UMEÅ UNIVERSITET

Computational methods for analyzing dioxin-like compounds and identifying potential aryl hydrocarbon receptor ligands

-multivariate studies based on human and rodent in vitro data

Malin Larsson

Department of Chemistry, 2017
Umeå University

Dissertation for PhD
ISBN: 978-91-7601-736-4
The electron density distributions of the highest occupied molecular orbital for a selection of dioxin-like compounds (Paper I). The calculations were performed using Gaussian 09 suite of programs. Reprinted with permission from Taylor & Francis.
Electronic version avaliable at: http://umu.diva-portal.org/
Printed at the KBC Service Center, Umeå University
Umeå, Sweden 2017

Nothing in life is to be feared,
it is only to be understood

–Marie Curie

Abstract

Polychlorinated dibenzo-*p*-dioxins/dibenzofurans (PCDD/Fs) and polychlorinated biphenyls (PCBs) are omnipresent and persistent environmental pollutants. In particular, 29 congeners are of special concern, and these are usually referred to as dioxin-like compounds (DLCs). In the European Union, the risks associated with DLCs in food products are estimated by a weighted sum of the DLCs' concentrations. These weights, also called toxic equivalency factors (TEFs), compare the DLCs' potencies to the most toxic congener, 2,3,7,8-tetrachloro-dibenzo-*p*-dioxin (2378-TCDD). The toxicological effects of PCDD/Fs and PCBs are diverse, ranging from chloracne and immunological effects in humans to severe weight loss, thymic atrophy, hepatotoxicity, immunotoxicity, endocrine disruption, and carcinogenesis in rodents.

Here, the molecular structures of DLCs were used as the basis to study the congeneric differences in *in vitro* data from both human and rodent cell responses related to the aryl hydrocarbon receptor (AhR). Based on molecular orbital densities and partial charges, we developed new ways to describe DLCs, which proved to be useful in quantitative structure-activity relationship modeling. This thesis also provides a new approach, the calculation of the consensus toxicity factor (CTF), to condense information from a battery of screening tests. The current TEFs used to estimate the risk of DLCs in food are primarily based on *in vivo* information from rat and mouse experiments. Our CTFs, based on human cell responses, show clear differences compared to the current TEFs. For instance, the CTF of 23478-PeCDF is as high as the CTF for 2378-TCDD, and the CTF of PCB 126 is 30 times lower than the corresponding TEF. Both of these DLCs are common congeners in fish in the Baltic Sea. Due to the severe effects of DLCs and their impact on environmental and human health, it is crucial to determine if other compounds have similar effects. To find such compounds, we developed a virtual screening protocol and applied it to a set of 6,445 industrial chemicals. This protocol included a presumed 3D representation of AhR and the structural and chemical properties of known AhR ligands. This screening resulted in a priority list of 28 chemicals that we identified as potential AhR ligands.

List of Papers

This thesis is based on the following papers:

I M. Larsson, B. Kumar Mishra, M. Tysklind, A. Linusson, and P.L. Andersson. *On the use of electronic descriptors for QSAR modelling of PCDDs, PCDFs and dioxin-like PCBs.* SAR and QSAR in Environmental Research, 24(6): 461–479, 2013.

II M. Ghorbanzadeh, K.I. van Ede, M. Larsson, M.B.M. van Duursen, L. Poellinger, S. Lucke-Johansson, M. Machala, K. Pencikova, J. Vondracek, M. van den Berg, M.S. Denison, T. Ringsted, and P.L. Andersson. *In Vitro and in Silico Derived Relative Effect Potencies of Ah-Receptor-Mediated Effects by PCDD/Fs and PCBs in Rat, Mouse, and Guinea Pig CALUX Cell Lines.* Chemical Research in Toxicology 27(7): 1120–1132, 2014.

III M. Larsson, M. van den Berg, M.B. van Duursen, K.I. van Ede, C. Lohr, S. Luecke-Johansson, M. Machala, S. Neser, K. Pěnčíková, L. Poellinger, D. Schrenk, S. Strapáčová, J. Vondráček, and P.L. Andersson. *Consensus Toxicity Factors for Polychlorinated Dibenzo-p-dioxins, Dibenzofurans, and Biphenyls Combining in Silico Models and Extensive in Vitro Screening of AhR-Mediated Effects in Human and Rodent Cells.* Chemical Research in Toxicology 28(4): 641–650, 2015.

IV M. Larsson, D. Fraccalvieri, C.D. Andersson, L. Bonati, A. Linusson, and P.L. Andersson. *Identification of potential aryl hydrocarbon receptor ligands by virtual screening of industrial chemicals.* Manuscript.

Author contributions:

I The author contributed extensively to the planning of the paper and the development of the electronic descriptors for DLCs. Moreover, the author performed the laboratory experiments, the calculations of 2D and semi-empirical 3D descriptors, the descriptor analyses, and the development of the quantitative structure-activity relationships (QSARs). The author had the main responsibility for writing the paper.

II The author was heavily involved in the planning of the project, did the initial QSAR modeling, and had the main responsibility for the descriptors and the interpretation of the QSAR models. The data evaluation was a joint effort among the authors.

III The author contributed extensively to the planning and execution of the paper, and the author performed the *in vitro* data evaluation and analyses (univariate analysis and principal component analysis), obtained the solution for the consensus scores development, and developed the QSARs. The author had the main responsibility for writing the paper.

IV The author contributed extensively to the planning of the paper, including the virtual screening procedure and its execution. The author performed all steps of the procedure, created the cut-offs for the initial filtration step and the ligand-based methods, made the structural comparisons to known AhR ligands for the outcome of all steps, and did the underlying literature searches and analysis of the *in vitro* data. The author had the main responsibility for writing the paper.

List of Abbreviations

AhR	Aryl hydrocarbon receptor
AM1	Austin Model 1
B3LYP	Becke–3–Lee–Yang–Parr
CAS	Chemical Abstracts Service
CB	Chlorinated biphenyl
CTF	Consensus toxicity factor
CYP1A1	Cytochrome P450, family 1, subfamily A, polypeptide 1
CYP1B1	Cytochrome P450, family 1, subfamily B, polypeptide 1
DD	Dibenzo-*p*-dioxin
DF	Dibenzofuran
DFT	Density function theory
DLC	Dioxin-like compound
ESP	Electrostatic potential
HF	Hartree–Fock
HOMO	Highest occupied molecular orbital
LUMO	Lowest unoccupied molecular orbital
MLR	Multiple linear regression
MO	Molecular orbital
mRNA	Messenger-ribonucleic acid
NCBI	National Center for Biotechnology Information
OECD	The Organization for Economic Co-operation and Development
OPLS	Orthogonal PLS
PBDD	Polybrominated dibenzo-*p*-dioxin
PBDE	Polybrominated diphenyl ether
PBDF	Polybrominated dibenzofuran
PC	Principal component
PCA	Principal component analysis
PCN	Polychlorinated naphthalene
PCB	Polychlorinated biphenyl
PCDD/Fs	Polychlorinated dibenzo-*p*-dioxins/dibenzofurans
PLS	Partial least squares projection to latent structures
QSAR	Quantitative structure-activity relationship
REP	Relative effect potency
SAR	Structure-activity relationship

SNFA	Swedish National Food Agency
TEF	Toxic equivalency factor
UV	Ultraviolet
WHO	World Health Organization
WHO-TEF	World Health Organization established TEF (in 2005)

DCDD	Dichloro-dibenzo-*p*-dioxin
TriCDD	Trichloro-dibenzo-*p*-dioxin
TCDD	Tetrachloro-dibenzo-*p*-dioxin
PeCDD	Pentachloro-dibenzo-*p*-dioxin
HxCDD	Hexachloro-dibenzo-*p*-dioxin
HpCDD	Heptachloro-dibenzo-*p*-dioxin
OCDD	Octachloro-dibenzo-*p*-dioxin
TCDF	Tetrachloro-dibenzofuran
PeCDF	Pentachloro-dibenzofuran
HxCDF	Hexachloro-dibenzofuran
HpCDF	Heptachloro-dibenzofuran
OCDF	Octachloro-dibenzofuran
PCB 74	2,4,4',5-tetraCB
PCB 77	3,3',4,4'-tetraCB
PCB 81	3,4,4',5-tetraCB
PCB 126	3,3',4,4',5-pentaCB
PCB 169	3,3',4,4',5,5'-hexaCB
PCB 105	2,3,3',4,4'-pentaCB
PCB 114	2,3,4,4',5-pentaCB
PCB 118	2,3',4,4',5-pentaCB
PCB 123	2',3,4,4',5-pentaCB
PCB 153	2,2',4,4',5,5'-hexaCB
PCB 156	2,3,3',4,4',5-hexaCB
PCB 157	2,3,3',4,4',5'-hexaCB
PCB 167	2,3',4,4',5,5'-hexaCB
PCB 189	2,3,3',4,4',5,5'-heptaCB

Contents

1. Introduction

New chemical compounds are constantly being introduced into society, either in new products or as unintentional by-products. Some of these compounds find their way into the environment and have negative impacts on living organisms. To assess the risks associated with these compounds, their level of exposure to humans from the environment is related to their toxicological effects. Due to the vast number of new compounds reaching the market, there are large data gaps to be filled regarding the assessment of a compound's toxicological effects. One way of dealing with these data gaps is, when possible, to use computational approaches to predict hazardous effects [143]. If we could understand how the structure of a molecule has a particular effect in a biological system, we would have key information for assessing the likelihood that compounds containing such molecules will induce the same effect. Relating an effect to certain structural features is usually called a structure-activity relationship (SAR) or qualitative read-across [91, 128, 149]. The structure can also be used to capture the chemical and physical properties of the molecule. Relating these kinds of properties to a biological effect is usually called a quantitative-structure activity relationship (QSAR) [31, 69, 91, 148]. Another computational approach is virtual screening. In a virtual screening, many computational techniques can be combined in the effort to identify compounds of concern for human health [48, 81, 85].

