was measured by Pierce assay with BSA as a standard. Following recommended procedures (WASHBURN, 2008), 100 µg of proteins was reduced with TCEP, alkylated (carboxamidomethylated) with IAA, and digested with trypsin (1:100) (overnight incubation at 37 °C). The reaction was quenched with formic acid.

2D-LC separation and mass spectrometry: The analyses were performed essentially as described previously (GROH et al., 2011; NESTLER et al., 2012). A three-phase MudPIT column was made in-house from a fused silica capillary (ID 100 lm, OD 375 lm) drawn to a fine tip with a needle puller (Sutter-2000, Science Products AG, Basel) and packed successively with C18 3 µm (analytical column), SCX on 5 µm beads, and C18 5 µm (precolumn) resins by using a pressurized custom-made filling unit. Following equilibration, the peptides were loaded and desalted on a column, which was then placed in-line between an HPLC system and a nano-electrospray ionization source on the LTQ-Orbitrap XL (Thermo Scientific, Bremen, Germany). The elution proceeded using a standard 11-step MudPIT protocol (DELAHUNTY and YATES, 2005). The machine was operated at 1.5 kV spray voltage in positive ion mode with the ion transfer capillary temperature set to 200 °C and the tube lens set to 110 V. One full scan FT mass spectrum (400–1600 m/z, resolution of 60’000) was followed by seven data-dependent MS/MS scans (triggered by the most intense peaks carrying 2 or 3 positive charges) acquired in the linear ion trap with normalized collision energy (setting of 35 %). This procedure was continuously repeated throughout each MudPIT step. Dynamic exclusion was activated for 120 s, and the automatic gain control (FT) was set to 10⁶.

Analysis of LC-MS/MS data: Automation of data processing and analysis was performed using AutoHotkey scripts (http://www.autohotkey.com/) written in-house. Peak listing from MS raw files was done with extract_msn.exe supplied with the XCalibur package (Thermo Fisher Scientific) followed by merging of the MS/MS peak lists to mgf format by the PERL script merge.pl (Matrix Science, Boston, MA, USA). The data were then searched for peptide hits with the OMSSA (Open Mass Spectrometry Search Algorithm), version 2.1.7 (GEER et al., 2004) against the RefSeq database of mouse (Mus musculus) protein sequences.
downloaded from NCBI and curated in-house. Redundant protein entries were partially removed by the software CD-HIT (Li and Godzik, 2006) with the threshold set to 0.85. Further, the database was supplemented with the sequences of common contaminants originating from sample processing, including trypsin and highly abundant human proteins such as keratins. A database of reversed sequences was then produced by the decoy.pl PERL script (Matrix Sciences, Boston, MA, USA) and added to the original database. The observed number of reversed sequences found among identifications was used to calculate false discovery rate (FDR) based on the target-decoy database approach (Elias and Gygi, 2010).

During database search, carbamidomethyl-Cys was set as a fixed modification. Variable modifications were oxidation of Met, deamidation of Asn and Gln, protein N-terminal acetylation and cyclization of N-terminal Gln to pyroGlu. The charge of precursors to be searched was 2 and 3, and minimum precursor charge to consider multiply charged products was set to 2. Tryptic specificity was set to a maximum of two missed cleavages. Product ion m/z tolerance was set to 0.3 Da. The precursor ion mass accuracy was initially set to 0.05 Da and finally to 0.005 Da after removal of systematic mass errors as done previously (Zubarev and Mann, 2007; Groh et al., 2011). For individual MS/MS spectra only the peptide hits found with best confidence (lowest e-value) and only discrete hits with peptide FDR less than 0.5 % were considered for further statistical analysis. Spectral counts were calculated from the number of discrete peptides observed for each identified protein (Wong and Cagney, 2010). Subsequently, two-column count tables were created for the control and exposure conditions. Count pairs having less than two peptides in the control and exposure condition were removed from the list. The resulting lists for control and exposures were analyzed by G-test as described previously (Zhang et al., 2006). The statistical platform R (http://www.r-project.org/) was used to calculate the p-values, and Benjamini-Hochberg FDR control (<0.05) was applied to account for multiple testing (Benjamini and Hochberg, 1995). For each protein, the proportion of its counts in the total counts of all proteins was determined for all experimental conditions. The protein fold changes were then estimated by dividing this value in the exposure conditions by the corresponding value in the control condition. For this condition, 0 counts in either condition were substituted with the minimum observable value 1 to avoid calculating fold changes with 0.
3.3 Results

The experiments were carried out in three different steps. First, the viability of the cells was determined in the 12-well plate exposure set up chosen for the silencing experiments. The markers for cell membrane integrity (CFDA-AM), lysosomal membrane integrity (neutral red) and metabolic activity (alamar blue) were used to verify cell viability. The second step was the establishment of the silencing protocol and observing gene expression down-regulation and associated phenotypic responses in the cells. Finally, different techniques (western blot and MudPIT) were applied to attempt to detect the endogenous tiPARP protein.

1) Cell-based phenotype (Cell proliferation and viability)

Cell viability (cell membrane integrity, lysosomal membrane integrity and metabolic activity) was measured after 2, 4, 12 and 24 h of BaP (Chapter 2, Figure 2.1C). Only a reduction for metabolic activity and lysosomal membrane integrity (~40 %) was observed after 24h.

2) siRNA silencing

Two silencing protocols were used for observing cytotoxic effects of the transfection reagents. Among those protocols the amount of siRNA probes were kept constant (25 nM) while a variation of volumes of the transfection reagent were tested (0.1 to 0.5 µL). The cells were incubated for 48 h with the transfection reagents and microscopy observation led to the choice of the protocol that caused less than 20 % mortality (data not shown).

The final set up for the siRNA experiments with all the end points were arranged in four twelve plates, where two plates, with each treatment being replicated in three wells, were used for RNA extraction and the other two plates were used for the phenotype characterization (Figure 3.4).
Figure 3.4. siRNA set up. A. General overview. B. siRNA for 48 h with a 24 h of BaP induction. C. siRNA for 72 h with a 24 h BaP induction. Untreated are vessels with cells but not treated with silencing probes or BaP; cells treated with siRNA cyclophilinB and negative control vessels are positive and negative control for silencing experiments, respectively; cells treated with siRNA tiPARP, siRNA tiPARP 50nM and siRNA tiPARP 5µM are cells treated with silencing probes with BaP (50 nM and 5 µM) or without.
**mRNA-abundance:** Functionality of siRNA protocol was checked by using cyclophilinB siRNA probes as positive control. *cyclophilinB* gene expression was reduced approximately 85% (Figure 3.5). Non-target siRNA probes was used as negative control and had a variable effect on the expression of *tiPARP*, affecting more the tiPARP expression after 72 h of silencing. siRNA *tiPARP* probes were functional with a reduction of 75% of the gene expression after 48 h silencing (Figure 3.5A), but less efficient after 72 h (Figure 3.5B). However, when the silenced cells were challenged with BaP, the levels of *tiPARP* mRNA expression increased, indicating that the stimulation by BaP overcame the silencing.

![Figure 3.5. mRNA levels of silenced cells. First box, in either panel A or B, shows untreated cells and cells exposed to cyclophilinB and tiPARP; second box shows cells exposed to BaP, either 50 nM or 5 μM and tiPARP silenced cells exposed to 50 nM or 5 μM; third box shows cells treated with non-target silencing probes, either 50 nM or 5 μM BaP or no chemical; A) 48 h of silencing, following protocol described on Figure 3.4B and B) 72 h of silencing, following protocol described on Figure 3.4C.](image)

**Phenotype characterization:** Viability of cells under 48 or 72 h of silencing showed no difference when compared to untreated cells (control) (Figure 3.6 A and B). Also no change in cell viability was observed for any of the viability tests in tiPARP silenced cells exposed to 50 nM or 5 μM BaP for 24 h compared to non-silenced cells treated with the same concentrations and time (Figure 3.6 A).
Figure 3.6. Phenotype characterization of cells silenced for cyclophilinB, tiPARP or non-target siRNA probes and silenced and non-silenced cells exposed to 50 nM and 5 µM BaP for 24 h. A) cells silenced for 48 h following protocol described in figure 3.4B and B) cells silenced for 72 h following protocol described in Figure 3.4C, however, fewer treatments were used due to the high mRNA variability.
3) **Protein Detection**

**Western blot**: Despite many variations of the protocol (variation in antibody concentrations and batches), western blot was unable to reveal the tiPARP protein in untreated cells and cells exposed to 50 nM or 5 µM BaP (Figure 3.7).

![Western blot](image)

**Figure 3.7.** Western blot of cells untreated or treated for 24 h with 50 nM or 5 µM BaP. Primary antibody was also loaded to the gel to check if the secondary was functional (right lane).

**MudPIT**: Global mass spectrometry-based proteomics was used in an attempt to detect endogenous tiPARP in the BaP treated and untreated cells. However, tiPARP could not be detected in the samples tested.
3.4 Discussion

Based on the available information about PARP proteins and the effects of BaP in cells, the hypothesis was that the main function of tiPARP was to protect cells against the damage effect of dioxin-like compounds. tiPARP would then have the function of controlling the damage being formed by the chemical or the metabolites at the DNA and protein level. Small interference RNA experiments were then designed to prove this hypothesis and the outcome expected was an early loss of viability compared to the cells not treated with siRNA probes.

Variations on the silencing protocol were done to minimize the impact of the transfection reagents on cell viability. The levels expected of silencing were 75 % or more, according to the manufacturers protocols. This was indeed observed for cyclophilinB (positive control) with a reduction of 85 % and for tiPARP (target gene) with a reduction of 75 %. No increase in the loss of viability (Figure 3.6) was observed in the cells treated with tiPARP siRNA probes and BaP. One explanation for this results can be the restoration of tiPARP mRNA levels when cells were challenged with both concentrations of BaP for 24 h (Figure 3.5). The qPCR results show an increase in mRNA level when compared to siRNA treated cells (Figure 3.5).

The second possible explanation is that the amount of protein present in the cell was still able to deal with the damage caused by BaP. To test this hypothesis Hepa1c1c7 cells we treated with BaP for different times and both BaP concentrations and it was attempt to detect the protein using western blot and MudPit techniques. Unfortunately it seems that this protein is present in those cells in a quantity that western blot techniques and MudPIT did not detect. A collaborator from our group also tried to detect tiPARP by western blot using an antibody against human tiPARP, however without any success (Trump, S., personal communication).

Western blot detection limits generally lies between 1 to 500 ng of the protein, varying according to antibody specificities and the visualization system. MudPIT however has a much lower limit of detection ranging from femtomol to picomol. It has not been completely ruled out that the commercially available antibodies are not able to detect tiPARP in cellular extracts, being only demonstrated in mouse spleen tissue lysates. Another important factor for protein detection is the protein solubility. tiPARP seems to be insoluble upon overexpression with a probability of 0.78 (SCRATCH, http://www.ics.uci.edu/~baldig/scratch/index.html).
Furthermore, we have tried to monitor tiPARP by targeted proteomics using single reaction monitoring (SRM). Several peptide traces were seen with a tendency to an increased intensity in some cases, usually at shorter exposure durations (2 or 4 h). However, the biological replicates showed very high variation and the identity of the obtained signals could not be unambiguously validated with available instrumentation (Groh K., personal communication).

Recently, MacPherson, L. & Matthews, J. (2012, personal communication) detected endogenous tiPARP protein using immunofluorescence. They suggested that the inability to detect tiPARP using commercially available antibody may be due to a tight regulation of this protein. They also suggest that the principal role of tiPARP is to negatively regulate AHR-mediated signalling.
3.5 References


Diani-Moore, S.; Ram, P.; Li, X.; Mondal, P.; Youn, D. Y.; Sauve, A. A.; Rifkind, A. B., Identification of the aryl hydrocarbon receptor target gene TiPARP as a mediator of suppression of hepatic gluconeogenesis by 2,3,7,8- tetrachlorodibenzo-p-dioxin and of nicotinamide as a corrective agent for this effect. Journal of Biological Chemistry 2010, 285 (50), 38801-38810.


Schirmer, K.; Dixon, D. G.; Greenberg, B. M.; Bols, N. C., Ability of 16 priority PAHs to be directly cytotoxic to a cell line from the rainbow trout gill. *Toxicology* 1998, 127 (1-3), 129-141.