Polychlorinated dibenzo-*p*-dioxins/dibenzofurans (PCDD/Fs) and polychlorinated biphenyls (PCBs) are examples of persistent organic pollutants. These compounds are bio-accumulative, which means that their concentration increases at higher levels in the food web. These pollutants are included in the global treaty called the Stockholm Convention, which lists compounds that, due to their serious environmental and health effects, should be banned from production or should have actions taken to reduce their risks to the environment and human health [152]. PCDD/Fs have no practical use in society, but are produced as by-products in, for example, combustion processes, while PCBs were used in the past for things like electronic insulators and sealants in building construction. Actions have been taken to reduce the emissions of PCDD/Fs, such as decreasing the use of chlorine for bleaching pulp and adjusting the procedures for waste combustion. The use of PCBs was heavily reduced in the 1970s, but many buildings still contain PCB sealants [152, 153]. Numerous actions have been taken, but due to the persistence of these pollutants and their continuous formation in society, exposure levels and toxicological effects need to be monitored. It will take many generations until the levels are no longer a concern for human health and the environment [13, 46, 72, 142]. The toxicological effects of PCBs and PCDD/Fs are diverse, but a selection of these

compounds share a toxicological pattern mediated through the aryl hydrocarbon receptor (AhR), a nuclear transcription factor present in vertebrates and many other organisms [35, 73, 123]. The subgroup of PCDD/Fs and PCBs that interact with the AhR are referred to as dioxin-like compounds (DLCs). To assess the risks of DLCs, rodent studies are often used along with epidemiological and occupational health studies [50, 55, 56, 62, 67, 87, 111, 120, 127]. Recently, questions have been raised regarding how applicable rodent information is for explaining DLCs' effects in humans. The latest reevaluation of the risk assessment procedures of DLCs concluded that there is a need for more research on assays based on human cells [123].

2. Aims and scope

For this thesis, DLCs were analyzed by using an extensive set of AhR-related *in vitro* data, including assays based on human and rodent cells. The overall objective was to develop and use computational approaches to increase our understanding of congeneric and species-specific differences in the biological activity of DLCs.

Electronic properties provide fundamental information for describing and assessing a molecule's interactions with other small molecules or with macromolecules like proteins [65], and in the case of DLCs these properties can describe DLCs' interactions with the AhR. Paper I provided novel ways of describing the electronic properties of DLCs in the search for common denominators that can explain their biological activity toward AhR. Papers II and III present species-specific differences and similarities in the activities of DLCs and PCBs based on a large battery of *in vitro* data. Paper III provides an analysis of human versus rodent cell responsiveness measured by 17 *in vitro* responses (see Section 3.4 of this thesis). Moreover, in Papers II and III, species-specific differences were studied by developing QSARs. These models predict the AhR-related activity of untested DLCs and facilitate the interpretations of the measured biological activities using the chemical descriptors derived in Paper I (see Section 3.1).

In Paper III, the main aim was to contribute to the reevaluation of the specific toxicity factors currently used in risk assessment of DLCs (see Section 3.3). This aim was addressed by developing a procedure that provides species-specific consensus toxicity factors based on *in vitro* human and rat responses from a large battery of tests. The development of computer-based procedures was also the focus in Paper IV in which we developed a virtual screening protocol with the aim of identifying industrial chemicals that could potentially interact with AhR and hence be of concern for society.

3. Background

This chapter gives an overview of how computational approaches can be used as a complement or replacement of *in vivo* and *in vitro* experiments in the evaluation of a compound's hazardous effects. This chapter also describes the context in which this type of approach has been applied in the work for this thesis. The main compounds involved—the DLCs—and their current risk assessment tools are defined and discussed. Furthermore, a brief background is given regarding how the AhR is involved in their toxicological outcomes. The chapter concludes with a description of the underlying techniques for the computational modeling used throughout this thesis.

3.1 Computational approaches as a complement or replacement for in vivo and in vitro testing

If one has a selection of compounds for which one would like to investigate their toxicity, then one can perform a variety of *in vivo* animal experiments after which biological endpoints (such as enzyme induction) or toxicological endpoints (such as carcinogenicity) can be analyzed [42, 59, 130]. In addition, *in vitro* experiments using cell-based systems can be used to investigate the toxicity. For instance, enzyme induction in cell-based systems can be seen as a biomarker related to the toxicity of the compounds. *In vitro* studies bring the toxicological testing down to a cellular level in which the applied concentration more directly causes the effect. Furthermore, *in vitro* testing can reduce the use of animals and thus reduce the time and costs for performing the experiment. Another possible way to further reduce time and costs is to use *in silico* methods. The term *in silico* was first used in public back in 1989 and was defined as "experiments solely carried out by a computer" [76]. In this thesis, *in silico* methods refer to computer-based methods used as a complement or replacement for *in vivo* or *in vitro* tests. The *in silico* approaches employed in Papers I–IV used the structures of the molecules as an indicator of their chemical properties. This information was then used to create models relating structural features or properties back to the biological activities of the compounds. A commonly used method for the latter is to develop QSARs (introduced in Chapter 1 of this thesis), and this method was employed in Papers I–III to predict *in vitro* responses to DLCs.

QSARs are regression or classification models that predict the biological activities of compounds based on their physical properties and chemical structures [39, 40, 64, 91, 97]. In these models, the biological response (Y) is modeled as a function

of the chemical structure (X). The X-data consist of descriptors for properties like molecular volume and hydrophobicity, and these descriptors can be of different levels of complexity. The simplest descriptors give information based on atom types, such as the number of halogen atoms or single bonds, and are referred to as one-dimensional (1D). However, to be able to convey more features of the structures and their chemical or physical properties, the descriptors are often based on the two-dimensional (2D) structure or the three-dimensional (3D) conformations of the structure [113]. Descriptors can also be derived from experimental data such as ultraviolet (UV) spectroscopy data, solubility measurements, gas chromatographic retention time, and dipole moment [7, 117, 118, 139]. In Paper I, descriptors were developed to reflect the differences in biological activity among DLCs focusing on the electronic properties provided by UV absorption spectroscopy and quantum-chemical calculations (see Chapter 4). Because QSAR models are empirical (based on observations), it is important to define their applications and limitations and to properly validate their performance as a model for making predictions of untested compounds. The Organization for Economic Co-operation and Development (OECD) has agreed on principles regarding what is required for a QSAR model to be used for regulatory purposes [148]. These principles comprise the use of a defined endpoint; an unambiguous algorithm; a defined domain of applicability; appropriate measures of goodness-of-fit, robustness, and predictivity; and, if possible, a mechanistic interpretation of the QSAR.

A QSAR could (depending on its applicability domain) be used to screen larger sets of chemicals for the modeled biological activity [131]. This can also be referred to as a type of virtual screening, which is, in its most general term, an *in silico* evaluation of a larger data set with the objective of identifying biologically active compounds regarding a specific target, usually a protein [47]. The concept is most commonly used in medicinal chemistry and is typically divided into ligand-based and structure-based virtual screening. The ligand-based approaches use the molecular structures of compounds that are known to bind to the protein in order to identify new potential binders. For example, structural fingerprints [134] provide information on the presence or absence of specific structural fragments, such as functional groups, that can be used to measure the structural similarity between a known binder and other compounds. Compounds can also be characterized by chemical descriptors and projected onto a multivariate space where new potential binders can be identified by their position relative to known binders. A common method is k-nearest-neighbors [112], which identifies the k closest compounds to a selected compound, in this case, a known binder.

In structure-based virtual screening, the studied compounds are docked into the structure of a macromolecule to identify plausible binding conformations. The different compounds are then ranked as potential ligands using scoring functions [23]. There are many types of scoring functions, e.g., energy-estimated, knowledge-

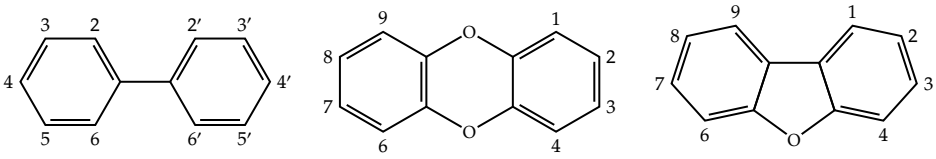


Figure 1: The general structural formulas and substitution positions of (left to right) polychlorinated biphenyls (PCBs), dibenzo-*p*-dioxins (PCDDs), and dibenzofurans (PCDFs).

based, and rule-based, and sometimes several of them are combined to get a comprehensive ranking of the molecular docking [6]. Crystal structures of the protein are often used for structure-based virtual screening. Because the X-ray crystal structure of AhR has not been determined, different homology models have been developed for use in protein-ligand interaction studies [24, 77, 70]. The homology models were built based on existing 3D-structures of similar proteins with the assumption that proteins with high similarity in their 2D amino acid sequence also adopt the same secondary structure, i.e., they fold in the same way. Paper IV presents a virtual screening protocol that combines ligand and structure-based approaches. In this protocol, an AhR homology model was used to assess the protein–ligand interactions between the rat AhR and potential ligands among a collection of small industrial chemicals.

3.2 Dioxin-like compounds (DLCs)

PCDD/Fs and PCBs are halogenated aromatic compounds, including a large range of congeners. These congeners differ from each other by the number of chlorine atoms substituted on the aromatic rings and their substitution pattern on the rings (Figure 1). Out of the theoretically possible 75 PCDDs, 135 PCDFs, and 209 PCBs, 29 congeners are referred to as DLCs (see Chapter 1) and these share toxic effects mediated through interaction with AhR and are of special concern for human health (see Section 3.3). Due to their high hydrophobicity, these compounds primarily partition to organic matter in soil and sediment and they accumulate in adipose tissues of organisms. DLCs bioaccumulate, and as a result they are found at high concentrations in top-level predators in the food chains, such as eagles, seals, and humans. Moreover, DLCs have been shown to have various negative impacts on organisms [56, 73, 90, 92, 135]. The toxicological effects of PCBs and PCDD/Fs are diverse, ranging from chloracne and immunological effects in humans to severe weight loss, thymic atrophy, hepatotoxicity, immunotoxicity, endocrine disruption, and carcinogenesis in rodents [92, 103, 122, 123]. The most potent and well-studied DLC is 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (2378-TCDD), which is a PCDD with four chlorine atoms positioned at the lateral carbons in the molecule (Figure 1).

Notably, all DLCs have chlorine atoms in four lateral positions on the rings, two on each side of the molecule. That is, chlorine atoms occupy the 2, 3, 7 and 8 positions of the PCDD/Fs and two adjacent positions among the carbons for the 3, 4, 5 and 3', 4', 5' positions of PCBs. In Paper I, the above-mentioned structural features were used as the foundation for developing novel descriptors for all DLCs (see Chapter 4).