4. Quantification in a time- and concentration-dependent manner of BaP uptake and DNA adduct formation

To be submitted to Chemical Research in Toxicology (in preparation).
Abstract

The work in this Chapter aimed to reveal in a time- and concentration dependent manner the intra-cellular concentration of BaP in relation to the amount of DNA adducts. As shown in Chapter 2, exposing Hepa1c1c7 cells to two different BaP concentrations, namely 50 nM and 5 µM, and for different time points, ranging from 2 to 24 h, indicated rapid uptake of BaP into cells along with transcription of genes related to BaP metabolism, cell cycle arrest and/or repair. Given that no impact on cell viability was seen for the low BaP concentration, while severe damage and ROS formation was observed in cells exposed to the high concentration of BaP, the hypothesis was that DNA adduct formation, caused by Benzo(a)pyrene diol epoxide (BPDE) metabolites, outperformed DNA damage repair in the high but not the low BaP exposure concentration. BaP cellular concentration was found to be dependent on the initial concentration. The number of BaP molecules within one cell after 2 and 4 h ranged between 1.2 to 1.9 \(10^6\) and 1.2 to 1.7 \(10^9\) for 50 nM and 5 µM, respectively. A total of 4 to 7% of total BaP was found associated to the plastic fraction (cell culture flask), indicating that most of the chemical is present in the culture medium. More than 90% of BaP were completely metabolized after 24 h for both (50 nM and 5 µM) exposure concentrations. Rates of BaP uptake and biotransformation were calculated, showing that uptake is faster than biotransformation (0.17 and 0.15 h\(^{-1}\) for 50 nM and 0.12 and 0.1 h\(^{-1}\) for 5 µM BaP, respectively). DNA adduct quantification showed the presence of adducts as early as 8 h of BaP exposure. Both DNA adduct isomers, t(+)BPDEdGuo and t(-)BPDEdGuo, were detected in Hepa1c1c7 cells exposed to 50 nM and 5 µM BaP, with the first isomer being more abundant. DNA repair was seen at the 24 h time point in the cells exposed to 50 nM BaP, while repair mechanisms seemed to be overwhelmed at the 5 µM BaP and DNA adducts continued to accumulate over time. These results are in agreement with the transcriptome analysis, performed in Chapter 2, where genes related to nucleotide excision repair (NER) were regulated at the high (5µM) BaP concentration and later (12 and 24 h) time points. The NER pathway is responsible to remove bulky DNA adducts, as caused by BPDE. Moreover, the rate of adduct formation indicates that t(+)BPDE-dGuo adducts are formed with a higher preference than the others isomers. The formation preference of these adducts might be the answer why this isomer is considered to be the most carcinogenic.
4.1 Introduction

The ultimate goal of exposure assessment is to estimate the target dose or biological effective dose, meaning the actual concentration and the time of exposure to a determined stressor. However, measuring the actual concentration at the target site is usually not feasible. Measurements of internal concentrations of a chemical are the best available way to predict actual concentrations at target sites, however not many publications reported measured internal concentrations, rather nominal concentrations are generally used. Consideration of internal concentration allows for a better comparison of variable exposure conditions (ESCHER and HERMENS, 2002; ESCHER and HERMENS, 2004). This measurement is also advantageous when predicting toxicity from mixtures. Landrum et al (2003) showed that the sum of the internal concentrations of a mixture of polycyclic aromatic hydrocarbons (PAHs) were similar to the concentration of each PAHs separately.

Benzo(a)pyrene (BaP) is a well characterized PAH not manufactured or used commercially. It is primarily a by-product of incomplete combustion but also occurs naturally in petroleum-based tars. BaP was identified as being the major tumour-producing agent in coal tar in 1932 and is classified as carcinogenic to humans by the International Agency for Research on Cancer (IARC) (COOK et al., 1932; KENNAWAY, 1955; STRAIF et al., 2005). Ingestion of boiled, roasted, or smoked meat and fish, eggs and excessively fertilized vegetables, or smoking cigarettes, represent major uptake routes of BaP for humans (FINLAYSON-PITTS and PITTS, 1997; KNIZE et al., 1999; BOSTROM et al., 2002). BaP crosses the air-blood and intestine blood-barrier as it has been detected in blood and various tissues (MARIE et al., 2010).

Previous studies have investigated the uptake of BaP within a cell. Bartley et al (1982), studying mammary and fibroblast cells, determined that the initial uptake of BaP was relatively fast and then depended upon the rate of conversion to metabolites. Sureau et al (1990), using human T47D mammary tumour cells along with microspectrofluorimetric techniques, showed that BaP uptake of one cell reaches a plateau after 10 min and that this plateau is related to the relative amount of lipoprotein of the cell. Barhoumi et al (2000), observing rat liver cells by microscopy, also showed that BaP was rapidly (within minutes) taken up. Exposing cells for 16 h to 10 μM of BaP and using specific probes for organelles, Barhoumi et al (2000) found indications of extensive BaP localization in Golgi, and in cytoplasmic membranes but not within the nucleus, the mitochondrial matrix, or within the lysosomes. In contrast, Allison et al (1966) had reported on the identification of BaP, based on its autofluorescence, in the lysosomes.
BaP metabolism, however, was never observed in the first hours of BaP uptake by different cell lines using different techniques (STAMPFER et al., 1981; RUĐO et al., 1986; SUREAU et al., 1990; BARHOUMI et al., 2000). Ruđo et al (1986) demonstrated that, when cell density or incubation time increases, the percentage of BaP converted to water soluble metabolites increases in a linear fashion.

It is widely accepted that BaP requires biological activation through oxidative metabolism to be carcinogenic. BaP induces its own metabolism by activating the Aryl hydrocarbon Receptor (AhR) pathway. This pathway leads to the regulation of expression of several phase I and phase II enzymes (Figure 4.1). During phase I, cytochrome P450 monooxigenases, e.g. Cyp1a1 and Cyp1b1, epoxide reductases and epoxide hydrolases induce the monooxygenation of BaP which leads to formation of oxidized metabolites. During phase II, the BaP metabolites are conjugated to hydrophilic molecules such as glucuronic acid or glutathione, making them more hydrophilic and for that reason easier to be excreted (GELBOIN, 1980). After conjugation, elimination occurs via transport proteins, like certain ATP binding cassette transporters (ABC-transporters) (MILLER and RAMOS, 2001).

BaP metabolism overall leads to production of a number of metabolites, which generally fall into the categories of quinones, diols, phenols and epoxides (Figure 4.1). The ultimate carcinogenic compound appears to be the 7,8-dihydroxy-9,10-epoxi-7,8,9,10-tetrahydrobenzo(a)pyrene (BPDE), a product of epoxide hydrolase activity. It is assumed that the formation of BPDE is responsible for the generation of the majority of bulky DNA adducts in cells (LI et al., 1995; GEACINTOV et al., 1997; SHUKLA et al., 1997; KOZACK and LOECHLER, 1999; KOZACK et al., 1999; PENNING et al., 1999; BOYSEN and HECHT, 2003).
Figure 4.1: General BaP metabolic process comprising phase I ([cytochrome] P450s and EH, epoxide hydrolases, NQO1, NAD(P)H dehydrogenase [quinone] 1) and phase II (GST, glutathione transferases; UGT, UDP-glucuronoyl transferases; and SULT, sulfotransferases) enzymes generating DNA adducts and detoxified products.

DNA adducts may lead to cancer by causing mutations in genes essential for key functions, including apoptosis, proliferation and differentiation (Chl et al., 2009). Unrepaired DNA adducts induce permanent mutations and when those appear in crucial regions of oncogenes, such as the tumour suppressor genes and genes which regulate cell growth cancer will be caused (Skipper et al., 1994). However, cellular defence mechanisms counteract carcinogenesis at multiple levels. Likely, the type and amount of DNA lesions play a role for which mechanism is at play. Optimally, cells repair the damage and survive, having only a transitory cell cycle arrest. However, in case of unsuccessful repair, damage might be manifested in mutations. As a last resort, damaged cells may undergo cell death to prevent expansion of mutated clones (Schreck et al., 2009). Nucleotide-excision repair (NER) is one of the most versatile and important pathways by which mammalian cells remove a variety of DNA lesions, such as bulky BPDE adducts, whereas base excision repair (BER) is critically involved in the repair of damaged bases induced by reactive oxygen species, as well by a variety of other lesions (Figure 4.2) (Peterson and Cote, 2004; Hanawalt and Spivak, 2008; Yu et al., 2012).
The results obtained in Chapter 2 gave indication of both metabolism and DNA repair (NER). Indication of metabolism came from Cyp1a measurements and from microarray data that showed an increase in transcripts of genes related to BaP metabolism (cyp1a1, nqo1, aldh3a1 and adh). Also, gene regulation indicative of DNA damage repair was seen on the microarrays together with reduced proliferation, indicating cell cycle arrest (regulation of histones, and some genes related to the NER pathway, polr2a, polr2f, polr2g, ddb1, rbx1, rad23b, rpa1, ercc5, pold1, pold2, rfe3, ligase1, pole4 and ubc). Moreover, the transcriptome returned to control levels at the low BaP concentration, and based on that it was hypothesized that DNA adducts were repaired at the low, but accumulated at the high, BaP concentration (with cells eventually dying).

Having this as background, this work aimed to quantify BaP internal concentration and DNA adducts in a time and concentration dependent manner. In addition, equilibrium constants, cellular uptake, biotransformation and adduct formation rates were calculated. In order to quantify the BaP, its metabolite concentration and adducts, Hepa1c1c7 cells were exposed to two different BaP concentrations, namely 50 nM and 5 µM, over different exposure times.
4.2 Material and Methods

Chemical and Reagents

Benzo(a)pyrene was purchased from Sigma-Aldrich (Switzerland) and 14C-Benzo(a)pyrene was purchased from American Radiolabeled Chemicals (USA). Acetonitrile, was purchased from ACROS organics (Switzerland), Dimethyl sulfoxide (DMSO) was purchased from Sigma-Aldrich (Switzerland) and glacial acetic acid from Merck (Switzerland). All test chemicals and solvents were of analytical grade (>96 %). Pure deionized water was prepared using a Millipore water purification system equipped with an organic-free kit (Thermo fisher, Switzerland). The standards [15N5]BPDE-N2-dGuo and BPDE-N2-dGuo were a kind gift from The Blair Lab, University of Pennsylvania and the Hecht Laboratory, University of Minnesota.

Enzymes used for the digestion mix, Benzonase, Phosphodiesterase I and Alkaline phosphatase were purchased from Sigma-Aldrich (Switzerland).

Cell culture and exposure conditions

The mouse hepatoma Hepa1c1c7 cell line was grown in Dulbecco’s modified medium without phenol red (Biochrom AG, Berlin) and complemented with 7 % of fetal calf serum (Biochrom AG), 1 % of penicillin/streptomycin (Sigma-Aldrich, Switzerland) and 2 mM of L-alanine and L-glutamine (Biochrom AG, Berlin).

A cell density of 10⁴ cells cm⁻² was used throughout. The cells were seeded in 75 cm² flasks on the first day, the medium was exchanged one day thereafter, and two days after seeding the cell were exposed to BaP in DMSO or to DMSO alone as solvent control. The final concentration of DMSO in the culture medium was 0.065 % (v/v) for the BaP quantification and 0.05 % (v/v) for the quantification of DNA adducts. BaP was used in two concentrations, 50 nM and 5 µM, and cells exposed for 2, 4, 12, and 24 h.

Upon termination of exposure, cells were processed in three ways for determination of i) BaP quantities in the exposure compartments (cells, medium, plastics of culture flask); ii) cell number to calculate internal BaP concentrations per cell; and iii) quantification of DNA adducts. Experiments related to BaP quantification were performed using a mixture of labelled and unlabelled BaP, while for the other experiments unlabelled BaP was used throughout.
**Liquid Scintillation Counter, fluorescence and radio-HPLC Analysis**

For the chemical analysis, cells were treated as described above for 2, 4, 12 and 24 h using a mixture of labelled and unlabelled BaP. After the exposure time, the samples were collected and extracted with acetonitrile, 12 and 24 h samples were acidified with acetic acid for better extraction of the metabolites. Liquid scintillation counter (LSC) was used to quantify total radioactivity of the samples while fluorescence and radio HPLC was used to differentiate between BaP and its metabolites. This experiment was done with three biological replicates.

For calculating total radioactivity in each sample, a liquid scintillation analyzer (LSC, Radiomatic™ 500TR-Series Flow Scintillation Analyzer from Packard Cranberra Company) was used. A volume of 10 mL of the LSC cocktail (Ecoscint A, Brunschwig, Switzerland) was added to 100 μL of each sample and measured three times for 10 minutes. Subsequently, the background of all the radioactive sample counts was corrected for background activity by using triplicates of blank controls (10 mL of Ecoscint A cocktail).

HPLC analyses were performed on a Hewlett-Packard system equipped with a HP1100 device fluorescence detector (Hewlett-Packard GmbH) and a flow scintillation analyzer (500 TR Series; Packard Cranberra Company). A Hypersil Green PAH column (150 x 2.1 mm, 5 μm, Thermo Scientific, Switzerland) was used for separation, the column was thermostated at 40 °C. The BaP and its metabolites were eluted using a gradient method in which the mobile phase consisted of A, water, and B, acetonitrile, held at 40 % A and 60 % B for 3 min, then ramped to 20 % A, 80 % B, for 3 min, 0 % A, 100 % B, for 5.10 min, 40 % A, 60 % B, 0.10 min and 40 % A, 60 % B for 6.2 min, the flow rate was 0.5 mL/min. BaP was monitored at excitation/emission 264/408 nm.

To determine the limit of detection (LOD) and quantification (LOQ) of BaP by all the analytical methods, signal to noise ratio of 3 and 10 was determined respectively. The LOD for LSC was 6 nmol L⁻¹ and the LOQ was 28.6 nmol L⁻¹; LOD for the fluorescence-HPLC was 2 nmol L⁻¹ and the LOQ was 15 nmol L⁻¹, while a LOD of 37 nmolL⁻¹ and LOQ of 181 nmol L⁻¹ was obtained for radio-HPLC.
**Cell number**

An additional flask with cells exposed to unlabelled BaP was used to determine the cell number after the specific times of exposure (2, 4, 12 and 24 h) and test concentrations (50 nM and 5 μM). After the respective time of incubation, the cells were suspended in 5 mL of culture medium and cell numbers were determined using a Neubauer chamber. The whole experiment was performed six times in order to obtain a more precise number of cells.

**Isolation and enzymatic digestion of DNA**

After 2, 4, 8, 12 and 24 h of exposure, DNA was extracted using the Wizard® SV Genomic DNA Purification System (Promega, Switzerland) and at least 20 μg DNA was digested using a digestion mix of benzonase (50 U), phosphodiesterase I (60 mU) and alkaline phosphatase (40 U) according to (QUINLIVAN and GREGORY, 2008). After digestion, the samples were spiked with 200 pmol \([^{15}N_3]BPDE-N^2\)-dGuo. The enzymatically digested samples were loaded onto a 50 mg Sep-Pak C18 cartridge (Waters) that had been activated using 1 mL of methanol and 2 mL of milliQ water. Upon loading, the cartridge was washed with 1 mL of water and 1 mL of 10 % methanol in water, and the DNA-adducts were eluted with 1 mL of methanol. Methanol was removed by rotary centrifugation, and dried samples were reconstituted in 20 μL of 50 % methanol / 50 % milliQ water for liquid chromatography–mass spectrometry (LC-MS) analysis.

**Quantitative analysis of DNA adducts by LC-ESI-MS/MS.**

The LC-MS/MS equipment consisted of a nanoAcquity UPLC system (Waters) connected to a tandem quadrupole mass spectrometer (Thermo LCQ Vantage) with an electrospray ionization interface (ESI). Quantitative analyses done in collaboration with Dr. Paul van Midwoud and Dr. Shana J. Sturla (Laboratory of Food and Nutrition Toxicology, ETH) were carried out in selected reaction monitoring (SRM) mode, and MS ionization parameters were optimized by tuning with 1 μM BPDE-N^2-dGuo solution. The ESI source was set in positive ion mode with the following parameters: capillary temperature, 330 °C; voltage, 3 kV; sheath gas pressure, 10; Q2 CID gas pressure, 1.5 mTorr; collision gas, argon; scan width, \(m/z\) 0.100; scan time, 0.050 s; collision energy, 17 V; Q1 peak width, 0.70 amu; Q3 peak width, 0.70 amu. Mass transitions monitored were \(m/z\) 570.0 to 454.0 for BPDE-N^2-dGuo, and \(m/z\) 575.0 to 459.0 for \([^{15}N_3]BPDE-N^2\)-dGuo. Xcalibur software (Thermo) was used for data acquisition and processing. Chromatography was performed with a Phenomenex Synergi Polar-RP 80 Å column (150 × 0.5 mm, 4 μm particle size). The HPLC flow rate was
10 µL / min and the mobile phases were 3 % acetonitrile and 0.1 % formic acid in H$_2$O (v/v) (mobile phase A) and 0.1 % formic acid in acetonitrile (v/v) (mobile phase B). The following gradient was used: 0 min – 10 % B, 37 min – 30 % B, 37.1 min – 100 % B, 40 min – 100 % B, 40.1 min – 10 % B, and 50 min – 10 % B.

LOD and LOQ were based upon a signal to noise ratio of 3 and 10, respectively. LOD was 5 fmol (~0.3 adducts per 10$^7$ nucleotides) and LOQ was 17 fmol (1 adduct per 10$^7$ nucleotides).
4.3 Results

Mass Balance

Measurements made by LSC, fluorescence and radio HPLC were used to calculate the mass balance of BaP, using the values obtained for the different fractions (medium, plastic and cells) for both concentrations over time (Table 4.1 and Figure 4.3). LSC measurements provide the total amount of radio activity present in the samples while HPLC measurements (both Fluorescence and Radio) reflect the amounts of parent BaP present. Major amounts of BaP stayed in the culture medium fraction while less than 10% of the BaP were recovered from each the plastic and cell fraction. A decrease of BaP amounts was observed at the 12 and 24 h exposure time for both concentrations (Fluorescence and Radio HPLC, Table 4.1 and Figure 4.3), indicating metabolism. BaP metabolites are also being excreted from the cells to the culture medium, as revealed by the difference in the amounts detected by LSC vs. HPLC (LSC and HPLC, Table 4.1 and Figure 4.3, A and D).

Cell internal concentrations (percentage of initial amount) reached a maximum in the first hours for both concentrations (Table 4.1 and Figure 4.3, A and D). Most of BaP was metabolized (more than 90%) for both concentrations after 24 h of exposure (Figure 4.3, B and E). For the lower concentration (50 nM), no BaP was found in the cell fraction at this time point, while for the higher concentration (5 µM), 1.2 ± 0.3% was still present (Table 4.1 and Figure 4.3, A and D). In the medium and the plastic fraction, it was still possible to detect BaP for both exposure concentrations (Figure 4.3, B and E). In the plastic fraction, almost only parental BaP was detected (Figure 4.3, C and F). Finally, there was a slight decrease (from 3.8 to 1.4% and 4.7 to 2.1% for 50 nM and 5 µM respectively) of BaP found in the plastic fraction over time.
Table 4.1: BaP mass balance over time and concentration determined by LSC, FLD and Radio HPLC as percentage of initial measured concentration. Mass balance was achieved by adding the values obtained in the different phases (culture medium, plastic and cells).

<table>
<thead>
<tr>
<th>Exposure Time (h)</th>
<th>Concentration</th>
<th>Fractions</th>
<th>Analytical Method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LSC(^1)</td>
<td>FLD HPLC(^1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50 nM</td>
<td>5 µM</td>
</tr>
<tr>
<td>2</td>
<td>medium</td>
<td>101.8 ± 5.5</td>
<td>85.5 ± 6.5</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>4.4 ± 1.2</td>
<td>3.8 ± 1.4</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>4.1 ± 0.9</td>
<td>0.9 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>96.0 ± 1.5</td>
<td>99.8 ± 3.9</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>4.7 ± 0.3</td>
<td>4.7 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>6.5 ± 0.7</td>
<td>6.5 ± 0.2</td>
</tr>
<tr>
<td>4</td>
<td>medium</td>
<td>88.7 ± 20.3</td>
<td>76.9 ± 13.9</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>3.2 ± 0.2</td>
<td>2.9 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>1.9 ± 0.1</td>
<td>1.2 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>76.2 ± 0.8</td>
<td>72.0 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>2.9 ± 0.03</td>
<td>2.8 ± 0.003</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>9.6 ± 0.1</td>
<td>7.2 ± 0.007</td>
</tr>
<tr>
<td>12</td>
<td>medium</td>
<td>65.8 ± 5.4</td>
<td>15.9 ± 8.1</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>2.2 ± 0.5</td>
<td>2.5 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>1.9 ± 0.7</td>
<td>1.3 ± 1.1</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>91.0 ± 7.1</td>
<td>34.1 ± 13.1</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>4.0 ± 0.9</td>
<td>2.7 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>4.5 ± 0.8</td>
<td>1.2 ± 0.4</td>
</tr>
<tr>
<td>24</td>
<td>medium</td>
<td>75.5 ± 3.7</td>
<td>1.7 ± 3.0</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>1.8 ± 0.6</td>
<td>1.4 ± 0.6</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>1.8 ± 0.6</td>
<td>&lt; LOD</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>84.6 ± 2.4</td>
<td>9.2 ± 1.9</td>
</tr>
<tr>
<td></td>
<td>plastic</td>
<td>3.1 ± 0.1</td>
<td>2.1 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>cell</td>
<td>3.8 ± 0.5</td>
<td>1.2 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>dead cells</td>
<td>0.3 ± 0.1</td>
<td>0.3 ± 0.3</td>
</tr>
</tbody>
</table>

\(^1\) Values are represented as percentage of initial measured concentration. LOD – limit of detection.
Figure 4.3. Percentage of BaP measured by the different methods for all the fractions for 50 nM and 5 µM A and D – cell; B and E - culture medium; D and F - plastic. LOD - limit of detection. Samples for the low BaP concentration were measured by Radio HPLC, however, the values were under the detection limit.
Chapter 4

Equilibrium constants

A first order kinetics model was set up to describe how BaP distributes among the different phases in the exposure set up (Figure 4.4). BaP is considered to partition between the culture medium, plastic (culture flasks) by reversible hydrophobic interactions. Uptake into the cells is by passive diffusion or active cellular mechanisms. Volatility was assumed not to play a role in this system, due to total recovery of the chemical (LSC, Table 4.1) and the low Henry’s law constant (4.9 $10^{-7}$ atm m$^{-3}$ mol$^{-1}$). The culture medium was used as loading system. The following differential equation was used to describe the change in BaP concentrations in the system over time: $dC_{BaP}/dT = -k_{bio}C_{BaP}$, where $C_{BaP}$ is the total concentration of BaP, $T$ is time, $k_{bio}$ is the constant of biotransformation. The rate of distribution ($K_d$) between culture medium and plastic is considered to be reversible but the uptake rate ($k_u$) and biotransformation rate ($k_b$) are unidirectional. The uptake was calculated as: $dC_{BaP}/dT = - k_u C_{BaP'} / (1+K_d)$ and $C_{BaP'}$ is the aqueous phase concentration.

Figure 4.4. BaP distribution model. $k_d$ is the distribution rate constant of BaP between the plastic and the culture medium, $k_u$ is the uptake constant and $k_b$ is the biotransformation constant. Based on the results of the mass balance and the physicochemical properties of BaP, evaporation was not accounted for.
**Rate of distribution (Kd)**

Partitioning of BaP between plastic and culture medium was calculated for both exposure concentrations based on all measurements performed for all time points considering only flasks without cells as the ratio of aqueous phase and sorbed phase concentrations ($C_{BaP, plastic}/C_{BaP, medium}$). For the lower concentration, 50 nM, the partition coefficient was calculated to be $0.04 \pm 0.01 \text{ L cm}^{-2}$, while for the high concentration, 5 µM was $0.07 \pm 0.03 \text{ L cm}^{-2}$. These values are not significantly different and can be interpreted as approximately 4 to 7% of BaP binds preferentially to the plastic.

**Rate of uptake (ku) and biotransformation (kb)**

Cell number was determined for each time point and concentration (Figure 4.5). It can be seen that the cells exposed to the high BaP concentration did not proliferate. Concentrations in the cell fraction were then divided by the average of cell number to determine the BaP concentration per cell (Table 4.2). For the first two hours, the cells exposed to the high BaP concentration took up $\sim$900 times more than the cells exposed to the lower concentration of BaP. At 12 h, this difference reduced to $\sim$180 times and at 24 h there was almost no difference.

Constant of uptake (ku) was calculate as $0.17 \text{ h}^{-1}$ and $0.12 \text{ h}^{-1}$ and constant of biotransformation (kb) was calculate as $0.15 \text{ h}^{-1}$ and $0.1 \text{ h}^{-1}$, for 50 nM and 5 µM BaP, respectively, indicating that uptake happens faster than biotransformation for both concentrations (Figure 4.3 and Figure 4.6 A and D). Also, when comparing among the concentrations, uptake and transformation are faster for the higher concentration of BaP.

**Table 4.2.** BaP concentration and number of molecules per cell at each time point and concentration determined by fluorescence HPLC.

<table>
<thead>
<tr>
<th>Concentration</th>
<th>Concentrations</th>
<th>2 h</th>
<th>4 h</th>
<th>12 h</th>
<th>24 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 nM</td>
<td>pg / cell</td>
<td>$5.2 \pm 3.1 \times 10^4$</td>
<td>$8.0 \pm 6.7 \times 10^4$</td>
<td>$4.8 \pm 5.0 \times 10^4$</td>
<td>$&lt; \text{LOD}$</td>
</tr>
<tr>
<td></td>
<td>N° molecules / cell</td>
<td>$1.24 \times 10^9$</td>
<td>$1.9 \times 10^9$</td>
<td>$1.14 \times 10^9$</td>
<td>$\text{does not apply}$</td>
</tr>
<tr>
<td>5 µM</td>
<td>pg / cell</td>
<td>$5.0 \pm 1.9 \times 10^1$</td>
<td>$7.1 \pm 2.9 \times 10^1$</td>
<td>$9.0 \pm 5.0 \times 10^2$</td>
<td>$8.0 \pm 5.0 \times 10^2$</td>
</tr>
<tr>
<td></td>
<td>N° molecules / cell</td>
<td>$1.19 \times 10^7$</td>
<td>$1.69 \times 10^7$</td>
<td>$2.14 \times 10^8$</td>
<td>$1.9 \times 10^8$</td>
</tr>
</tbody>
</table>
DNA adducts

Only the isomer t(+)BPDE-dGuo was detected in the cells exposed to 50 nM BaP while both isomers, t(-)BPDE-dGuo and t(+)BPDE-dGuo were detected in the cells exposed to 5 µM BaP. The amounts of DNA adducts present for each concentration differed immensely (Table 4.3 and Figure 4.7). For the lower concentration, the adducts formed between 4 and 12 h disappeared after 24 h of BaP exposure, while for the high concentration, DNA adducts continued to accumulate and reached twice the numbers at 24 h compared to 12 h. Adducts were detected at 4 h of exposure for the lower concentration, however only in one of the triplicates.
Table 4.3: BPDE-dGuo adduct per $10^7$ nucleotides detected at each time point and concentration in Hepa1c1c7 cells.

<table>
<thead>
<tr>
<th>BaP Concentration</th>
<th>Time (h)</th>
<th>t(-)BPDE-dGuo</th>
<th>t(-)BPDE-dGuo</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 nM</td>
<td>2</td>
<td>0.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.0 ± 0.0</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.0 ± 0.0</td>
<td>0.7 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.0 ± 0.0</td>
<td>1.4 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>0.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td>5 µM</td>
<td>2</td>
<td>0.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.0 ± 0.0</td>
<td>0.0 ± 0.0</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.6 ± 0.2</td>
<td>22.4 ± 8.0</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2.8 ± 0.1</td>
<td>69.4 ± 23.2</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>4.2 ± 0.4</td>
<td>197.4 ± 47.9</td>
</tr>
</tbody>
</table>

Rate of DNA adduct formation ($k_f$)

The rate of adduct formation of each isomer was calculated for each concentration of BaP. The following differential equation $dC_{TP}/dT = + k_{bio} C_{BaP} \theta - k_{bio,TP} C_{TP}$, was used where $k_{bio}$ is the constant of biotransformation of BaP, $C_{BaP}$ is the chemical concentration, $\theta$ is the formation of adduct formation, $k_{bio, TP}$ is the constant of biotransformation of the measured transformed product, and $C_{TP}$ is the concentration of the measured transformed product, in this case BPDE adducts. The rate of constant of formation is the product of the constant of biotransformation of BaP and the fraction of formation of the measured transformed product.