3.3 The toxic equivalency factor (TEF) concept

The toxic equivalency factor (TEF) concept was developed during the 1980s as a method for risk assessment of PCDD/Fs [2, 147]. Today, TEFs provide a unique tool for assessing the toxic potency of mixtures of DLCs in food, feed, human populations, and wildlife [122, 123].

The TEF concept is based on the assumption of additivity [15, 96]. Each DLC is assigned a relative value that relates the DLC's potency to the potency of 2378-TCDD. A congener with a TEF value of 0.1 means that ten times as much of that DLC is needed to induce the same effect as 2378-TCDD. To calculate the total contamination in "2378-TCDD units" in a matrix, such as food, feed, or human blood, each concentration of a DLC in the matrix is multiplied by its TEF. After that, the products are summed up to become the total toxic equivalent (TEQ) of the matrix. Hence, a TEQ is calculated as

$$\mathrm{TEQ} = \sum_{i=1}^{N} c_i \mathrm{TEF}_i,$$

where c_i is the concentration of compound i in the matrix, TEF_i is the assigned TEF for compound i, and N is the number of compounds in the matrix. The estimated risk associated with DLCs is now expressed by its TEQ concentration.

TEQs are frequently used to assess and compare exposure levels of potentially contaminated sites or to monitor levels of DLCs in humans, but they are primarily intended for estimating exposure and risks via oral ingestion [123]. That is, the TEQs are intended to assess risks from exposure of DLCs in food, which is the most common route of exposure for the general public. The European Union (EU) [144] has established the tolerable daily intake of DLCs at 2 pg TEQ/kg body weight, and the EU also provides specific limits (also expressed in TEQ) for certain food items, such as fish, meat, milk, eggs, and baby food. The current EU limit for DLCs is 6.5 pg TEQ/g fresh weight of the food item [151], thus changing the TEF of a congener or assigning a TEF to a new compound could have great regulatory implications on the national level. For instance, it might allow higher or lower contamination levels in food products or lead to changes in the recommendations on how often certain food items should be consumed.

Table 1: The toxic equivalency factors (TEFs) for risk assessment of dioxin-like compounds (DLCs) assigned in 2005 at the expert meeting organized by the World Health Organization (WHO) [123].

	Compound	WHO-TEF
Chlorinated dibenzo-*p*-dioxins	2,3,7,8-TCDD	1
	1,2,3,7,8-PeCDD	1
	1,2,3,4,7,8-HxCDD	0.1
	1,2,3,6,7,8-HxCDD	0.1
	1,2,3,7,8,9-HxCDD	0.1
	1,2,3,4,6,7,8-HpCDD	0.01
	OCDD	0.0003
Chlorinated dibenzofurans	2,3,7,8-TCDF	0.1
	1,2,3,7,8-PeCDF	0.03
	2,3,4,7,8-PeCDF	0.3
	1,2,3,4,7,8-HxCDF	0.1
	1,2,3,6,7,8-HxCDF	0.1
	1,2,3,7,8,9-HxCDF	0.1
	2,3,4,6,7,8-HxCDF	0.1
	1,2,3,4,6,7,8-HpCDF	0.01
	1,2,3,4,7,8,9-HpCDF	0.01
	OCDF	0.0003
Non-*ortho*–substituted PCBs	3,3',4,4'-tetraCB (PCB 77)	0.0001
	3,4,4',5-tetraCB (PCB 81)	0.0003
	3,3',4,4',5-pentaCB (PCB 126)	0.1
	3,3',4,4',5,5'-hexaCB (PCB 169)	0.03
Mono-*ortho*–substituted PCBs	2,3,3',4,4'-pentaCB (PCB 105)	0.00003
	2,3,4,4',5-pentaCB (PCB 114)	0.00003
	2,3',4,4',5-pentaCB (PCB 118)	0.00003
	2',3,4,4',5-pentaCB (PCB 123)	0.00003
	2,3,3',4,4',5-hexaCB (PCB 156)	0.00003
	2,3,3',4,4',5'-hexaCB (PCB 157)	0.00003
	2,3',4,4',5,5'-hexaCB (PCB 167)	0.00003
	2,3,3',4,4',5,5'-heptaCB (PCB 189)	0.00003

In the beginning of their use in the 1980s, different TEF concepts that only included congeners among PCDD/Fs [2, 147] were developed in parallel. During the early 1990s, specific PCB congeners were reported to exhibit dioxin-like toxicity and were thus included in the TEF concepts [4, 92, 94]. In 1993, the World Health Organization (WHO) organized an expert meeting to establish TEFs for regulatory purposes [3]. Since then, the WHO has held another two expert meetings [122, 123] where the TEFs have been reevaluated and new compounds discussed as candidates for receiving a TEF. To be included in this TEF concept, a compound must: a) show structural similarity with PCDD/Fs, b) activate AhR, c) induce AhR-mediated effects (see Section 3.4), and d) be persistent and bioaccumulative [123]. The resulting TEFs from the most recent of these expert meetings are the ones currently being used in the EU-established limits for DLCs in food [144]. These TEFs are henceforth referred to in this thesis as the WHO-TEFs (Table 1).

To establish the WHO-TEFs, an extensive database with relative effect potency (REP) values for the various DLCs based on *in vivo* and *in vitro* data was used [51]. The REP of a DLC is the ratio between the effective concentrations of 2378-TCDD and the DLC. The statistical REP distributions were studied with no preferences for the used toxicological endpoint in each study, i.e., it was an unweighted distribution. That information was combined with point estimates from selected studies and expert judgment and included considerations of the environmental abundance of the compounds, especially regarding PCBs. Highest priority was given to *in vivo* studies of chronic exposure in rats. However, due to the very limited information provided in existing *in vivo* data, *in vitro* data were often used as a complement or as the only source of information for the reevaluation. When establishing the WHO-TEFs [123], *in vitro* data based on human cells were also discussed for certain DLCs. However, those data were not part of the statistical analysis of the distribution of toxic responses to DLCs because it was unclear how comparable human and rodent systems really are [123]. In general, it was concluded that more research is needed for assessing the REPs from human systems. In Paper III, a set of *in vitro* data concerning both human and rodent cell responses related to AhR was presented and analyzed. In addition, an alternative way to derive consensus values based on REPs from a battery of tests was presented and implemented so as to provide additional *in vitro*-based data for future reevaluations of the WHO-TEFs.

3.4 Aryl hydrocarbon receptor (AhR)

AhR is an evolutionarily old protein that is encountered in many different organisms such as mammals, birds, fish, insects, and worms, but the two latter groups do not respond to DLCs [89]. AhR has been found in many different tissues at different concentrations; rats have been shown to have the highest concentrations of

AhR in the lung, thymus, liver, and kidneys, whereas humans have been shown to have the highest concentrations in the placenta followed by the lung, heart, pancreas, and liver [27, 28, 38]. AhR was discovered in the 1970s to be a mediator of the measurable (adaptive) responses to DLC exposure. Since that time, AhR has been intensely studied regarding its mechanism of action and its interplay with other mechanistic pathways [35]. The mechanism of action of DLC-induced AhR activity is believed to be the same in all mammalian species and involves ligand activation. Recently, the activation of AhR has been linked to many immune response processes that are associated with various diseases [22, 25, 53, 105].

AhR is a ligand-activated transcription factor whose activation results in the transcription of a battery of genes, including the drug-metabolizing enzyme cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) [62]. In Paper III, different endpoints were investigated to analyze DLC exposure, and these endpoints included gene expression assays targeting CYP1A1; cytochrome P450, family 1, subfamily B, polypeptide 1 (CYP1B1); and the aryl-hydrocarbon receptor repressor. The studied endpoints also included luciferase reporter gene assays and CYP1A1 activity bioassays (see Section 6.1). The data from these experiments, which involved rat, mouse, guinea pig, and human cell systems and which targeted different tissues, were used to study potential species differences in response to DLCs. Induction of CYP1A1-related responses is the most studied and generally accepted biomarker for DLC exposure. Many structurally different compounds have been shown to induce CYP1A1-related responses and have been identified as potential AhR ligands [34, 79]. Ligands are compounds that have been shown to bind to the receptor in question but not necessarily to activate the receptor. AhR has been shown to cross talk with other nuclear receptors, which makes the situation even more complex [63, 82, 108]. In Paper IV, a distinction was made between AhR binders (i.e. ligands) and AhR modulators. The latter were defined as compounds that—in one way or another—suppress or enhance the AhR activity.

3.5 Chemometrics

Chemometrics is a scientific field that combines chemistry with multivariate statistics and mathematics to capture relevant information from a given data set [137]. In this thesis, the chemometric tools of principal component analysis (PCA) [57] and partial least squares projections to latent structures (PLS) [138] were used to capture trends and to find relationships within the data.

3.5.1 Principal component analysis (PCA)

PCA is a multivariate projection method commonly used in chemometrics to investigate how objects are related to each other in a data set. In this thesis, the objects for this data matrix, typically denoted $\mathbf{X}$, are different compounds, and the variables consist of the compounds' individual properties as reflected by a number of different descriptors. In PCA, this data matrix is projected onto a space that has as many dimensions as the number of variables in the data set. PCA simplifies the interpretation of the data by condensing the information in the original data set into a new set of variables called principal components (PCs). The first PC defines the direction of the largest variation in the data; that is, it gives the direction in which the objects differ from each other the most. The second PC captures the largest variation in the subspace orthogonal to the first PC. This is commonly referred to as the second largest variation. Mathematically, the original data $\mathbf{X}$ can be described by the three new matrices $\mathbf{T}$, $\mathbf{P}$, and $\mathbf{E}$ as

$$\mathbf{X} = \sum_{i=1}^{N} \mathbf{t}_i \mathbf{p}_i' + \mathbf{E},$$

where $\mathbf{t}_i$ is the ith column of $\mathbf{T}$ and also the score vector of the ith PC, $\mathbf{p}'$ is the transposed ith column of $\mathbf{P}$ and also the transposed loading vector of the ith PC, N is the number of PCs, and $\mathbf{E}$ is the residual matrix. The residual matrix $\mathbf{E}$ is what is left of the variation in the data when the N PCs have been extracted. The results from a PCA are shown as scores and loadings and are usually depicted in two-dimensional plots that are studied at the same time. For instance, one can put the scores of the first and second PCs, i.e. $\mathbf{t_1}$ versus $\mathbf{t_2}$, as one plot and the loadings of the same two PCs, i.e. $\mathbf{p_1}$ versus $\mathbf{p_2}$, as another plot. Scores and loadings are related to each other; the scores describe how similar (or dissimilar) the objects are, and the loadings describe which variables (i.e. properties) make them so. Moreover, objects close to each other in the score plot share similarities, while increased distance between objects means increased dissimilarity between objects. The loading plot is related to the score plot and shows the reasons for the observed similarity (or dissimilarity) between objects. For instance, objects that are located in a certain area of the score plot are characterized by the variables located in the same area in the loading plot.