At the lower BaP concentration, the rate constant of formation for the t(+)-BPDE-dGuo adduct was calculated as $1.5 \times 10^{-6}$ h$^{-1}$, while for the high concentration was $1.0 \times 10^{-6}$ h$^{-1}$. For the t(-)-BPDE-dGuo, there was no detectable formation at the low concentration and for the high concentration adduct formation rate was $2.6 \times 10^{-8}$ h$^{-1}$, indicating that the t(-)-BPDE-dGuo isomer is formed slower than the t(+)-BPDE-dGuo. The fraction of formation (amount of epoxide over the amount of BaP measured in the culture medium by HPLC) for t(+)-BPDE-dGuo was $1.0 \times 10^{-5}$, for the low and high BaP concentration respectively, and for t(-)-BPDE-dGuo was $2.6 \times 10^{-7}$ for the high concentration (Figure 4.6, B, C, E and F). These results indicate that a very small fraction of DNA adducts is formed from the number of BaP molecules present.
Figure 4.6. Model of the measured data. A and D - BaP concentrations in the different compartments over time for the 50 nM and 5 µM BaP, respectively. B and E - total BaP concentration (black points and line) over t(+)BPDEdGuo (red points and line) for 50 nM and 5 µM BaP. C and E - total BaP concentration (black points and line) over t(-)BPDEdGuo (red points and line) for 50 nM and 5 µM BaP. C medium, BaP concentration measured in the culture medium; C plastic, concentration bound to the plastic; C cell, concentration measured within the cells.
Figure 4.7. Amount of the different DNA adduct isomers per $10^7$ nucleotides present over time and at both BaP concentrations (50 nM and 5µM).
4.4 Discussion

As previously demonstrated by different mammalian cells and techniques, the uptake of BaP by cells happens fast (BARTLEY et al., 1982; SUREAU et al., 1990; BARHOUMI et al., 2000). In this study, highest cell internal BaP concentrations were observed within the first 4 h for both concentrations, 50 nM and 5 µM. Based on total vs. parent BaP and DNA adduct formation, BaP metabolism starts between 4 and 8 h of exposure, which is in accordance with the results presented in Chapter 2 where Cyp1a catalytic enzyme activity was measured.

Major amounts of BaP stayed in the culture medium, presumably due to the complexity of the serum component. Schirmer et al (1997), using a confluent fish cell line monolayer, exposed for 24 h to fluoranthene, also showed that when the culture medium was supplemented with fetal bovine serum (FBS) the amount associated with this fraction was between 60-70 %, while the culture medium without FBS had only between 9 - 15 %. Bopp et al (2006), also using a fish cell line, showed that without FBS most of BaP binds to the plastic. Hestermann et al (2000) showed that serum altered the uptake of TCDD by a fish cell line, leading to a reduction of Cyp1a induction when compared to cells exposed to TCDD in culture medium without serum.

With the conditions used in these experiments, it seems that biotransformation, not uptake, was the rate limiting step because uptake for both concentrations occurred slightly faster than the biotransformation. However, for both concentrations, more than 90 % of BaP was bio-transformed, even though the cells at the high exposure concentration were damaged (Figure 4.5 and Chapter 2, Figure 2.1).

The number of BaP molecules within one cell after 2 and 4 h were between $1.2 - 1.9 \times 10^6$ and $1.2 - 1.7 \times 10^9$ for 50 nM and 5 µM, respectively. Holmez and Pollenz (1997) determined that the number of AhR and ARNT molecules in Hepa1c1c7 cells were $3.2 \times 10^5$ and $3.3 \times 10^4$ respectively. This shows that the number of BaP molecules within a cell, already at the lower concentration, is much higher than the number of transcription factors. Nevertheless, one also has to consider that, based on the physicochemical properties of BaP and available data, not all BaP molecules would be available as activator of the AhR/ARNT transcription factor. Rather, a fraction of BaP will likely be associated with lipophilic regions in the cell, like plasmatic membrane, mitochondria, lysosomes, endoplasmic reticulum and nucleus (ALLISON et al., 1966; BARHOUMI et al., 2000). This would suggest that the activation of the AhR pathway might be partially limited by the cell and not by the amount of chemical.
The amount of DNA adducts found in our experiments for the high BaP concentration are in accordance with previous reported data for Hepa1c1c7 using $^{32}$P-postlabelling analysis and a different cell density (Holme et al., 2007). Extrapolating the results obtained for the high BaP concentration, we measured 95 adducts $10^{-7}$ base pairs while Holmes et al detected approximately 72 adducts $10^{-7}$ base pairs. Besides that, Holme et al (2007) using a cell density of $2 \times 10^4$ cells cm$^{-2}$ exposed to different concentrations of BaP for 15 h observed a concentration dependent relationship between 1 and 5 µM but not between 5 and 10 µM. The induction and posterior repair of adducts at the low BaP concentration however was never reported in cells before. Repair of adducts, although not complete, was previously shown when cells were kept an extra 72 h in a BaP-free culture medium (Bartley et al., 1982).

As shown in Chapter 2, the low BaP concentration led to a slight increase in doubling time so as to allow for DNA damage repair while cell viability was not affected according to the endpoints observed. In contrast, cell proliferation was completely abolished at the high BaP concentration. BaP metabolism is known to form ROS as one of the most abundant by-products. ROS was indeed detected at 5 µM BaP exposure (Chapter 2 Figure 2.3) indicating extensive BaP metabolism at this concentration. Holme et al (2007) also showed the presence of ROS at 5 µM or higher BaP concentrations. The phenotype state of the cells at the high concentration (Chapter 2 Figure 2.1) can be mainly explained due to the toxic effect of the metabolites and the accumulation of DNA adducts over time.

Two different isomers, t(+)BPDE-dGuo and t(-)BPDE-dGuo, were detected in the DNA in different quantities, the t(+)BPDE-dGuo being the most abundant. Rojas and Alexandrov (1988) also observed in mouse and rat epidermis a higher amount of t(+)BPDE-dGuo, 8 and 5 times more, than the t(-)BPDE-dGuo. However, the same was not observed in the mouse dermis. Pruess-Schwartz et al (1986) also demonstrate that t(+)BPDE-dGuo was the major adduct formed in mammary epithelial and fibroblastic cells.

The difference in the amount of isomers can be explained mostly by the rate of formation of the BPDE precursor, ($\pm$)7,8 dihydrodiol BaP. Yang et al (1976) and Thakker et al (1977) demonstrated that liver microsomal enzymes from rats pre-treated with methylcholanthrene metabolized BaP to 24 times more ($-$)7,8 dihydriodiol BaP than ($+$)7,8 dihydriodiol BaP. The same pool of enzymes further metabolized the ($-$)7,8 dihydriodiol BaP to 6 times more ($+$)7β8α diol 9α10α (which is a pre-cursor for t(+)BPDE-dGuo) epoxide than ($-$)7β8α diol 9β10β epoxide (the pre-cursor for t(-)BPDE-dGuo), while ($+$)7,8 dihydriodiol BaP was transformed 22 time more to ($+$)7α8β diol 9α10α epoxide (c(+)BPDE-dGuo) than ($-$)7α8β diol 9β10β epoxide (c(-)BPDE-dGuo). However, the cis
adducts were not detected in our samples. Slaga et al (1979) showed that t(+)BPDE-dGuo adducts caused 60 to 70 % of tumour initiating activity in mouse skin while the three other isomers were weak or essentially inactive.

The t(+)BPDE-dGuo adducts bind strongly to the DNA and are more effective in causing DNA unwinding than the other isomers. Besides that, t(±)BPDE-dGuo adducts adopt an external conformation that accommodate the benzo(a)pyrenyl moiety into the minor groove and retain Watson-Crick alignment, however, there is a reduced stability of the GC. Nonetheless, c(±)BPDE-dGuo adopt an internal conformation characterized by the benzo(a)pirenyl insertion into the double helix and concomitant disruption of Watson-Crick hydrogen bound. As a consequence, these helix insertions cause displacement of the covalently modified guanine and its cytosine partner into the major or minor groove. Interestingly, BPDE- dGuo are processed by human nucleotide excision repair (NER) about 1 order of magnitude less efficiently in the trans configuration than in cis (HESS et al., 1997). This might indicate that the conformational arrangement of the adduct might influence its repair.

Microarray results indicated a regulation of the NER pathway at the high but not the low BaP concentration (Chapter 2). NER is known to promote genome stability by removing a wide range of carcinogen-DNA adducts, such as BPDE. NER cuts damaged strands on either side of the target lesions, followed by excision of oligonucleotide segments 24 to 32 residues in length. Further, double helical integrity and the correct nucleotide sequence are re-established by DNA repair synthesis and DNA ligation. Excision activity, however, is highly non-uniform in the context of mammalian chromosome. Frequently, sites that are slowly repaired coincide with hot spots of BPDE mutations (HESS et al., 1997).

There is considerable evidence that adduct formation is a non-random process that occurs with certain specificity (WARPEHOSKI and HURLEY, 1988; HEMMINKI, 1993). BPDE has preference to bind to regions reach in GC and runs of continuous guanines in the genome (MATTES et al., 1986; MATTES et al., 1988; SAID and SHANK, 1991). Hot spots for BPDE adduct formation have been described in p53, ras and hprt (WEI et al., 1991; DENISSENKO et al., 1997; HUSSAIN et al., 2001; FENG et al., 2002). Histone genes have a high GC content, which puts them among the probable places for adducts. However, no studies have shown that histone genes are hotspot for adduct formation although BPDE has been shown to form covalent adducts with histone proteins in vitro and in vivo (KUROKAWA and MACLEOD, 1985; KUROKAWA and MACLEOD, 1988; MANN et al., 1997). Microarray analysis (Chapter 2) showed a down regulation of histones at the high, 5 µM, BaP concentration. Hockley et al
(2006) also showed a down regulation of histones in human cells exposed to BaP. The repression of histone genes are normally considered an indication of cell cycle arrest. This is likely a response to create time for the cell machinery to repair DNA damage (SCHRECK et al., 2009). Further research needs to address the question whether histone genes are hot spots for BPDE adducts or if the histones are just regulated during cell cycle arrest.

BaP has been demonstrated to induce apoptosis in Hepa1c1c7 cells through activation of intrinsic caspase cascade, mitochondrial dysfunction and p53 activation (KO et al., 2004). Holme et al (2007) suggested that the rate of metabolism and type of reactive metabolites formed influenced the resulting balance of pro and anti-apoptotic cells signalling, and hence the mechanisms involved in cell death and the chances of permanent genetic damage. In the exposure set up chosen, cell number was also determined for 48 h for the high BaP concentration (Chapter 2, Figure S1), however, no increase in cell number was observed, meaning that the cells were still arrested. Dautel et al (2011) showed an arrest at the S-phase of Hepa1c1c7 cells exposed to 5 µM for 48 h. This might indicate that the extension of the damages within the cells where too high to recover even if most of the BaP present was bio-transformed already after 24 h.

Cell internal concentration is dependent on the external exposure concentration. The number of BaP molecules in cells, for both concentrations, are higher than the amount of the transcription factors AhR and ARNT. Metabolism of BaP induces adduct formation. Amount of adducts are related to the concentration used and repair is only observed at the low BaP concentration. This is in line with the results observed in Chapter 2 where cells exposed for 24 h to 50 nM BaP show no disturbance (for the end points tested). For the high BaP concentration, the number of DNA adducts increases over time and it is not repaired in the course of 24 h. The accumulation of adducts maintain the cells at an arrest phase (increasing its doubling time) and can induce apoptosis. Arrest at S-phase was observed when the cells were kept for 48 h at the high BaP concentration (DAUTEL et al., 2011).

BPDE-adducts repair is normally processed by NER, however, this system was shown to be selective to the type of adducts, being more efficient against cis than trans adducts, due to the position of the adducts in the DNA molecule. t(+)-BPDE-dGuo adducts seems to be formed with a higher preference than the others isomers, besides that, it seems also to be less efficiently repaired. The formation preference of these adducts added to the lower efficiency of repairing might be the answer why this isomer is considered to be the most carcinogenic.
4.6 References


Kurokawa, M.; MacLeod, M. C., Differences in the binding of stereoisomeric benzo(a)pyrene-7,8-diol-9,10-epoxides to histones in rat liver nuclei. *Cancer Res* **1985**, *45* (11 Pt 2), 5741-5.


Rojas, M.; Alexandrov, K., Comparison of DNA adducts in mouse and rat epidermis versus dermis after topical application of (±)-t-7,t-8-dihydroxy-t-9,10-oxy-7,8,9,10-tetrahydrobenzo(a)pyrene and (±) benzo(a)pyrene-4,5-oxide. *Journal of Cancer Research and Clinical Oncology* 1988, 114 (1), 110-112.


Yang, S. K.; McCourt, D. W.; Roller, P. P.; Gelboin, H. V., Enzymatic conversion of benzo(a)pyrene leading predominantly to the diol-epoxide r-7,t-8-dihydroxy-r-9,10-oxy-7,8,9,10-tetrahydrobenzo(a)pyrene through a single enantiomer of r-7, t-8-dihydroxy-7,8-dihydrobenzo(a)pyrene. *Proc Natl Acad Sci U S A* **1976**, *73* (8), 2594-8.

5. Conclusion & Outlook

The overall aim of this thesis was to investigate the molecular response of the murine hepatoma cell line, Hepa1c1c7, on BaP exposure, combining high throughput technique, bioinformatics, phenotype characterization and chemical analysis in a time and concentration dependent manner. BaP was chosen as a relevant environmental contaminant, widely distributed and classified as carcinogenic to humans, Group 1, by the International Agency for Research on Cancer (IARC). Exposure time varied from 2 to 48 h and exposure concentration from 0.5 nM to 5 µM BaP. In this Chapter, the main findings are summarized and discussed in a System Biology perspective (Figure 5.1).

Transcriptome analysis indicates that BaP enters the cells rapidly, already inducing transcriptome changes of AhR regulated genes 2 h after the onset of BaP exposure. Quantification of cell internal concentration by Liquid Scintillation counter (LSC) revealed that, based on the time points analysed, the maximum cellular uptake for the 50 nM BaP was at 2 h (~4 % of initial concentration) and for the 5 µM BaP was at 4 h (~10 % of initial concentration) (Figure 5.1a).

It was shown in this thesis that the early regulation of genes (2 and 4 h) did not differ among the two different BaP concentrations (50 nM and 5 µM) used. Genes involved in BaP metabolism, like cyp1a1, was present among those genes (Figure 5.1b). Based on Cyp1a activity, quantification of adducts and reduction of parent BaP amount, metabolism was induced between 4 and 8 h and more than 90 % of BaP was metabolized for both concentrations within 24 h, although most of BaP was initially present in the culture medium.

Early gene expression (cyp1a1, tiparp, rnf39 and mpp2) reaches a maximum induction at the 2 or 4 h time point. Gene expression at 4 h could be broadly classified in biotransformation, cytoskeleton/membrane repair/transport, cell cycle/survival, TF/TF activator and miscellaneous. From the early regulated genes, the least known is tiparp, TCDD-inducible poly(ADP-ribose) polymerase. The hypothesis generated in this thesis regarding tiPARP was that this gene is involved in damage repair caused by dioxins and dioxin-like compounds through activation of the AhR. Attempts to define its function through a silencing technique revealed no cellular changes in the silenced compared to control cells. Besides that, when the tiPARP silenced cells were exposed to BaP, the tiparp mRNA levels increased again, indicating that this gene is very responsive to the treatment. Attempts to detect endogenous tiPARP protein failed using both western blot and MudPIT techniques. One explanation for this might be due to the low quantities of this protein in cells or due to its
Chapter 5: Conclusion & Outlook

low predicted solubility. However, based on personal communication with MacPherson, L. & Matthews, J., endogenous tiPARP protein has now apparently been detected by immunofluorescence and it was suggested that its cellular function might be related to regulation of the AhR.

On the phenotype level, DNA adducts were detected already at 8 h of BaP exposure for both concentrations (Figure 5.1c). Regulation of genes related to nucleotide excision repair (NER) was observed after 12 and 24 h at the high BaP concentration. Complete DNA-adduct repair was seen only at the low concentration, while at the high BaP concentration the number of adducts increased over time. Cell cycle arrest, observed based on the transcriptome, a hugely increased cell doubling time, and S-phase arrest, as shown for the same experimental set-up by Dautel et al (2011), is in accordance with the increase in DNA adducts for the high BaP concentration.

One type of adduct, t(-)BPDEdGuo was found to be the most abundant adduct in BaP exposed Hepa1c1c7 cells, followed by t(+)BPDEdGuo. The cis adducts were, however, not detected. This was to be expected since the DNA quantities analysed were at maximum 50 µg and cis adducts have been only reported when analysing milligrams of DNA. Besides that, t(±)BPDEdGuo has been shown to be more resistant to repair because it does not disrupt Watson-Crick hydrogen bounds while the cis adducts do (HESS et al., 1997).

Reactive oxygen species (ROS) was another phenotype end point detected as early as 12 h of 5 µM BaP exposure (Figure 5.1d). BaP metabolism produces several by-products, with ROS being one of them. The detection of ROS at this time point and concentration indicates that cell mechanisms to remove it were insufficient against the amount of ROS produced. Regulation of aldh3a1, sod2, gst1, prdx2, among other genes, was observed at the high BaP concentration already after 4 h. Proteomic studies from one partner from this system biology project confirmed that proteins related to oxidative stress were a main group regulated by BaP (DAUTEL et al., 2011).

Later gene regulation (12 and 24 h) shows induction of several pathways for the high BaP exposure concentration. AP-1, TGFβ and glucocorticoid receptor are among the regulated pathways that are relevant to the cell due to their broad control of the cell machinery, like cell cycle and cell proliferation. Cell cycle arrest was also observed in Hepa1c1c7 cells exposed to the 5 µM BaP for 24 h or more (Figure 5.1e). Cytoskeleton remodelling was another strongly regulated pathway already at 4 h of BaP exposure (Figure 5.1f). Proteomics of Hepa1c1c7 cells exposed to BaP likewise indicated cytoskeleton remodelling as one of the major regulated groups of proteins (DAUTEL et al., 2011). Tekpli et al (2010) showed that
BaP-induced membrane remodelling was an early apoptotic event leading to an intracellular alkalisation in F258 cells.

The presence of metabolites has been shown previously to impact on cell viability (LANDOLPH et al., 1976). The results presented in Chapter 2 corroborate this findings. Cellular damage, like lysosome integrity, metabolic activity and membrane permeability, increased over time for the high BaP concentration, especially for 24 and 48 h (Figure 5.1g). As has been showed in Chapter 4, by high pressure liquid chromatography (HPLC) most of the BaP has been metabolized at the 24 h time point. The low BaP concentration showed only a slight increase in doubling time. At the high concentration, however, a strong increase in doubling time together with a decrease in cell viability was observed. This is in agreement with the results obtained on DNA adducts: adducts had been completely removed in cells exposed to 50 nM BaP for 24 h, while adducts accumulated in cells exposed to 5 µM for 24 h. This result indicates that cells arrested in both concentrations for DNA repair to proceed but this repair was accomplished only at the low concentration of BaP. Cells need to stop growing in order to fix the possible DNA errors to avoid mutations that can be deleterious to themselves.

Gene regulation of p53 (trp53inp1, up regulated at 4, 12 and 24 h), bax (up regulated at 12 h) and bcl-2 (bcl2l11 and bcl2l1 down regulated at 12 and 12 and 24 h, respectively) at the high BaP concentration might indicate that cells at these conditions are entering apoptosis (Figure 5.1h). P53 plays a role in apoptosis and Bcl-2 proteins recruit mitochondria into the apoptotic pathway (GROSS et al., 1999). Dautel et al also observed a decrease in anti-apoptotic proteins and an increase in proteins involved in the apoptosis process (DAUTEL et al., 2011).
Figure 5.1. General scheme of the most relevant findings of this thesis. (a) Cell internal concentration reaches maximum levels within hours although only a fraction of the BaP added is initially bioavailable; (b) induction of genes involved in metabolism and Cyp1a catalytic activity, seen as early as 2 h of exposure; (c) formation of adducts between 4 and 8 h, transient (50 nM BaP) or permanent (5 µM) cell cycle arrest and regulation of genes involved in NER; (d) formation of ROS (5 µM), regulation of genes (5 µM) (proteins related to oxidative stress response was also observed by Dautel et al. (2011)); (e) regulation of pathways that influence cell cycle progression and development; (f) regulation of cytoskeleton remodelling at the transcriptome level, which confirms the findings on the proteomic level (DAUTEL et al., 2011); (g) decrease in cell viability (5 µM, metabolic activity, lysosome membrane stability and plasma membrane permeability); (h) Cell recovery (50 nM) or death by apoptosis (5 µM); (i) up regulation of abcc4, a known transport of BaP metabolites.

Further research needs to address the importance of the early regulated genes. One possibility is to monitor those genes in a more detailed time interval to better understand the pattern of regulation, especially focusing on early time points (until 8 h of exposure). Protein levels and localization should also be addressed. Endogenous localization of the protein might give an evidence of its function. For these studies wild type cells should be compared to AhR mutant cells or cells silenced for specific genes/proteins.

Another interesting approach would be to monitor the transcriptome, not only for the early genes but for all genes, and phenotype in a continuous exposure situation. This can be achieved by using a partition-controlled dosing system where cells are grown on the bottom
of the insert facing a PDMS membrane that is responsible to continuously load the exposure medium (Kramer et al., 2010). Another passive dosing method available are the O-rings, where silicone rings are loaded with the chemical of interested and the rings are then placed in the bottom of the well (Smith et al., 2010). With these methods it would be possible to track BaP concentrations and metabolism under continuous exposure. Detection of metabolic products being formed in a time and concentration dependent manner could also be addressed by such a dosing system. This would provide an answer as to which metabolites is most cytotoxic.

Another interesting avenue for further investigations relates to hot spots for benzo(a)pyrene diolepoxide (BPDE) DNA adduct formation. BPDE DNA adduct hot spots are normally associated with regions reach in GC. It was previously reported that p53, ras and hprt genes are hot spot regions for BPDE adducts. Histone genes are also reach in GC and can be potential hot spots, however, no research has been done in this direction. Combination of GC content with regions showing slower repair may contribute to accumulation of BPDE DNA adducts. Thus, research is needed to address if histones are hot spot areas for adducts. Herewith, one can define if the down regulation of histones, as observed in this thesis, is a consequence of cell cycle arrest or if the adducts are a physical barrier to the transcription of cell cycle genes.

DNA damage focused on bulky DNA adducts because these are formed by one of the most important BaP metabolites. However, DNA damage by oxidation caused by the presence of ROS can also be considered. The type of DNA damage can be assessed by differentiating bulky DNA damage from the damage caused by oxidation of bases (8-oxo-7,8-dihydroguanine) induced by ROS, for example. Some methods are available for studying DNA damage and repair. For example, BD Biosciences described a way to quantify DNA damage and repair by following the evolution of H2AX foci formation (accumulation of phosphorylated H2AX at the damage site) and recruitment of 53BP1 (P53 binding protein involved in DNA repair) (http://www.bdbiosciences.com/documents/Bioimaging_AppNote_DNADamage.pdf). Both trp53bp1 (53bp1) and h2afx (h2ax) were down regulated at the transcriptome level for the high BaP concentration for 12 and 24 h, respectively. This type of investigation could be used to explore the differences among the different types of DNA damages being present in BaP exposed cells, and would provide a more detailed viewed on DNA repair.

Studying cellular responses at different biological levels over time helps to comprehend the dynamics of this response. Generally, toxicological studies address cell-
chemical interactions in a more static way, focusing on one time point and effects assessment on one level only, e.g. cell phenotype. This thesis demonstrates that a more integrative approach sheds light on the intricate network of chemical-cell interaction to different chemical concentrations over time. The toxicogenomics approach used in this thesis gives indications of the cellular state before any of the phenotype endpoints, demonstrating that effects in cells can be detected on a sub-cellular level prior to changes to the phenotype.

In the context of an adverse outcome pathway (AOP), the work presented here adds to a more detailed understanding of the relationship between the molecular initiating events and the cell response (Figure 5.2). What is more, the time resolved information and rate associated constants can support building of computational models to predict cellular responses in the Hepa1c1c7 and other cell models to BaP and other chemicals with similar molecular targets. For example, the quantitative data provided are valuable for calculating the distribution of a chemical and the adduct formation rates can be incorporated into a mechanistic model to define a threshold potentially leading to cancer and/or organ failure for different BaP exposure scenarios. Thus, the knowledge gained here can contribute to increasing the predictive power of risk assessment.

![Figure 5.2. Overview of a conceptual AOP integrating the data obtained in this thesis. Secondary targets of the AhR are the genes or proteins regulated by the by-products of BaP metabolism, like BPDE and/or ROS. The rates calculated for BaP uptake and biotransformation and the rates calculated for DNA adduct formation can be integrated into a quantitative AOP model to improve risk assessment predictions.](image-url)
References


Landolph, J. R.; Bartholomew, J. C.; Calvin, M., Quantitative studies of the toxicity of benzo(a)pyrene to a mouse liver epithelial cell strain in culture. Cancer Research 1976, 36 (11 I), 4143-4151.


Appendix

Transcriptional signatures of regulatory and toxic responses to benzo-[a]-pyrene exposure

Jacob J Michaelson1†, Saskia Trump2†, Susanne Rudzok2, Carolin Gräbsch2, Danielle J Madureira3, Franziska Dautel4, Juliane Mai5, Sabine Attinger5, Kristin Schirmer3,6, Martin von Bergen4,7, Irina Lehmann2* and Andreas Beyer1*

* Correspondence: irina.lehmann@ufz.de; andreas.beyer@biotec.tu-dresden.de
† Contributed equally
1Cellular Networks and Systems Biology, Biotechnology Center, TU Dresden, Dresden, Germany.
2Dept. of Environmental Immunology, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany.
3EAWAG Aquatic Research, ETH Swiss Federal Institute of Technology, Zürich, Switzerland.
4Dept. of Proteomics, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany.
5Dept. of Computational Hydro Systems, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany.
6EPF Lausanne, Lausanne, Switzerland. 7Dept. of Metabolomics, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany.

BMC Genomics 2011, 12:502-
Transcriptional signatures of regulatory and toxic responses to benzo-[a]-pyrene exposure

Jacob J Michaelson1, Saskia Trump2†, Susanne Rudzok2, Carolin Gräbsch2, Danielle J Madureira3, Franziska Dautel4, Juliane Mai5, Sabine Attinger5, Kristin Schirmer3,6, Martin von Bergen4,7, Irina Lehmann2* and Andreas Beyer1*

Abstract

Background: Small molecule ligands often have multiple effects on the transcriptional program of a cell: they trigger a receptor specific response and additional, indirect responses (“side effects”). Distinguishing those responses is important for understanding side effects of drugs and for elucidating molecular mechanisms of toxic chemicals.

Results: We explored this problem by exposing cells to the environmental contaminant benzo-[a]-pyrene (B[a]P). B[a]P exposure activates the aryl hydrocarbon receptor (Ahr) and causes toxic stress resulting in transcriptional changes that are not regulated through Ahr. We sought to distinguish these two types of responses based on a time course of expression changes measured after B[a]P exposure. Using Random Forest machine learning we classified 81 primary Ahr responders and 1,308 genes regulated as side effects. Subsequent weighted clustering gave further insight into the connection between expression pattern, mode of regulation, and biological function. Finally, the accuracy of the predictions was supported through extensive experimental validation.