3.5.2 Partial least squares projections to latent structures (PLS)

PLS is a method to identify linear relationships between different sets of data. Applied to QSARs, PLS can be used to study correlations between the measured biological response and the factors that influence the response. The data to describe these factors, i.e., descriptors, are stored in a matrix, $\mathbf{X}$. PLS can handle multiple response variables and this information is stored in another matrix, $\mathbf{Y}$. The same as

for PCA, directions in the multivariate space are calculated to follow the variation in the data in the $\mathbf{X}$ and $\mathbf{Y}$ matrices. However, in PLS the directions are computed to also maximize the covariance between $\mathbf{X}$ and $\mathbf{Y}$ [138]. For PLS models that only use one response variable, the response data is a vector, $\mathbf{y}$, but the procedure is the same. The PLS model can be expressed as:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{F},$$

where $\mathbf{B}$ is the PLS regression coefficients and $\mathbf{F}$ is the Y-residual matrix [138]. The most influential factors for the model have large positive or negative coefficients. Hence, the coefficients can be used to identify the factors that are of most importance for the biological response.

PLS can deal with co-linearity in the data, which means that PLS can model descriptors that are correlated. However, the more descriptors one includes in the model, the harder it is to interpret. To simplify the interpretation of the model, i.e., to more easily identify influential descriptors, orthogonal-PLS (OPLS) [114] can be used. A PLS and an OPLS model, based on the same input, give identical predictions. However, in OPLS the systematic, non-random variation in $\mathbf{X}$ is divided into two parts, one part that is related to $\mathbf{Y}$ and one part that is unrelated, i.e. orthogonal, to $\mathbf{Y}$. This gives the opportunity to further study what descriptors, and in the long run what features or properties, are important for the biological response.

4. Chemical characterization of DLCs with a focus on their electronic properties (Paper I)

The DLCs comprise three different chemical classes, where the structural differences within each class might be as small as moving one chlorine atom from one aromatic carbon to another (see Section 3.2). Nevertheless, these small changes can be of great importance for the biological activity of DLCs. To determine the effect of these seemingly small changes in the chemical structure, the DLCs studied in Paper I were characterized based on their electronic properties. We used 2D and 3D descriptors along with spectroscopic data to describe the electronic properties of the DLCs. The descriptor analysis using PCA showed that many of these descriptors are highly class dependent; that is, the descriptor value was more influenced by the chemical class of the DLC than by the substitution pattern of the DLC. To overcome this issue, we developed descriptors that are atom-specific and tuned to the lateral positions of the DLCs (Figure 1). With this as a basis, we calculated molecular orbital (MO) densities and partial charges of the lateral carbons (see Section 4.1). This approach generated descriptors that revealed new class-independent properties of DLCs. UV absorption (200–400 nm) was measured for all DLCs because PCBs and PCDFs have been shown to produce distinct absorption patterns within the UV range [7, 117]. UV absorption reflects electronic transitions within the aromatic system, i.e., UV reflects properties of the conjugated π-electron system, and our measurements showed that each DLC had a unique spectrum. The measured upper range (250–400 nm) was found to be class dependent, while we concluded that the lower range (200–250 nm) reflected the same transitions for all DLCs (see Section 4.3), and as such was class independent.

4.1 Molecular and novel atom-specific calculated electronic descriptors of DLCs

MO energies were calculated to estimate the reactivities of the DLCs. More specifically, the MO energies of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) were calculated, and from those the gap in energy between the HOMO and the LUMO energies (GAP) [113] was calculated. According to calculated MO energies, PCDFs are the strongest electron acceptors, i.e, they have the lowest LUMO energies, among the DLCs. Moreover, PCDFs and PCBs are stronger electron donors, i.e., they have lower HOMO energies, than PCDDs, and PCBs are the most stable chemical class among the DLCs based on their higher GAP values compared to the other DLCs. This high chemical class dependence suggests that the MO energies are not sufficient to

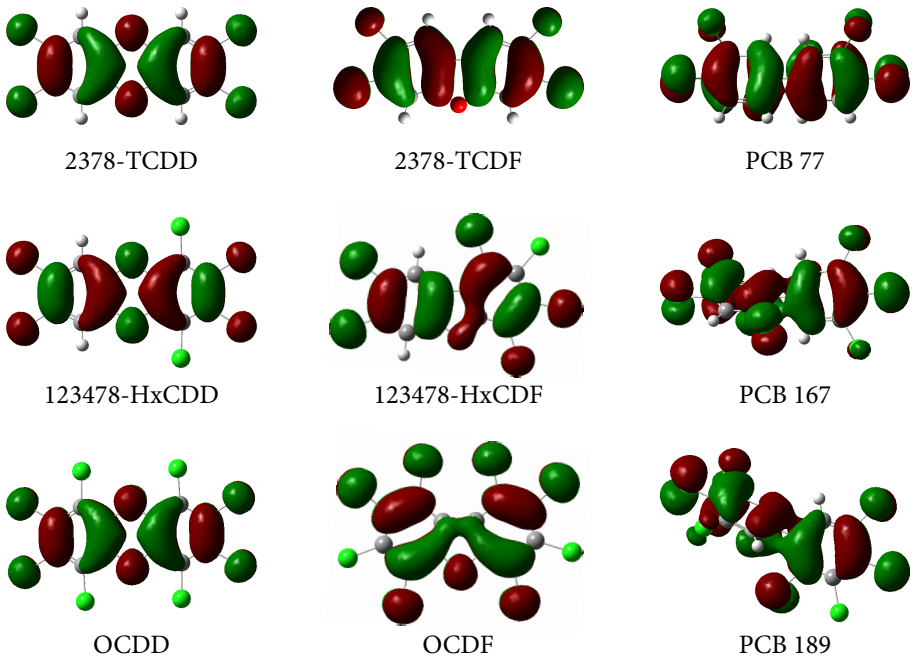
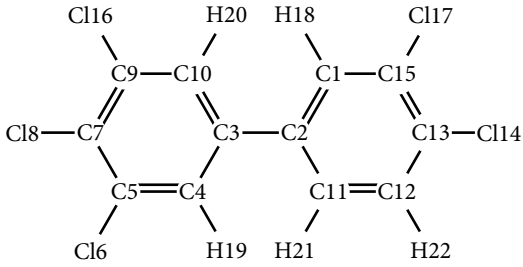


Figure 2: The electron density distributions of the HOMO for different congeners of PCDDs, PCDFs, and PCBs calculated by using the DFT hybrid functional B3LYP with the 6-31G** basis set. Further information on the corresponding molecular structures is given in the supporting information of Paper I.

explain the congener-specific differences seen in biological activity among DLCs. Because of this, we instead focused on the MO densities (Figure 2) for the lateral carbons of the DLCs. By focusing on just the densities of the MOs, a new property appears that reflects the symmetry of the MO density distribution (Figure 2). In practice, the electron densities for the lateral carbon positions of the DLCs are recorded and ranked for each molecule, and the highest density gives the highest ranking. The lateral carbon positions' ranking is used instead of the actual positions as descriptors because PCBs have six lateral positions compared to the four of PCDD/Fs (Figure 1).

Figure 3 shows the workflow for calculating the atom-specific descriptor Ha2/Ha1 for PCB 126. Ha2/Ha1 is the ratio between the second highest and the highest lateral carbon HOMO density. This descriptor gives high values for 2378-TCDD, 2378-TCDF, PCB 77, and PCB 169. All of these structures have a symmetrical plane (such as 2378-TCDD) or can achieve a symmetrical conformation (such as PCB 77), unlike for instance PCB 167 (Figure 2). However, there are DLCs that do not have a symmetrical plane but still have a symmetrical MO density distribution and thus also receive a high value for Ha2/Ha1. One example of this

atom number	HOMO density
C1	0.027
C2	0.126
C3	0.106
C4	0.045
C5	0.030
Cl6	0.011
C7	0.134
Cl8	0.060
C9	0.029
C10	0.045
C11	0.071
C12	0.019
C13	0.144
Cl14	0.065
C15	0.054
Cl16	0.011
Cl17	0.022
H18	0.000
H19	0.000
H20	0.000
H21	0.000
H22	0.000



Compound	Ha1	Ha2	Ha3	Ha4
PCB126	0.144	0.134	0.054	0.030

Compound	Ha2/Ha1	Ha3/Ha1	Ha4/Ha1
PCB126	0.930	0.375	0.208

Figure 3: To develop the molecular orbital density ratios based on the lateral carbon positions, the atom-specific densities were calculated and ranked. For PCBs, here illustrated by PCB 126, the four highest values were taken out and then compared to each other by ratios. The selected positions always corresponded to carbons also being attached to chlorine atoms.

is 123478-HxCDD (Figure 2). The HOMO density ranking of the lateral carbons also reflects the corresponding density ranking of the chlorine atoms. For example, the highest ranked lateral carbon of PCB 126 (C13) is directly connected to the highest-ranked chlorine atom (Cl14) (Figure 3).

In addition to the MO density ratios, the partial charges are also quite different from the other calculated 3D descriptors. This is illustrated in Figure 4 by their low covariance (descriptors 18, 23, 24, and 26). These charges were derived by calculating the electrostatic potential (ESP) [65] of the lateral carbons and by using the same ranking system as the MO density ratios. The second-highest atomic partial charges of the lateral carbons (QC_a2_ESP) are very high for non-*ortho* PCBs. That is, by calculating partial charges on the lateral carbons, it is possible to differentiate between non-*ortho* and mono-*ortho* PCBs, which is an important substitution pattern known to separate more and less potent PCBs (Paper I). Atom-specific descriptors based on MO densities and partial charges have previously been applied to predict the photolytic half-lives of polybrominated diphenyl ethers (PBDEs) [54], another chemical class consisting of many congeners. In contrast,

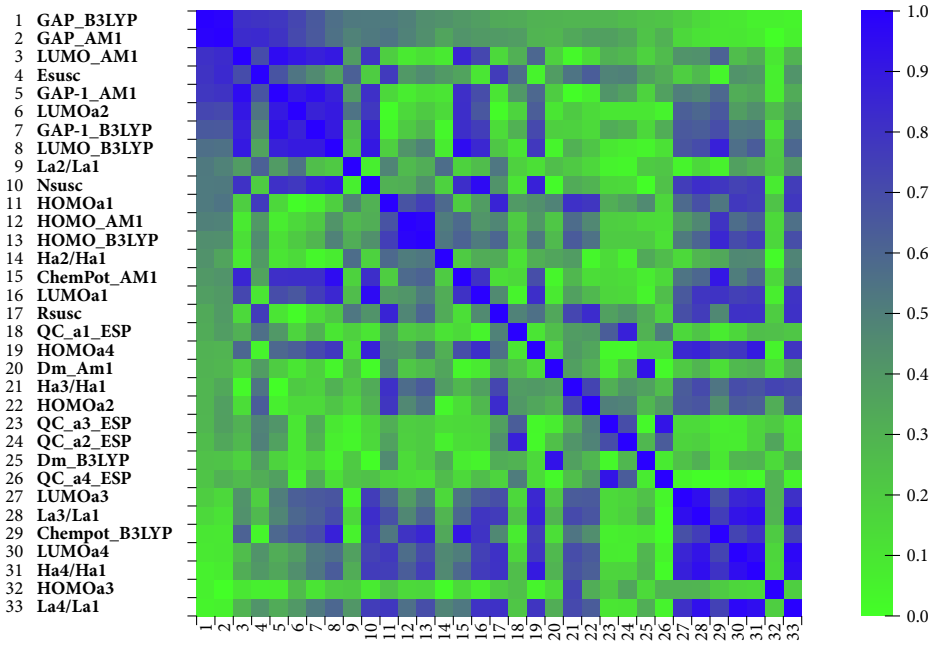


Figure 4: The covariance matrix of calculated molecular and atom-specific descriptors related to the electronic properties of the DLCs. The data were sorted by increasing covariance with the difference between the HOMO and LUMO energies (GAP) and normalized, i.e. they show the absolute values of each covariance entry. Further information on the descriptors is given in the supporting information of Paper I.

we compared congeners of different chemical classes and focused on a specific region of the molecules.

4.2 2D descriptors

The calculations using the 2D structures of the molecules showed that congeners within each chemical class of DLCs having the same number of chlorine atoms, i.e. homologues, usually received the same numerical value. That is, most of these descriptors did not differentiate between substitution patterns. Furthermore, many of the 2D descriptors correlated with size. For example, the size-related property van der Waals surface area (descriptor 1, Figure 5) was highly correlated with descriptors reflecting the hydrophobic surface area (descriptors 2 and 9), density (descriptor 16), hydrophobicity (descriptor 6), flexibility (descriptor 8), and polarity (descriptor 20). In general, the 2D descriptors differentiated between chemical classes of DLCs, but not between homologues.

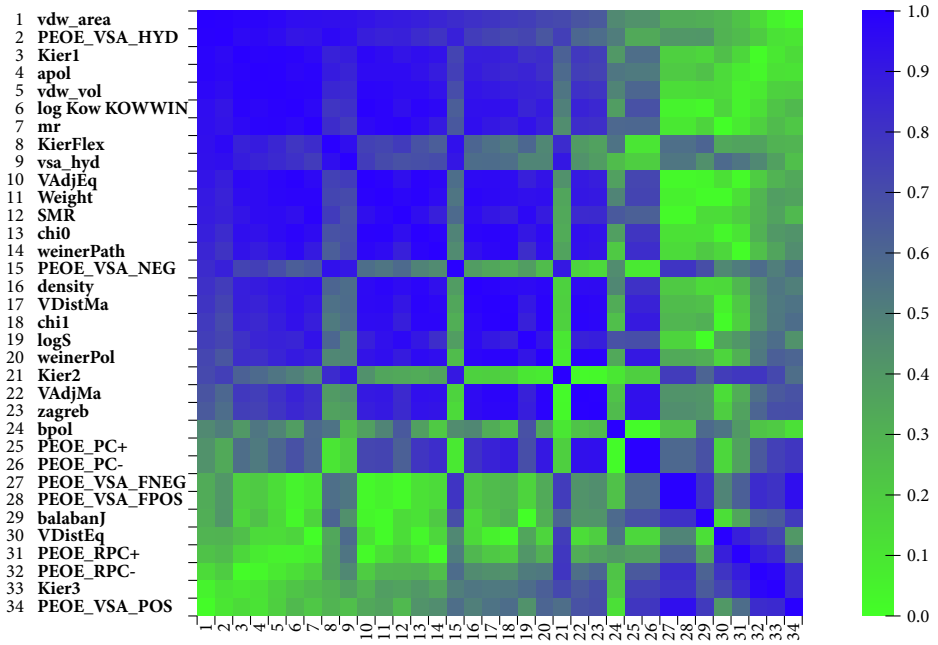


Figure 5: The covariance matrix of a selection of 34 2D descriptors regarding, for example, size, shape, polarity, hydrophobicity, and surface characteristics of the DLCs [146]. The data were sorted by increasing covariance with the size-related property van der Waals surface area (vdw_area) and normalized, i.e. they show the absolute value of each covariance entry. Further information on the descriptors is given in the supporting information of Paper I.

4.3 Ultraviolet absorption spectra and experimentally derived descriptors

In the lower wavelength region (200–250 nm), the UV absorption spectra of all measured DLCs showed clear smooth peaks (Figure 6), most likely originating from the same kinds of electronic transitions. Peaks in this region have been identified as electronic transitions from bonding MOs to antibonding MOs. That is, the peaks arise from p to p* transitions in the conjugated π-system [44]. Furthermore, the absorption at 200–215 nm reflects the resonance over the benzene rings [7, 32]. Regarding the higher wavelength region (250–350 nm), the UV absorption spectra are class dependent as indicated by our observations that the spectra included oxygen lone-pair excitations at roughly 300 nm for PCDD/Fs (e.g. 2378-TCDF and 2378-TCDD, Figure 6bc), that there was no absorption of PCBs above 300 nm (e.g. for PCBs 81 and 123, Figure 6ad), and that there were very peak-rich and specific patterns of PCDFs as illustrated by OCDF (Figure 6e). The greater complexity in spectral pattern among PCDFs was concluded to be due to the fact that PCDFs abide by Hückel's rule of aromaticity, which means that they only have one aromatic

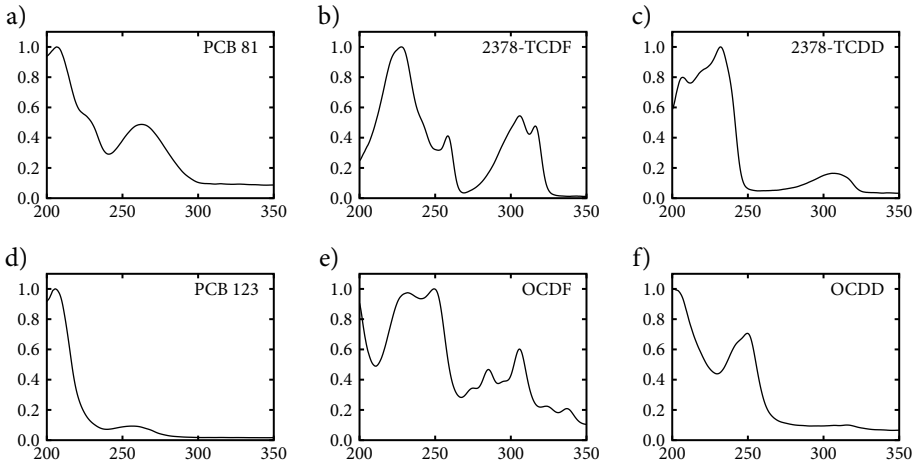


Figure 6: The ultraviolet absorption spectra of selected PCBs (a,d), PCDFs (b,e), and PCDDs (c,f). The structures of the compounds are given in the supporting information of Paper I.

system covering all rings in the skeleton, and thus have more potentially excited states [44].

Measuring the UV absorption (200–400 nm) for the DLCs generated partial spectral data that have not previously been accounted for in the literature. Tysklind et al. [117] measured all congeners of PCDFs, but they did not include the lower wavelength interval (200–275 nm). For our UV experiments in Paper I, we used isooctane as the solvent, and this enabled us to measure the lower wavelength interval (200–275 nm). Moreover, the PCDDs have to my knowledge not been extensively studied before, i.e., by addressing an extended number of substitution patterns and degrees of chlorination. Kende et al. [61] studied four PCDDs, but only reported information regarding one peak. They reported peak maxima located at 296 nm, 299 nm, 304 nm, and 306 nm for 13-DCDD, 23-DCDD, 237-TriCDD, and 2378-TCDD, respectively. Funk et al. [44] studied five DLC congeners—2378-TCDD, 12378-PeCDD, OCDD, 2378-TCDF, and OCDF—but those experiments were conducted in gas phase, which resulted in rather rough spectra in the 200–400 nm range.