Conclusion: Using a combination of machine learning followed by extensive experimental validation, we have further expanded the known catalog of genes regulated by the environmentally sensitive transcription factor Ahr. More broadly, this study presents a strategy for distinguishing receptor-dependent responses and side effects based on expression time courses.

Background

Elucidating the transcriptional response of cells to xenobiotic compounds like drugs or environmental contaminants is of primary importance for understanding the physiological effects of such compounds. However, exposure to xenobiotic compounds often induces a complex transcriptional response comprised of specific (i.e. transcription factor (TF) activated programs) and unspecific regulatory mechanisms. Dissecting these responses and identifying the transcriptional profiles associated with each individual (sub-)effect is essential for explaining specific and possible side effects of drugs or for predicting toxic responses of environmental contaminants.

One of the most studied TFs involved in the response to environmental pollutants or xenobiotic compounds in general is the aryl hydrocarbon receptor (Ahr). The Ahr has been studied for decades mainly because of its critical role in xenobiotic-toxicity and carcinogenesis. In its inactive state, Ahr resides in the cytoplasm in a chaperone complex together with the X-associated protein 2 (Xap2, also known as Aip, Ara9) and heat-shock protein 90 (Hsp90). After ligand binding, the receptor translocates to the nucleus where it associates with its cofactor Arnt (Ahr nuclear translocator) yielding a competent TF. This heterodimer binds to a DNA binding motif called the xenobiotic response element (XRE), which functions as an enhancer in the regulatory domain of a wide range of genes commonly referred to as the Ahr gene battery [1,2]. Some of these genes, such as the cytochrome P450 enzyme Cyp1a1, NAD(P)H:quinone oxidoreductase (Nqo1), aldehyde dehydrogenase (Aldh3a1), UDP glucuronosyltransferase (Ugt1a2) and glutathione-S-transferase (Gsta1), are involved in Phase
I/II metabolism. As previously mentioned, this activation of metabolizing enzymes through Ahr may lead to the formation of toxic metabolites of the activating ligand itself. This is particularly true for benzo-[a]-pyrene (B[a]P), a classical Ahr agonist. Only after the self-induced metabolism of this procarcinogen is the ultimate genotoxic metabolite anti-benzo-[a]-pyrene-trans-7,8-dihydroxy-9,10-epoxid (BPDE) formed. Several studies have examined the transcriptional effects of Ahr activation in different species and cell types [3-6]. However, deciphering the Ahr-specific transcriptional response is not a trivial task, considering that Ahr activation might trigger the activation of other TFs or the generation of toxic metabolites which will add side effects to the observed differential gene expression (Figure 1). Therefore, the overall transcriptional response directly related to Ahr binding is incompletely elucidated, and the number of well-defined Ahr specific genes still remains small.

One strategy to assess Ahr-dependence is to compare gene expression of cells or tissues that have the wild type Ahr with those of Ahr-null cells in a matched genetic background, as was shown by Tijet et al. [7]. In their study they compared the effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in Ahr +/- and Ahr -/- mice after long term exposure. This experimental setup, as the authors themselves conceded, does not allow the discrimination of genes directly regulated through Ahr as a primary response and secondary, downstream effects: both classes would register as being differentially expressed. A time course design with early measurements has the potential to distinguish primary responders, which are likely to change first, from indirect responses that are likely to show up later.

In an elegant experimental setup, Hockley et al. [8,9] sought to separate the primary effects of Ahr activation from the side effect caused by the genotoxic metabolite BPDE. They compared the effects of B[a]P, BPDE and

![Figure 1](image-url)
TCDD exposure in two different human cell lines. Unfortunately, the first time point they investigated was not until six hours after exposure. Considering that it was shown previously that Ahr translocation and nascent transcription is already induced 1 h after TCDD exposure [10], we believe that identification of primary Ahr responders is only possible by including early time points of exposure in gene expression studies.

In this work, we investigate the hypothesis of whether time-resolved transcriptional signatures of genes that are primary Ahr targets differ from the profiles observed for genes responding to the toxic metabolite BPDE. We demonstrate that machine learning can be used for identifying these characteristic signatures and for subsequently classifying genes as to whether they are primary Ahr-dependent targets or indirectly affected (BPDE-dependent) genes. This general strategy of using time course gene expression data to predict transcriptional regulatory roles has been previously explored [11-14], although primarily in lower organisms such as bacteria and yeast.

We expect that because such learning methods are less encumbered by methodological assumptions (compared to traditional statistical comparisons), they are more able to find subtle but meaningful patterns in the data. For example, an important assumption of previous attempts to cluster Ahr-centric expression data [3,15-17] is that co-regulated genes should also be co-expressed. Hence, clustering of genes based on expression patterns should identify sets of genes subject to the same regulatory program. However, in time courses such co-expression may only be present during certain phases. In the case of Ahr we expect co-expression during early time points, whereas expression may diverge later when the influence of Ahr diminishes. The analysis presented here anticipates and effectively deals with this scenario.

Here we employ machine learning techniques coupled to a straightforward yet robust experimental design in order to more clearly define genes that are under the direct transcriptional control of Ahr. This is accomplished by training a Random Forest [18] (RF) classifier to learn the difference between genes responding to B[a]P exposure and side effects caused by the B[a]P metabolite BPDE. The trained classifier is then applied to all genes found to be significantly differentially expressed as a result of B[a]P exposure, and their roles as primary responders or side effects are predicted. In addition, the patterns learned by the classifier are used as a basis for performing weighted clustering. These clusters facilitate a better understanding of the functional relatedness of the perturbed genes. Finally, we support predictions with our own experimental follow-up, as well as with data from independent studies.

Results

Extensive transcriptional response

The transcriptional response due to Ahr activation by 50 nM and 5 μM B[a]P was investigated in murine hepatoma cells (Hepa1c1c7). Exposure effects were examined in time-course data for 2, 4, 12 and 24 h after treatment, together with corresponding vehicle (DMSO) controls.

A total of 2,338 genes were perturbed significantly (FDR <0.05) by exposure to B[a]P and had at least a 2-fold change (with respect to DMSO-exposed cells) at some time point over the course of the experiment. Compared to previous studies of Ahr-mediated temporal gene expression, this represents a very substantial transcriptional response (see Additional File 1, Table S1). These genes were highly enriched for a host of biological processes (summarized in Additional File 1, Table S2), including mRNA transport, control of the cell cycle, apoptosis, and development.

Prediction of primary vs. side effects

The overall analytical framework used here is summarized in Figure 2. Using a matrix of time-resolved gene expression values as predictors (interpolated as described in methods), we trained a Random Forest classifier in a two-class scenario (Ahr primary and side effect). Training labels were assigned based on the significant perturbation of a gene in conditions that suggest being either a primary Ahr responder or responsive to the presence of BPDE (side effect). This yielded 28 genes as primary responders and 559 genes as side effects (Additional File 1, Figure S2), before filtering for outliers. The final classifier had an estimated misclassification rate of 7%. Performance of the classifier on out-of-bag (OOB) data is depicted as a receiver operating characteristic (ROC) curve in Additional File 1, Figure S3, panel A.

We then used this trained classifier to predict on all of the 2,338 differentially expressed genes. The predictions have varying degrees of confidence, indicated by the proportion of votes cast for the predicted class. To establish a threshold above which we could be confident that the classifier was predictive, we permuted the original training labels randomly, trained a Random Forest with these labels, and predicted on all 2,338 genes. In general we found that in this “null” scenario, the Random Forest did not predict with a proportion of votes greater than 0.8. Therefore, we consider a class prediction with a proportion of votes greater than 0.8 to be a reliable prediction (Additional File 1, Figure S3, panel B). After filtering, 81 genes were predicted as primary responders to Ahr (Table 1), 1,308 genes were predicted as side effects, and 949 genes could not be reliably classified (see Additional File 2).
Characterization of transcriptional response programs

To characterize the expression patterns that underlie the classifier’s decision rules, we used the RF proximity measure as an input to PAM (partitioning around medoids) clustering - this combination is a form of weighted clustering. This yielded three coherent clusters, depicted in Figure 3. Clusters 1 and 2 are comprised of genes predicted to be side effects of Ahr activation by B[a]P, while cluster 3 contains genes predicted to be primary responders to Ahr. Clusters 1 and 2 are characterized by undulating expression profiles in the low (50 nM) B[a]P exposure, with the mean behavior of each cluster strongly anticorrelated to the other. The high (5 μM) B[a]P exposure shows less cohesive expression patterns, but with the same general trend of anitcorrelation between clusters 1 and 2. In both cases, time points in the 50 nM B[a]P series are more important for the identity of the clusters than time points in the 5 μM B[a]P series. Cluster 3 is characterized by punctuated expression induction at 3 hours in the 50 nM B[a]P time series, and a slightly extended phase of induction in the 5 μM B[a]P time series. Other time points are unimportant for the cluster’s identity; indeed, the expression of these genes is fairly divergent outside of the common phase of induction. Although cluster 3’s “identity phase” is generally between 3-4 hours after exposure, where all genes in the cluster show elevated expression, several genes (such as Cyp1a1 and Tiparp) in the cluster are highly expressed well before this window.

Using the Kolmogorov-Smirnov (KS) test, we evaluated the clusters for enrichment of genes perturbed by an Ahr mutation (Figure 4). By using data from previous studies [7,19], we performed a 2-way ANOVA and took P values from the genotype*ligand interaction; these P values were used as indicators of genes under the direct influence of Ahr. Genes belonging to the training set were excluded when calculating the enrichment. Cluster 3 was the only cluster to show enrichment for genes perturbed by an Ahr mutation. This result further supports the assertion that cluster 3 contains true Ahr primary responders, and that the classifier is predictive in practice. We similarly checked the three clusters for overrepresentation of known XRE motifs, using UCSC 5 kb upstream promoter sequences and motifs from TRANSFAC (release 2009.3). We found only borderline (P = 0.056) enrichment of an XRE motif among genes in cluster 3 and no enrichment in the other clusters. The lack of significant enrichment among the predicted primary Ahr responders suggests that our knowledge of
the sequence-level requirements for functional binding of Ahr is currently far from complete.

**Experimental confirmation of Ahr dependency**

Two independent experimental approaches were chosen to confirm Ahr-dependency for a subset of representative genes: direct comparison of the transcriptional response of Ahr-expressing Hepa1c1c7 and mutant tao BprC1 cells deficient in endogenous Ahr, as well as confirmation of binding of Ahr in the corresponding promoter regions by chromatin immunoprecipitation (ChIP).

B[a]P is likely to induce side effects caused by B[a]P metabolites independent of direct Ahr activation,
therefore we included TCDD - a non-metabolized Ahr ligand - in our confirmation experiments. Differential expression of Tiparp, Tnfaip2, Cdkn1a, Cdkn1b, Cyp2s1, Nfe2l2, Mpp2, and Klf9 after treatment with B[a]P or TCDD at different concentrations was investigated by quantitative real-time PCR (qPCR). After B[a]P and TCDD exposure, the expression of all genes was induced as soon as 1 h after the start of treatment in Hepa1c1c7 cells, while there was no significant induction compared to vehicle control samples detectable in tao BpRc1 cells up to four hours after exposure (Additional File 1, Figure S4). To complement these experiments, the effect of BPDE treatment on the predicted primary Ahr targets was investigated. After 2 h of

---

**Figure 3** Clustering with the RF proximity measure. PAM clustering was performed with a supervised, weighted distance measure, derived during the classification of Ahr primary responders and side effects. Three distinct programs were found, depicted here as clusters (1-3). Color saturation indicates the importance of the time points for the identity of the cluster. To further emphasize these important time points, this same information is shown again for each cluster (black to yellow scale). The classification of each gene is shown as the proportion of RF votes.
exposure to 5 μM BPDE, a time point for which pronounced induction with B[a]P was observed, no significant upregulation of these genes by BPDE exposure was found (see Additional File 1, Figure S9).

Enrichment of Ahr binding in the promoter region of all chosen genes could be confirmed by ChIP, with fold changes (compared to vehicle control samples) ranging from 7.2-152.2 (Additional File 1, Figure S7). This finding suggests a continuous transcriptional activation by Ahr.

Confirmation of BPDE dependency
To confirm that the differential expression observed for genes of either cluster 1 or 2 is indeed dependent on BPDE, we used the previously mentioned in vitro set-up (Hepa1c1c7 vs. tao Bprc1 cells). The relevant BPDE concentration was determined by comparing the effect of 5 μM B[a]P on Hepa1c1c7 cell proliferation to that of different BPDE concentrations in the same cell line. While BPDE concentrations up to 1 μM had only a marginal effect on proliferation, 5 μM BPDE induced an effect very similar to that seen with 5 μM B[a]P (Additional File 1, Figure S1). Therefore, subsequent qPCR experiments for representative genes were performed with 5 μM B[a]P or BPDE respectively. The chosen genes showed transcriptional responses in Ahr-deficient cells only when exposed to BPDE itself. In Hepa1c1c7 cells B[a]P treatment induced effects similar to BPDE, however with a pronounced time lag (Additional File 1, Figure S8).

Discussion
Exposing cells to xenobiotic compounds like drugs or environmental pollutants often induces a complex transcriptional response, made up of both specific and unspecific regulatory mechanisms. Distinguishing the transcriptional profiles associated with the primary target effect from those acting in parallel is essential for understanding possible side effects of such chemicals.

As an example of such a case, we investigated Ahr, one of the most prominent ligand-activated TFs involved in xenobiotic-induced signaling. The cellular response to Ahr activation can be seen as a mixture of a primary response and side effects. The side effects are due in part to stress caused by the formation of active metabolites of the Ahr ligand, while the primary response is related to Ahr binding to gene regulatory sequences. A subsequent transcriptional cascade downstream of Ahr might be activated by other TFs, which are themselves regulatory targets of Ahr (Figure 1).

We have employed a time-course design involving early and late time points to capture both primary and downstream effects. These effects are separated on the time axis, but it is not obvious a priori where to draw the line, i.e. up to which time point expression changes reflect primary responses. The use of machine learning allowed us to identify the relevant time points in a data-driven way. In addition to weighting time points with respect to their relevance for distinguishing Ahr target genes, this analysis also identifies specific expression patterns that are characteristic of primary Ahr targets.

Ahr target genes
Previously well-described members of the Ahr gene battery like Cyp1a1, Nqo1 Cyp2s1, Aldh3a1, Aldh4a1 and Cyp1b1 [1,20,21] were predicted as primary responders to Ahr. In addition to this qualitative confirmation of the effectiveness of our computational approach, we
could demonstrate Ahr dependency experimentally by chromatin immunoprecipitation (ChIP) and qPCR. Further, we found the set of predicted targets to be enriched for genes that showed significant Ahr genotype-ligand interaction (i.e. 2-way ANOVA) effects based on previously published data [7,19] (Figure 4).

B[a]P is likely to induce side effects caused by B[a]P metabolites independent of direct Ahr activation, therefore we included TCDD - a non-metabolized Ahr ligand - in our confirmation experiments. All of the genes chosen for the qPCR verification confirmed the predicted Ahr-dependency (Additional File 1, Figure S4).

We performed a GO enrichment analysis for a functional evaluation of the predicted target genes. The regulated genes in cluster 3 were enriched for 15 different biological functions including terms related to cell cycle control and proliferation. This influence on the cell cycle is also manifested on the protein level, as we were able to show in a previous study [22]. Experimental confirmation of two of these genes, the cyclin-dependent kinase inhibitors Cdkn1a and Cdkn1b, showed an exclusive induction in wild type cells, together with an enrichment for Ahr binding at the corresponding promoters. Another gene known to be involved in cell cycle regulation, but less well-defined, is the palmitoylated membrane protein 2 (Mpp2). Mpp2 was also strongly induced by TCDD and B[a]P in Ahr-expressing cells, while no differential expression was elicited in the mutant tau BpRc1 cells. A more indirect effect on cell cycle regulation originates from the TNF alpha activated signaling cascade. Five genes (Tnfaip2, Tnfaip8, Traf5, Casp3, Ddx58) involved in this pathway were predicted to be primary responders to Ahr.

Tnfaip2 and Casp3 were investigated in our independent experimental confirmation. For both genes induction of expression was only detectable in Hepa1c1c7 cells, while the Ahr-deficient counterparts showed no significant differential regulation. Actual binding of Ahr to the promoter sites could be confirmed by ChIP. Primary regulation by Ahr of the important regulators of the cell cycle Cdkn1a, Cdkn1b as well as Mpp2 together with targeting of the TNF alpha signaling pathway emphasizes the impact of Ahr on endogenous cellular functions outside of xenobiotic metabolism. Further, these findings suggest that the observed reduction in proliferation after exposure to B[a]P is not only a response to DNA damage, but is also, at least in part, a direct consequence of Ahr activation.

The early time points proved vital in distinguishing Ahr targets from genes induced as side effects (Figure 3), emphasizing the importance of planning experiments such that the immediate effects are captured. Although perturbation at early time points determined the Ahr primary response for both B[a]P concentrations, the consistency of expression between the concentrations diverged later in the time course (Additional File 1, Figure S6). We investigated if indeed Ahr itself might be important for this difference. Comparing the translocation behavior of Ahr we could show a persistent nuclear localization of Ahr for high B[a]P concentrations for up to 24h of exposure, while for low concentrations of B[a]P cells showed fewer and less pronounced translocation events. Obviously many mechanisms might be responsible for the concentration-dependent differences in the transcriptional pattern, like the balance between mRNA production and decay. Nevertheless, persistent Ahr translocation suggests persistent mRNA production, thereby shifting this balance.

An Ahr transcriptional cascade

Twelve of the genes in cluster 3 (i.e. the Ahr target cluster) are known transcriptional regulators. These regulators could constitute a transcriptional cascade that begins with the activation of Ahr.

In a recent study, Dere et al. [23] integrated ChIP-chip and transcriptional data of murine liver tissue after TCDD exposure. Interestingly, over 70% of the genes we predicted by our approach as primary Ahr responders were also identified in their study to be located in regions of Ahr enriched binding. More importantly, eleven out of the twelve transcriptional regulators identified by our method were also found in such Ahr enriched binding regions. This not only underlines the quality of our Random Forest classifier, but suggests a more general transcriptional network initiated by Ahr, independent of the activating ligand.

Ahr has been connected to hormone-induced signaling as was reinforced by our GO enrichment analysis that identified “regulation of hormone levels” as one of the biological functions. Crosstalk with the estrogen receptor has been studied extensively [24-26] and glucocorticoid receptor (GR)/Ahr crosstalk has also been suggested [27,28]. Our classifier predicted the glucocorticoid receptor (Nr3c1) itself as an Ahr target together with Sgk1, a GR-regulated kinase. In addition, the TF Klf9, known to be induced by GR and involved in adipogenesis, was predicted to be a direct Ahr target. Besides Klf9, further Ahr targets were predicted with an involvement in lipid synthesis and lipid transport, i.e. the transcriptional regulators Ppard and Lpin play a role in mammary lipid synthesis, and Npc1, Osbpl2, and Pitpnc1 are involved in lipid transport. The role of GR in lipid homeostasis and metabolism is well-established [29-31]. From our analysis we can deduce a possible Ahr-activated network of genes directly influencing lipid status and its regulation by the glucocorticoid receptor.

The interaction of Ahr with another TF Nfe2l2 (aka Nrf2) might also have an influence on lipid status, specifically on adipogenesis [32]. A bidirectional regulation of
these two pathways has been described previously [33]. Both TFs have been shown to bind in the other’s promoter region, thereby directly influencing transcription [32,34]. Therefore, the prediction of Nfe2l2 being an Ahr target is very well corroborated by previous studies and was indeed verified by our experimental follow-up. In addition, a recently described interaction of Nfe2l2 and Ahr confirms one other predicted Ahr target gene: Abcc4. Xu et al. showed that this multidrug resistant protein is directly activated by Ahr and Nfe2l2 in liver [35].

In our analysis we were only able to reliably classify 1,389 of the 2,338 regulated genes as either primary Ahr targets or as genes responding to BPDE stress. We found that the unclassified genes were enriched (P = 0.019) for genes perturbed by an Ahr mutation. A possible explanation for this enrichment is that there are genes that are downstream targets of Ahr (e.g. via the other transcriptional regulators that are primary responders to Ahr; see Figure 1, panel C) among this set. Since the classifier was not trained on such examples of downstream Ahr targets, we expect that it would not reliably classify these genes.

Side effects
Genes in clusters 1 and 2 are predicted to be perturbed not as a result of Ahr regulation, but by the presence of the metabolite BPDE. This genotoxic metabolite of B[a]P is known to cause DNA damage by DNA-adduct formation [36,37]. DNA repair processes are initiated, followed by re-initiation of DNA replication (one of the eleven GO categories enriched in cluster 1). Further, many MAP kinases were differentially regulated, and all of them are members of clusters 1 or 2. The idea that MAP kinases are Ahr-independent is supported by Tan et al. [38], who could show that Ahr ligands could activate MAPKs independent of Ahr.

To further support the predictions of our classifier, we selected some representative genes from clusters 1 and 2 (Agg1, Anape1, Njkb, and Parp1) and measured their expression in response to exposure to B[a]P or BPDE in wild type (Hepa1c1c7) or mutant (tao BpRc1) cells (Additional File 1, Figure S8). These experiments demonstrate that BPDE causes differential expression with and without Ahr, while B[a]P only perturbs expression in the presence of Ahr, i.e. when metabolism of B[a]P to BPDE is made more efficient by a functional Ahr pathway. These results demonstrate, as predicted, that these genes are affected by the presence of BPDE and are not a primary response regulated by Ahr.

Utility of weighted clustering
One unique and desirable aspect of the type of learning approach applied here is a side effect of the learning process - the proximity measure. The RF proximity is a type of similarity measure between subjects (in this case genes), based on how often two genes take the same path down the decision trees of the forest. It is in effect a weighted similarity measure because only time points that are useful in the learning process are used in the calculation of the proximity. This is in contrast to the widely used Euclidean distance or Pearson correlation, in which all features make an equal contribution.

A weighted (dis)similarity measure is advantageous in clustering gene expression time series, especially in complex transcriptional responses of higher eukaryotes, as presented in this work. Additional systems are present in higher eukaryotes that influence the synthesis, stabilization, and degradation of mRNA. These additional systems make it less likely that functionally related genes share precisely the same characteristic expression profile over time. For instance, functionally related genes, induced by a common TF, may share similar expression patterns shortly after induction, but may then diverge as other factors come into play, such as microRNAs. A supervised, weighted metric such as RF proximity de-emphasizes the diverging time points while emphasizing the common phase of induction, resulting in the grouping of the functionally related genes. Conversely, such expression profiles are unlikely to fall into the same cluster when using e.g. the Euclidean distance, and could be a contributing factor to the mixed success of past attempts [3,15-17] to cluster Ahr-induced gene expression time courses in a way that is biologically interpretable.

One technique that is frequently used to address problems such as those described here is biclustering [39-43]. Briefly, biclustering is a strategy that seeks to cluster in two dimensions simultaneously, e.g. genes and time points. The goal is to find genes that show similar expression in some (though not necessarily all) conditions. There are many algorithms and heuristics that implement biclustering. Strengths and weaknesses of the approach vary by implementation, but in general most biclustering methods are unsupervised and are non-deterministic. Without alleviating assumptions it can become a computationally intractable problem. It can be difficult to judge the quality of the resulting clusters, and clusters are often redundant. In the work presented here, clustering with the RF proximity presented fewer potential pitfalls compared to biclustering, since we had a means of performing supervised learning and the RF proximity was obtained “for free” since it was part of the learning process. In addition, the clusters were non-redundant and judging their quality was fairly straightforward by using another Random Forest to predict the assigned cluster labels of the genes (as described in the methods section). In addition to the work presented
here, clustering with an unsupervised RF proximity has been described in Shi and Horvath [44], and an example using multivariate response Random Forests to examine transcriptional programs in yeast can be found in Xiao and Segal [45]. We have found that PAM clustering with the RF proximity measure works well in scenarios where weighted clustering is desirable, and is an alternative to biclustering that is worth considering. However, one obvious limitation for any supervised method - including our use of RF here - is the need for a training set. In some situations a training set may be difficult or impossible to assemble - this is an important consideration when selecting a clustering method.

**Conclusion**

We explored the time-resolved transcriptional response induced by exposure to the environmental pollutant B[a]P and mediated by the transcription factor Ahr. As with many microarray experiments involving cellular stress, we observed an immense degree of differential expression, which often complicates biological interpretation. However, by using machine learning approaches, we successfully teased apart the specific, receptor-driven transcriptional response from the more general toxic response. Genes predicted to be part of a primary receptor-driven response were validated by extensive experimental work, further supporting the predictive power of our classifier. In addition to the specific results that further characterize the Ahr regulatory battery, our work here offers a useful strategy for distinguishing receptor-dependent responses and side effects based on expression time courses.

**Methods**

**Cell culture and sample preparation**

Murine hepatoma cells, Hepa1c1c7 as well as the mutant tao BpRc1 cells (both LG Standards GmbH, Wesel, Germany), deficient in endogenous Ahr, were used for all experiments. Cells were cultured in phenol red-free DMEM supplemented with 7% FCS, 1% glutamine and 1% penicillin/streptomycin. Ahr translocation was investigated in a stable cell line based on tao BpRc1 cells expressing GFP-Ahr under tetracycline control. Cells were stimulated with different concentrations of benzo-[a]-pyrene (B[a]P; Sigma Aldrich, Steinheim, Germany), BPDE (Midwest Research Institute, NCI Chemical Repository, Kansas City, MO, USA) and TCDD (Sigma-Aldrich, Steinheim, Germany) dissolved in DMSO respectively.

**Microarrays**

To investigate the differential kinetic behavior of the transcriptome after B[a]P exposure, and to identify the primary Ahr response, we used two different setups: (1) short term exposure, Hepa1c1c7 cells were treated with 50 nM B[a]P for 0, 1, 2, and 4 hours and (2) long term exposure, Hepa1c1c7 cells were treated with 50 nM or 5 μM B[a]P for 2, 4, 12 and 24h. Corresponding time-matched vehicle controls were generated. All experiments were performed in triplicate. Cells were lysed in Trizol reagent (Invitrogen, Darmstadt, Germany) and RNA extracted using RNAeasy kits (Qiagen, Valencia, CA, USA). RNA was quantified and integrity verified on a Bioanalyzer (Agilent Technologies, Palo Alto, CA). Sample preparation for Affymetrix GeneChip Mouse Exon 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA) was performed following the manufacturer’s recommendations. Microarray data was deposited in the Gene Expression Omnibus (GEO) under the identifier GSE29188.

**Detection of differential expression**

Microarrays were normalized using RMA and the University of Michigan custom CDF file (version 12.1.0) with mappings to Ensembl exon IDs. After normalization, but before proceeding with the analysis, we subtracted the (log2) DMSO expression values from the corresponding time point and batch of each of the B[a]P treatments. Exon expression values were then summarized to their corresponding Ensembl gene IDs, with the summarized gene expression value being the mean of its constituent exons. A 2-way ANOVA analysis was performed on each gene, with time and concentration as the factors. We then corrected for multiple testing by using the FDR. We considered only genes with an FDR <0.05 for any of the main effects or time-concentration interaction. In addition, we admitted genes with an FDR <0.05 from a simple t-test each B[a]P concentration (all time points pooled) vs. DMSO. Of these, we only considered genes that achieved 2-fold (or greater) differential expression at at least one time point. This left us with a total of 2,338 genes regulated in the long-term exposure (24 hour) data set. We interpolated the expression between the measured time points by averaging the simple linear interpolation with the spline interpolation. Since we have no measurement at time 0 hours, we assume equivalent expression with the DMSO samples, i.e. the expression ratio at time 0 hours is 0 on the log2 scale. The interpolation gave us a total of 25 values per gene, 1 value every hour from 0 to 24 hours. Whether or not the expression values were interpolated did not significantly affect the results of the classification and clustering, but we opted to use interpolated values to aid in visualization and interpretation.