The new measurements of all of the DLCs confirmed the peaks found by Funk et al. [44], and the high resolution of these experiments also revealed congener-specific peak patterns in the lower wavelength interval (200–250 nm) (Figure 6). In particular, for the region of 230–245 nm there was a specific two-peak pattern for the PCDFs where the relative magnitudes of the peaks varied with the substitution pattern. For OCDF these peaks were similar in magnitude (Figure 6e), while the first or second peak decreased if the congener had a 678- or 789-chlorination

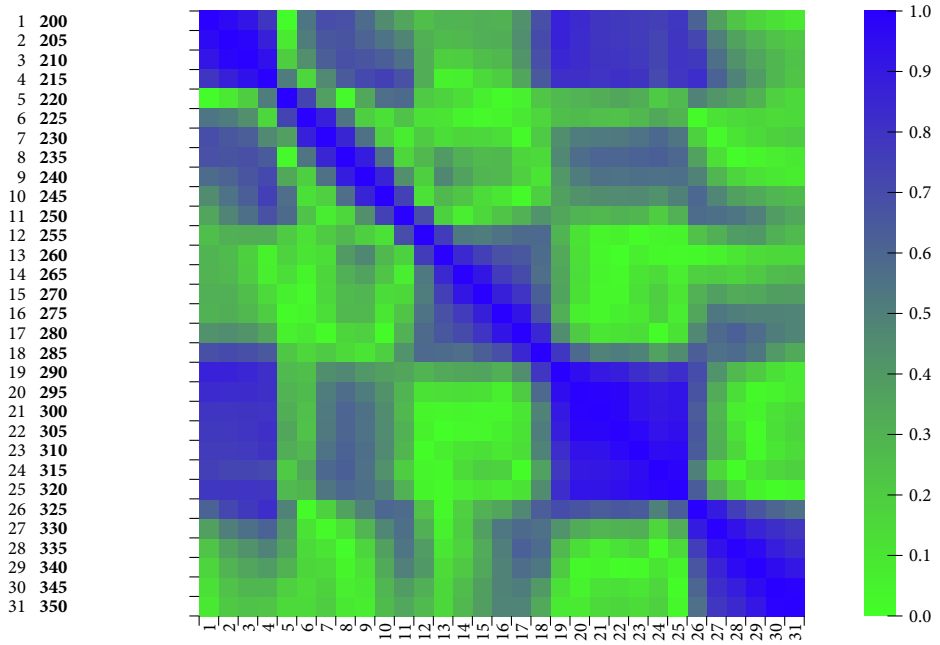


Figure 7: The covariance matrix of the 31 UV descriptors comparing the DLCs at every 5th nm. The covariance data were normalized, i.e. they show the absolute value of each covariance entry. Further information on the descriptors is given in the supporting information of Paper I.

pattern (123678-HxCDF, 1234678-HpCDF, 123789-HxCDF, and 1234789-HpCDF) (see the supporting information of Paper I). In this region, the spectral trend of the seven measured PCDDs was less prominent compared to the PCDFs, but it also can be interpreted as a two-peak pattern with a peak (visible at 230 nm for 2378-TCDD, Figure 6c) that gradually shifts to higher wavelengths with increasing chlorination (OCDD, Figure 6f, close to 250 nm). Furthermore, for 1234-chlorinated DLCs among the PCDD/Fs, the spectra showed that a peak maximum at or just below the lower wavelength limit. This can, for instance, be seen in the spectra of OCDF and OCDD (Figure 6ef).

From the measured UV spectra, we used the absorption at every 5th nm to derive descriptors for each DLC. This resulted in 31 UV descriptors, named by the measured wavelength. Because the peaks of the spectra are usually quite wide, around 10–15 nm, the correlation between descriptors is high. This can, for instance, be seen for the wavelength range 200–215 nm (Figure 7). The UV descriptors carry similar information block-wise, and this is illustrated in the PCA in Paper I where adjacent UV descriptors, such as 230, 235, and 240 in the second PC, have similar loadings. This particular wavelength region corresponds to a local minimum lo-

cated right before the last peak of the PCBs, which is located at roughly 250 nm (Figure 6a, and d). This peak refers to the so-called kappa band [7, 32] and shows high absorption for non-*ortho* PCBs, i.e., the more potent PCBs according to the WHO-TEFs (Table 1).

4.4 Applicability, future adjustments, and remarks on the developed electronic descriptors for DLCs

In addition to the DLCs, the atom-specific descriptors are applicable to other chemical classes of small halogenated aromatic compounds with many congeners. Examples of such classes are polychlorinated naphthalenes (PCNs), PBDEs, and brominated or mixed brominated/chlorinated analogues of the DLCs. These are chemical classes that currently are not included in the WHO-TEFs, but where specific congeners have been shown to activate AhR. For instance, among PCNs, certain penta-, hexa-, and hepta-chlorinated congeners have been shown to be active in *in vitro* tests of AhR mediated effects, while other congeners with the same number of chlorine atoms have been shown to be inactive [83].

The most accurate approach to calculating electronic properties is to use *ab initio* methods, such as coupled cluster methods [18], however, such calculations are very time consuming. In Paper I, we calculated MO energies and electron densities using the quantum mechanical method density functional theory (DFT) and a semi-empirical method. The DFT-based method derives properties of the molecule (e.g. calculates the energy) based on a determination of the electron density of the molecule [65]. Semi-empirical methods are faster than DFT, but have a higher extent of approximations (i.e. simplifications) in their calculations. To correct the deficiencies introduced by the approximations, semi-empirical methods use experimental data from studied molecules' electronic properties and embed them as empirical parameters in the calculations [58].

We chose the semi-empirical method AM1 (Austin Model 1) [36] for the task of optimizing the geometry and calculating the energy of the DLCs because it gives low-energy 3D conformations of the molecules that are consistent with the available crystal structures of PCBs [45, 60, 68, 100]. In addition, this method provides structures with non-planar low-energy 3D conformations of PCDD/Fs, i.e. chlorine–chlorine twists for adjacent carbons, which are suggested to be highly plausible conformations for these compounds [9, 30]. For the DFT calculations, we used the Becke–3–Lee–Yang–Parr (B3LYP) [19, 66, 104, 129] functional together with the basis set 6-31G**; a basis set defines how the MOs are represented mathematically. Zou et al. [141] earlier used the quantum mechanical method Hartree–Fock, together with a similar basis set (6-31G*), to calculate ESP-based

descriptors to differentiate between PCDF congeners. In that case, those descriptors proved to be useful in QSAR modeling of AhR binding affinities and *in vitro* responses to PCDFs. Our studies revealed that the MO energies based on AM1 and B3LYP/6-31G** are highly correlated, and that the MO density maps show similar density distributions. This implies that, for DLCs, it is sufficient to use the semi-empirical approach to calculate MO densities and to generate atom-specific descriptors based on these MOs.

To experimentally derive descriptors—for example, by using spectroscopy—takes more time and entails greater costs compared to using calculated descriptors. However, it would be interesting to measure the absorption spectra for classes of compounds other than PCDD/Fs and PCBs that are known to have congeners that activate AhR. Examples of such classes are the PCNs, PBDEs, and brominated or mixed-halogenated analogues of PCDD/Fs and PCBs. This would provide meaningful insights in how applicable the lower UV range is for predicting the AhR activity generated by halogenated aromatic compounds.

5. QSAR modeling of DLCs based on in vitro data from different species (Papers I–III)

For risk assessment of DLCs, the main source of data has for a long time been whole-animal rat studies or rat cell studies [123]. An unanswered question is how different the results might be if one were to measure *in vitro* responses not based on rat cells. The set of AhR-related *in vitro* data for DLCs presented in Paper II consisted of data obtained by bioassays of rat, mouse, and guinea pig cells performed in the same laboratory and by the same person. This provided us with a unique opportunity to closely study potential species differences. However, to save time and costs, not all DLCs were tested in these experiments. The study left 11 of the 29 DLCs untested (see Table 1, Section 3.3), and hence with no data regarding their potency. To give a better picture of the study, QSARs were developed to fill the data gaps. Moreover, by developing QSARs, we had the opportunity to make interpretations of the models, which might carry information on species-specific differences. These studies found that among the most influential descriptors were several that were in common in all of the QSAR models, indicating that rat, mouse, and guinea pig cells have a similar overall SAR. In the *in vitro* data for the tested PCDD/Fs, the mouse and guinea pig REPs were often higher than the rat REPs (Paper II). The developed QSARs picked up this trend. For example, the predicted REPs of the untested 12378-PeCDF were 0.6, 0.7, and 0.09 for the QSARs of the guinea pig, mouse, and rat *in vitro* responses, respectively (Table 2, Section 5.1). This thesis also includes QSARs based on consensus toxicity factors (CTFs), i.e., condensed information from several *in vitro* responses from human and rat cells (Paper III). The descriptor analysis of these models is included in this chapter (see Section 5.2). The human and rat CTFs are further discussed in Chapter 6.

5.1 Predictions of rat, mouse, and guinea pig *in vitro* REPs for untested DLCs

In Paper II, we developed QSARs using REPs from rat, mouse, and guinea pig *in vitro* responses to study potential species differences. In general, the REPs of the tested PCDDs were similar for the three cell systems. However, for the tested PCDFs, the guinea pig and mouse cell systems generally generated higher REPs than the rat cell study (Paper II). The guinea pig cell study also generated higher REPs for specific non-*ortho* and mono-*ortho* PCBs than the rat cell study. Our QSARs picked up these differences in REPs between the three studied cell systems (Table 2). Below follows a more detailed discussion of the REP predictions from the QSAR models.

Table 2: Summary of predicted REPs from the QSARs based on guinea pig, rat, and mouse *in vitro* responses for the untested DLCs (Paper II). For comparisons, the corresponding WHO-TEFs [123] are also shown. REPs that are at least one order of magnitude higher than the corresponding WHO-TEF are marked in bold text. REPs marked with *ex* were outside the applicability domain, which means that these REPs came from extrapolations.

Compound	guinea pig	rat	mouse	WHO-TEF
1,2,3,4,7,8-HxCDD	0.2	0.1	0.3	0.1
1,2,3,7,8,9-HxCDD	0.2^{ex}	0.2	0.4^{ex}	0.1
OCDD	-	-	-	0.0003
1,2,3,7,8-PeCDF	**0.6**	0.09	**0.7**	0.03
1,2,3,6,7,8-HxCDF	0.2	0.04	0.2	0.1
1,2,3,7,8,9-HxCDF	0.8	0.1	**1**	0.1
OCDF	$\mathbf{0.03}^{ex}$	$\mathbf{0.01}^{ex}$	$\mathbf{0.04}^{ex}$	0.0003
3,4,4',5-tetraCB (PCB 81)	**0.002**	0.0001	0.0005	0.0003
2,3,4,4',5-pentaCB (PCB 114)	**0.0002**	0.00003	0.00004	0.00003
2',3,4,4',5-pentaCB (PCB 123)	0.000009	0.000001	0.000001	0.00003
2,3,3',4,4',5'-hexaCB (PCB 157)	0.00004	0.00001	0.000008	0.00003

Table 2 shows the predicted REPs of untested DLCs based on the three QSARs of the guinea pig, mouse, and rat *in vitro* responses. These REPs are hereafter referred to as the guinea pig, mouse, and rat REPs. Some of the predicted REPs were extrapolated from the QSAR models. Regarding OCDD, we did not report any results because this compound was clearly out of the domain for all three QSAR models; the probability that OCDD belonged to the models was outside the 95% confidence level (Paper II). All predicted guinea pig and mouse REPs were similar or higher than the rat REPs and the WHO-TEFs (Table 2). For PCDDs, the predicted guinea pig and mouse REPs were similar. For example, the guinea pig and mouse REPs of 123478-HxCDD were 0.2 and 0.3, respectively, while the rat REP had a value of 0.1, which is the same value as the corresponding WHO-TEF. For 12378-PeCDF and 123789-HxCDF, the guinea pig and mouse REPs were substantially higher than the corresponding rat REPs and WHO-TEFs (Table 2).

The predicted REPs for OCDF were similar for all three cell systems, but notably they were approximately two orders of magnitude higher than the corresponding WHO-TEF (0.0003). An earlier *in vitro* study [21] of OCDF with a similar rat cell system and method have reported a REP of 0.012, i.e., a value closer to our predicted values than to the corresponding WHO-TEF. However, the WHO-TEFs were primarily based on *in vivo* data, and there can be differences of several orders of magnitude between *in vitro* and *in vivo* REPs [51, 123]. The big difference between the WHO-TEF and our data might be explained by large differences in toxicokinetics between 2378-TCDD and OCDF [121]. That is, for OCDF there

can be large differences (compared to 2378-TCDD) in how much of and at what rate the administered dose reaches the target where the toxic effect is measured.

Unlike for the PCDD/Fs, the REPs of the tested PCBs were generally below the corresponding WHO-TEFs for all three cell systems. For instance, the REPs of PCBs 169 and 118 were between half an order and one order of magnitude lower than the corresponding WHO-TEFs (Paper II). However, while the rat and mouse cell systems rendered REPs that consistently were on the low side, the guinea pig REPs were for several compounds higher than the WHO-TEFs, e.g., for PCBs 77, 105, and 156 (Paper II). This could explain the varied outcome in the predictions of the untested non-*ortho* and mono-*ortho* PCBs for the guinea pig cell system. More specifically, in contrast to the other predictions, the predicted REPs of PCBs 81 and 114 were one order of magnitude higher for the guinea pig cell system than the corresponding WHO-TEFs (Table 2).

5.2 Descriptor analysis of the developed QSARs based on rat, mouse, guinea pig, and human cells

In this thesis, QSARs of DLCs were developed by using data from *in vitro* studies of 1) a rat cell system from Behnisch et al. [21] (Paper I), 2) individual rat, mouse, and guinea pig cell systems (Paper II), and 3) several rat and human *in vitro* responses (Paper III). The QSAR studies in Paper I showed that combining blocks of 2D descriptors, 3D descriptors, and spectroscopic data for the model building is superior to only using one or two of the three types of variables. Therefore, all three blocks were used as a starting point for all subsequent QSAR modeling in Papers II and III. All QSAR models developed by rodent-based bioassays (rat, mouse, and guinea pig) showed similar features regarding what types of descriptors that contributed the most to the models (Papers I–III). In all cases, a combination of structural 2D descriptors, calculated electronic descriptors, and experimentally derived electronic descriptors from UV spectra were among the most influential variables. When developing a QSAR for the *in vitro* data of human responses, it was, however, a disadvantage to include many of the 2D descriptors that were used for the rodent models (Paper III). In contrast to the rodent-based QSARs, the QSAR for the human *in vitro* data was developed based on only a small set of descriptors, where several of them were atom-specific electronic descriptors.

The most significant descriptors for the QSARs of rodent data (Papers I–III) indicated that highly potent DLCs have large positive van der Waals surface areas and high UV absorption at both 210–215 nm and 235–240 nm. These wavelength regions indicate the importance of the resonance of the benzene rings and of properties due to the specific chlorine substitution patterns on the rings (see Sec-

tion 4.3). The chemical stability of the DLCs is also important for these models. Highly potent DLCs have lower GAP values, which indicates that these DLCs are chemically less stable than the less potent DLCs. Moreover, the shape of the DLCs, described primarily by the Kier3 and Balaban J indexes [113], is also important for the potency of DLCs.

The significant descriptors for the QSAR of the human responses included atom-specific descriptors that reflected the electronic properties of the lateral positions of the structures. DLCs that were highly potent often had a high ratio between the second highest and the highest-ranked lateral carbon HOMO density (Ha2/Ha1). These DLCs also had a high ratio between the second highest and the highest-ranked lateral carbon and the corresponding for the LUMO density (La2/La1) (see Section 4.1). Moreover, for the DLCs, our results showed that the first and second highest HOMO and LUMO densities for a lateral carbon were always on different rings. This means that highly potent DLCs have similar highest HOMO and LUMO densities on both rings regarding their lateral carbons. Furthermore, the MO densities of the lateral carbons also reflected the ranking of the MO densities of the connected chlorine atoms (Figure 3). That is, the ranking of the MO densities of the lateral chlorine atoms corresponded to the ranking of the MO densities of the lateral carbons. This suggests that highly potent DLCs have two equally reactive sites regarding lateral chlorine atom positions and that these sites are located on different rings.

In the QSAR model of the human responses, the partial charges of the lateral carbon positions were also important in order to differentiate between more potent and less potent DLCs. Especially QC_a3_ESP, i.e. the partial charge at the third highest ranked lateral carbon (see Section 4.1), was a significant descriptor. In general, our calculations indicated that highly potent DLCs have higher partial charges at their first, second, and third highest ranked lateral carbon compared to the less potent DLCs. Assuming that positions with high partial charges can be thought of as reactive, as suggested by Olivero-Verbel et al. [83], this could be a measure of the ability of the DLCs to interact with AhR.

5.3 Applicability, future adjustments, and remarks on the developed QSARs of DLCs based on *in vitro* data

The QSARs in this thesis were developed according to the OECD principles (see Section 3.1), and this makes the models suitable to be used for regulatory purposes. More specifically, these QSARs predict REPs that could be used for future reevaluation of the WHO-TEFs. Our *in vitro* studies and developed QSARs showed that for several DLCs the REP differed depending on whether it was based on rat,

mouse, or guinea pig responses (Paper II). If one were to apply our REPs to the REP distributions used at the last expert meeting for establishing the WHO-TEFs [123], the guinea pig and mouse data would promote higher TEFs for several DLCs for which there are little or no *in vivo* data. Our data would, in particular, have a great impact on the TEF reevaluation for DLCs that lack or have very little available *in vivo* and *in vitro* data. An example of such a DLC is 123789-HxCDF, for which there were no *in vivo* data and only two *in vitro* studies for the reevaluation [123]. For 12378-PeCDF and OCDF, some of the predicted REPs were based on extrapolations (Table 2). In such cases, we suggest that the data should not be used for regulatory purposes even if the results seem reasonable.

In this thesis, the QSARs were built on *in vitro* data based on rat, mouse, guinea pig, and human responses to the three chemical classes of DLCs—PCDDs, PCDFs, and PCBs. In contrast, all models that I have encountered in the literature that combine congeners from at least two of these chemical classes are built on rat-based information. More specifically, these models have used the rat-based [^{3}H]2378-TCDD competitive binding data or AhR *in vitro* rat hepatocyte assay data by Safe and followers [16, 17, 74, 92, 93, 95]. Another difference compared to previously reported mixed chemical-class QSAR studies regarding PCDD/Fs and PCBs is that we primarily focused on DLCs. That is, we used a considerably smaller number of compounds compared to earlier reported mixed-class models [10, 20, 37, 71, 84, 99, 102, 115, 116].

To validate the models (Paper I–III), approximately one fourth of the data was removed and used as an external validation set. For the split into training and validation sets, we used the X-data to define the chemical space of the studied compounds and applied PCA. By using PCA, a training set was picked so that it covered and spanned the full space. To avoid extrapolation, the training set for the model was set to span the full Y-range of the tested biological response. For example, in the QSAR of the human CTFs (see Chapter 6), 1234678-HpCDF had the lowest CTF of the tested DLCs and was thus explicitly put as part of the training set (Figure 8). The resulting QSAR (Figure 8, left) only included PCDD/Fs due to the low responsiveness to PCBs in the human cells (Paper III). The results from the modeling of the human responses, that is, the QSAR based on human CTFs (Figure 8, left), are discussed in Section 6.2. In the QSARs based on rodent *in vitro* data (Papers I–III), PCB 126 was consistently underestimated. For example, in the QSAR based on rat CTFs, the underestimation was more than two orders of magnitude (Figure 8, right). When this compound instead was part of the training set (Paper I), the underestimation was less prominent. However, overall this showed that the key chemical properties of PCB 126 related to its biological response were not covered.

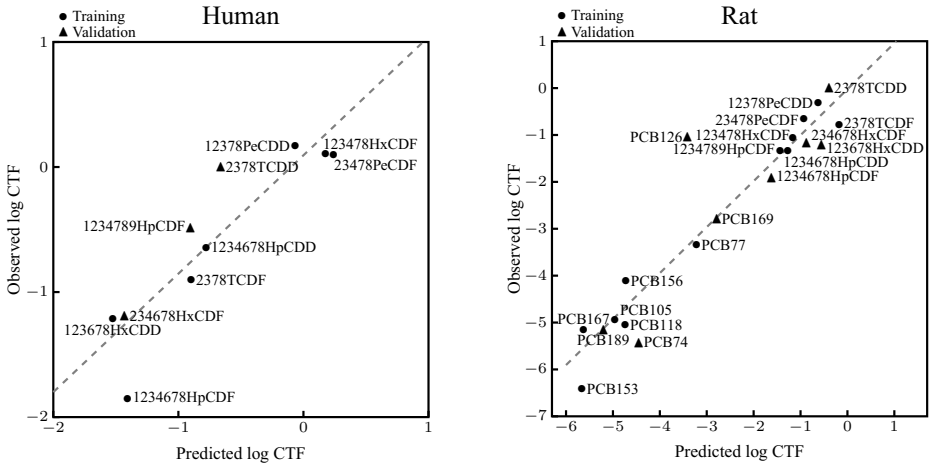


Figure 8: The observed versus predicted log-transformed consensus toxicity factors (CTF) of the developed human and rat QSAR models (Paper III).