**Classification with Random Forests**

We used the R implementation of Random Forests [46] to perform the two-class classification (Ahr primary
response vs. side effects), using the time course expression measurements of significantly regulated genes as predictors. To derive training labels (Additional File 1, Figure S2), we used data available from two BPDE studies in human cell lines [47,48], combining the P values from the studies using Fisher’s method. We labeled mouse orthologs of genes with BPDE-perturbed expression (FDR <0.05) as “secondary” since BPDE does not bind Ahr, but indicates affected genes further downstream of Ahr. We labeled genes as “primary Ahr” that showed differential expression (FDR <0.05) in an independent gene expression time course of cells exposed to 50 nM B[a]P from 0 to 4 hours, with the additional condition that they were not significantly regulated in the BPDE data (i.e. orthologs had FDR >0.05). These criteria led to 28 “Ahr-primary” labeled genes and 559 “side effect” labeled genes.

With this training set we ran RF with mtry set to 5, and ntree set to 5,000. We used the built-in outlier measure and removed genes in the 95th percentile of outlier scores (resulting in 27 primary response and 530 side effect training cases), then re-ran RF, this time with 1,000 trees. In both cases, to avoid biased predictions (since there are far more “secondary” samples) we randomly sampled 20 genes from each class for the construction of each tree in the forest. The overall misclassification rate for the final forest was 7% (out of bag error estimate).

Predictions were made for all 2,338 differentially expressed genes, and genes with a proportion of class votes greater than 80% were retained for further analysis. This cutoff was chosen because when the training labels were permuted randomly and a RF trained, no prediction had a proportion of votes greater than 80%. Using these criteria, a total of 82 genes were predicted to be responding to Ahr directly, and 1,365 genes were predicted to be side effects (e.g. regulated through the presence of B[a]P metabolites). In addition to predictions, the RF proximity measure was calculated for all significant and confidently classified genes, yielding a 1,447 by 1,447 matrix.

Clustering
The RF proximity matrix was used as a distance measure by the transformation $D = \sqrt{1 - P}$, where $P$ is the original proximity matrix and $D$ is the distance matrix. This distance matrix was then used as the input for PAM clustering, available in the R cluster package. We tested a range of $k$ values and found that specifying 3 clusters gave the best average silhouette.

To assess the degree of confidence in cluster assignment for each gene, an RF was fit to predict cluster label using the gene expression measurements. The proportion of votes for the correct cluster is an indication of how well a gene fits in the cluster. Genes that were given a lower proportion of votes for the correct class than expected under the null hypothesis (labels permuted randomly) were excluded. When including this additional filtering criterion, the final number of genes classified as primary responders was 81, with 1,308 genes as side effects. In addition, the importance measurements obtained in the construction of this RF give an indication of which time points and which concentrations are important parts of the cluster’s identity.

GO enrichment was performed for each cluster (Additional File 1, Table S3) using the topGO package [49]. Enrichment of the clusters for genes perturbed by an Ahr mutation was performed using the Kolmogorov-Smirnov test, using $P$ values derived from differential expression of genes from [7,19]. $P$ values were calculated for each study separately, then combined using Fisher’s method. Genes used to train the RF classifier were removed prior to calculation of enrichment, to ensure that the results reflected the actual predictive ability of the classifier.

Cell proliferation
Long-term exposure studies in Hepal1c1c7 cells treated with B[a]P versus BPDE were performed using the xCELLigence System (Roche Diagnostics, Mannheim, Germany). This system measures electrical impedance across micro-electrodes integrated on the bottom of 96-well tissue culture E-plates (Roche Applied Science, Germany). Shifts in impedance are measured in real time, indicating changes in cell proliferation. Cells were monitored every 15 min for up to 24 h after treatment with 50 nM, 500 nM, 1 μM, 2.5 μM or 5 μM of B[a]P or BPDE respectively. Each experiment was performed in triplicate.

qPCR
In a separate experiment Hepal1c1c7 and tao BpRc1 cells were exposed to B[a]P (50 nM, 5 μM), BPDE (50 nM, 5 μM) and 1 nM TCDD for 0.5, 1, 2, and 4 h. mRNA was extracted and isolated using the MagNA Pure LC System (Roche Diagnostics GmbH, Mannheim, Germany). 50 ng of mRNA was reverse transcribed according to the protocol provided with the AMV reverse transcriptase (Promega, Madison, WI, USA). Resulting cDNA was diluted 1:5 and 4 μl of template used in a 12 μl PCR reaction. qPCRs were performed for the following example genes: Tnfaip2, Tiparp, Cdkn1a, Cdkn1b, Mpp2, Cyp2s1, Nfe2l2, Klf9, Lig3, Myst2, Axin2, Afgf4, Anapc1, Nfkb1, Parp1, and the housekeeping genes 18S rRNA and Gapdh (primer sequences, Additional File 1, Table S1). All qPCR experiments were carried out on a
LightCycler®480 system (Roche Diagnostics GmbH, Mannheim, Germany) with the following settings: touchdown amplification with an initial step of 96°C for 10 min; followed by the first cycle at 95°C for 10 sec. The annealing step started at 68°C for 20 sec (decrease of 0.5°C/cycle with a step delay of 1 cycle) and reaching the annealing temperature of 58°C for the last 25 cycles, followed by 72°C for 20 sec for extension. A total of 45 cycles were performed in each experiment.

**ChIP**

Hepa1c1c7 cells were exposed to 50 nM B[a]P or DMSO as the vehicle control for 1 h. Subsequently, cells were exposed to 50 nM B[a]P or DMSO as the vehicle control for 1 h respectively. Subsequently cells were cross-linked for 10 min at 37°C in 1% formaldehyde followed by a quenching step for 10 min with 150 mM glycine. After cross-linking, chromatin DNA was sheared into 200-500 bp fragments by sonication using a Bioruptor® Next Gen (UCD-300, Diagenode SA, Liege, Belgium). Sonicated, soluble chromatin was immune-precipitated with 2.5 μg of an anti-AhR antibody (Enzolifesciences/Biomol, Lörrach, Germany) or anti-Pol II (Millipore, Billerica, MA, USA). Control IPs were performed using rabbit IgG (Millipore, Billerica, MA, USA) corresponding to our specific antibodies. DNA isolates from immunoprecipitates were used as templates for real-time quantitative PCR amplification using the primer pairs listed in Additional File 1, Table S2. All ChIP experiments were performed at least two times.

**AhR translocation**

Stably transfected tao BpRc1c cells, expressing a GFP-tagged AhR under tetracycline control, were used to investigate the differences in translocation behavior for different concentrations of B[a]P. Cells were seeded in 96-well imaging plates (BD, Franklin Lakes, NJ, USA) and taken off tetracycline 24 h before exposure to allow for sufficient GFP-AhR expression. Final B[a]P concentrations were 50 nM and 5 μM respectively, including a corresponding DMSO control (0.05%). After treatment, cells were fixed using 3.7% formaldehyde, and the nuclei stained with Hoechst 33342 (Invitrogen, Darmstadt, Germany). Imaging was performed on a BD Pathways™-mager 855 in a non-confocal mode using a 20X U-Apo 340 objective (Olympus, NA 0.75). Images were binned 2 × 2 and montaged 2 × 2. Further analysis of fluorescence intensity was performed using the Attovision software (BD, Franklin Lakes, NJ, USA). After segmentation of the nucleus and the cytoplasm, the ratio of the nuclear and cytoplasmic fluorescence was calculated for each cell. Ratios were confined in 0.01 intervals and relative frequencies determined. To allow for comparability the measurements were standardized so that the mean of the negative control equals 1. For the statistical analysis, more than 250 cells/treatment were considered.

### Additional material

**Additional file 1: supplementary information** A PDF containing additional details on the experiments and analysis.

**Additional file 2: cluster assignment and regulatory predictions for differentially expressed genes** An XLS file containing the results of our RF classifier.

### Acknowledgements

This publication is based on studies performed as part of our project “From contaminant molecules to cellular response: system quantification and predictive model development” funded by the Helmholtz Alliance on Systems Biology. F. Dautel and J. Mai were supported by the Helmholtz Young Investigator Group “Interdisciplinary Graduate School for Environmental Research (IGRADE). Technical support by the UZH/ETH FG C2 Functional Genomics Centre for microarray processing is greatly appreciated.

### Author details

1. Cellular Networks and Systems Biology, Biotechnology Center, TU Dresden, Dresden, Germany. 2. Dept. of Environmental Immunology, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany. 3. EAWAG Aquatic Research, ETH Swiss Federal Institute of Technology, Zurich, Switzerland. 4. Dept. of Proteomics, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany. 5. Dept. of Computational Hydro Systems, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany. 6. EPF Lausanne, Lausanne, Switzerland. 7. Dept. of Metabolomics, UFZ, Helmholtz Center for Environmental Research, Leipzig, Germany.

### Authors’ contributions

JM and ST wrote the manuscript. ST coordinated and performed the experimental work. JM performed the computational work. FD and DM performed cell culture experiments in connection with the microarray data. Validation experiments were performed by SR (qPCR, ChIP) and C. Gräbsch (qPCR). JMa performed image analysis. KS, IL, MvB, SA, and AB conceived of the experimental design and functioned in an advisory capacity. All authors approved the final manuscript.

### Received: 19 July 2011 Accepted: 13 October 2011 Published: 13 October 2011

### References

6. Carney SA, Chen J, Burns CG, Xiong KM, Peterson RE, Heideman W: Aryl hydrocarbon receptor activation produces heart-specific transcriptional


doi:10.1186/1471-2164-12-502

Cite this article as: Michaelson et al: Transcriptional signatures of regulatory and toxic responses to benzo-[a]-pyrene exposure. BMC Genomics 2011 12:502.
Acknowledgements

I would like to express my gratitude to my advisor Kristin Schirmer for the opportunity to perform this thesis under her supervision. I am very thankful for all that you thought and for the scientific freedom that you gave me. I am also grateful to Dr. J. Legler, Prof. Dr. M. Ackermann and Prof. Dr. R. I. Eggen to accept being part of my committee.

My appreciation goes to Anita Hintermeister, Christina Otto, Nadine Bramaz and Hana Mestankova for the amazing time that we had not only in the lab but also in “real life”. I am also thankful for my lab colleagues for the constant challenges. To my amazing office mates, Mirjam, Regine, Judith, Flavio, Nynke and Nadia for an incredible time together.

Many thanks goes to my “close by” friends, Carolina, Dimos, Dilum, Johannes, Hana, Irène, Mirjam, Nadine, Ruedi, Viviane and my two favourite boys in the world, Christian and Flavio, for these amazing four years. It was indeed an experience. Also to my “far-off” friends that even the distance cannot make it apart. Cristina and Ilana, you guys are always in my heart. To my adorable flatmates, Stefan, Christian and Peter for a fun home. A special thanks goes to my work colleagues at UTOX and at Eawag specially to Bettina, David, Smitha and Samy for the advices and imperturbability during this time. I also would like to thank Frederik Weiss, Dr. S. Sturla, Dr. P. van Midwoud, Dr D. Helbling and Dr K. Groh for the help during this work.

I would like to express my never ending gratitude for the professors and colleagues that passed through my life. Life is a never ending journey and you made me a better person.

I am eternally in doubt to my whole family, specially my parents, Carla and Noé, my sister and my brothers for all the love, support and joy that you gave me all these years. My love also goes to my aunt Lucy and uncle Glauco for all the funny and serious times we had, especially in the past four years. To my dearest Thomas thank you for your patience and love throughout. Your never ending support gave me serenity to finish my PhD.

Finally, special thanks go to the Helmholtz Association, Alliance on System Biology and Eawag for their financial support.
**Curriculum vitae**

<table>
<thead>
<tr>
<th>Year</th>
<th>Degree/Role</th>
<th>Institution/Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008 - 2012</td>
<td>Ph.D.</td>
<td>ETH, Zürich, Switzerland</td>
</tr>
<tr>
<td></td>
<td>Ph.D. thesis</td>
<td>Eawag - Swiss Federal Institute of Aquatic Science and Technology, ETH - Zurich</td>
</tr>
<tr>
<td>2006 - 2008</td>
<td>Master</td>
<td>State University of Campinas (UNICAMP), Campinas, Brazil</td>
</tr>
<tr>
<td></td>
<td>Master thesis</td>
<td>Expression analysis of genes that encodes for transporter proteins in <em>Acidithiobacillus ferrooxidans</em> and <em>Acidithiobacillus thiooxidans</em> kept in the presence of different metal sulphide</td>
</tr>
<tr>
<td>2000 - 2004</td>
<td>Bachelor</td>
<td>Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil</td>
</tr>
<tr>
<td></td>
<td>Bachelor thesis</td>
<td>DNA sequencing of mobilization genes from pRJ6 and its oriT regions from <em>S. aureus</em></td>
</tr>
<tr>
<td>1997 - 1999</td>
<td>High School</td>
<td>Rio de Janeiro (RJ), Brazil</td>
</tr>
<tr>
<td>1987 - 1996</td>
<td>Primary School</td>
<td>Rio de Janeiro (RJ), Brazil</td>
</tr>
</tbody>
</table>