When developing the descriptors for the DLCs (Paper I), our aim was to capture the ability for the molecule to make non-covalent interactions to other small molecules and proteins. Specifically, the descriptors were calculated and selected to capture subtle electronic features of the lateral atoms (see Section 4.1) or to describe the impact on the π-system due to different substitution patterns (see Section 4.3). Previously reported QSARs of mixed classes concerning PCDD/Fs and PCBs [10, 20, 37, 84, 99, 102, 115, 116] have in common that the models are, partially or fully, built on electronic properties. Some of the earlier mixed models [10, 84, 102] were 3D QSARs; to develop a 3D QSAR, the studied molecules are aligned and spatial fingerprints are calculated to describe, for instance, steric effects, lipophilicity, and electrostatic interactions [91]. A model based on 78 congeners, including PCDD/Fs and PCBs, was developed by Turner et al. [115] based on infrared vibration frequencies. However, to my knowledge, no earlier reported mixed-class model includes UV-data or the interpretation of specific regions of the obtained spectra.

6. Species-specific consensus toxicity factors of DLCs (Paper III)

At the reevaluation meeting for establishing the WHO-TEFs [123], there were questions regarding how comparable human and rodent cell systems really are for measuring AhR-mediated effects (see Section 3.3). To obtain insights into this matter, in Paper III we analyzed 18 DLCs for a total of 17 different *in vitro* responses. The initial analysis showed that the responsiveness and sensitivity of the DLCs were lower in the human systems than in the rodent systems. These results agree with previous *in vitro* studies of PCB 126 and 2378-TCDD in human cells [125, 140]. However, our studies also showed that there is a difference between human and rodent responses regarding how the DLCs relate to each other in terms of potency. Based on the REP data, we provided a new approach, the CTF, to condense information from a battery of screening tests. A CTF for a DLC is based on multivariate analysis of many DLCs' REPs from the applied assays (Figure 9). More specifically, we performed PCA using the *in vitro* data of one target species and used the first PC, the score values, to establish the toxicological ranking in the data. This approach directly addressed the ranking of the DLCs seen for each applied response in the PCA. By using PCA, the information is condensed and noise is removed (see Section 3.5.1). In this case, the noise consisted of data that did

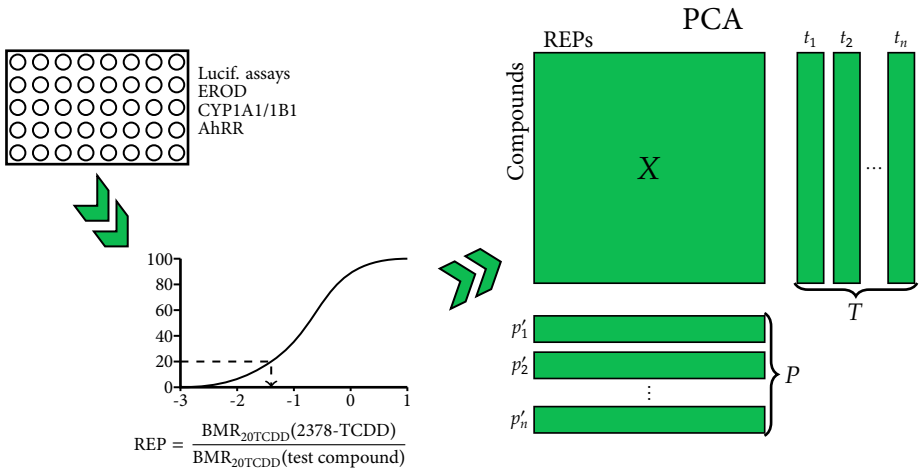


Figure 9: The REPs for the *in vitro* data are based on the concentrations needed to reach a benchmark response (BMR) equal to 20 % of the maximum 2378-TCDD response. The REP data were log-transformed, scaled, and then used for the PCA. The data included studies of luciferase reporter gene and ethoxyresorufin-O-deethylase (EROD) assays, as well as gene expressions of CYP1A1, CYP1B1, and aryl hydrocarbon receptor regulator (AhRR) (see Section 6.1).

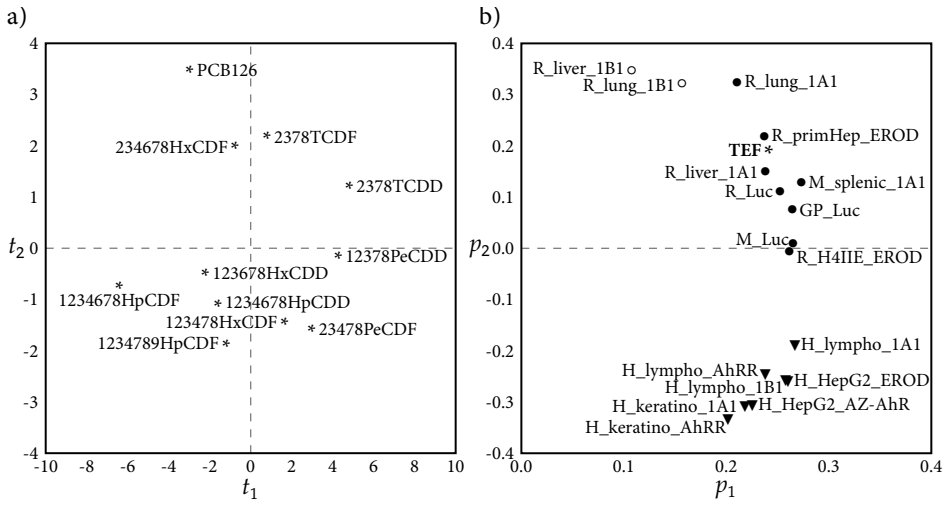


Figure 10: PCA results based on the REPs of 17 *in vitro* responses from human and rodent bioassays and the WHO-TEFs (here called "TEF"). (a) The score plot for the first and second PCs explaining 64 % and 17 % of the variation in the data, respectively, and (b) the corresponding loading plot. In (b), the human responses are shown as triangles, the rodent responses are shown as dots, and the rodent CYP1B1 responses are shown as open circles.

not contribute to the overall toxicological ranking. Moreover, by performing a PCA, we were able to see that a large part of the information was systematic variation for human and rodent *in vitro* responses (Figure 10). This chapter includes a case study where CTFs and WHO-TEFs were applied to representative concentrations of DLCs in fish from the Baltic Sea region (see Section 6.2). The case study showed that the TEQ based on CTFs in the fish samples did not substantially change, even though the PCBs were assigned a lower impact by the human CTFs compared to the WHO-TEFs (Table 3).

6.1 Applied AhR mediated biological responses

In 2009, the project "The development, validation and implementation of human systemic Toxic Equivalencies (TEQs) as biomarkers for dioxin-like compounds" (SYSTEQ) was initiated to extensively study the toxicological effects of DLCs [154]. Within this project, a large group of DLCs—18 of the 29 compounds defined as DLCs in this thesis (see Section 3.2)—plus two additional PCBs were studied in parallel for different *in vitro* AhR mediated endpoints. In Paper III, we performed a comprehensive study of seven rat, two mouse, one guinea pig, and seven human *in vitro* responses measured in SYSTEQ. These 17 responses comprised luciferase gene expression assays, CYP1A1/CYP1B1/aryl hydrocarbon receptor regulator

mRNA expression assays, and ethoxyresorufin-O-deethylase induction, i.e., CYP induction, assays. Furthermore, these assays targeted several types of organs and tissues, such as liver, lung, spleen, and skin. The slope and efficacy, i.e., the top level, of the dose-response curves of the tested DLCs and 2378-TCDD were not always comparable. For this reason, we calculated the concentration needed for a tested DLC to reach a benchmark response equal to 20% of the maximum 2378-TCDD response in order to establish *in vitro REPs* (Paper III). The advantage of a benchmark response at the lower part of the curve is that the agreement in curve shape is less essential, and we believed that calculating REPs based on 50% effect concentrations might add a significant uncertainty due to differences in slope and efficacy between the tested DLCs and 2378-TCDD.

6.2 The rat and human CTFs

The rat CTFs corresponded well with the WHO-TEFs. This was rather expected because the WHO-TEFs are primarily based on data from rat studies (see Section 3.3). Among the PCDD/Fs, most rat CTFs were within one order of magnitude from the corresponding WHO-TEFs. For example, 123789-HxCDD has a CTF of 0.3 and a WHO-TEF of 0.1 (Table 3). The exceptions were OCDF and PCB 169. The rat CTF for OCDF is 23 times higher than the corresponding WHO-TEF. As discussed in Section 5.1, the higher CTF of OCDF compared to the WHO-TEF is likely due to the fact that CTFs are based on *in vitro* information, which does not consider absorption, distribution, metabolism and excretion processes; OCDF might have considerably lower bio-availability *in vivo* compared to 2378-TCDD. Toxicokinetics are accounted for in *in vivo* experiments and by the expert judgment underlying the WHO-TEF of OCDF. Moreover, the *in vivo* and *in vitro* databases [51] used when deciding on the WHO-TEFs [123] show that there are many *in vitro* REPs that are higher than the *in vivo* REPs regarding OCDF. PCB 169 has a rat CTF more than ten times lower than the corresponding WHO-TEF; the CTF is 0.002 and the WHO-TEF is 0.03. The CTF of PCB 169 (0.002) is similar to reported *in vitro* REPs in the literature (0.0019 and 0.0034) [21] obtained by using luciferase reporter gene and ethoxyresorufin-O-deethylase induction assays, respectively. On the whole, there are large differences in reported *in vivo* REPs for PCB 169, from less than 0.001 to almost 1 [51]. These difference are believed to be due to species differences and differences in the studied endpoints [123].

The human CTFs indicate that there are differences in potency within the groups of hexa-chlorinated and hepta-chlorinated DLCs among PCDD/Fs (Table 3). In contrast, in the WHO-TEFs, all of the hexa-chlorinated and hepta-chlorinated DDs and DFs are assigned a value of 0.1 and 0.01, respectively (Table 3). In particular, 123789-HxCDD has a very low CTF (0.002) and 123478-HxCDF has a very high

Table 3: Summary of the CTFs calculated from the *in vitro* data (marked with a "t") and the CTFs for the untested DLCs (unmarked) predicted by QSAR modeling based on existing CTFs (Paper III). For comparisons, the corresponding WHO-TEFs [123] are also shown. CTFs that are at least one order of magnitude higher than the corresponding WHO-TEF are marked in bold text. CTFs marked with *ex* were outside the applicability domain, which means that these CTFs came from extrapolations.

	Compound	rat CTF	human CTF	WHO-TEF
Chlorinated dibenzo-*p*-dioxins	2,3,7,8-TCDD^{t}	1	1	1
	1,2,3,7,8-PeCDD^{t}	0.5	1	1
	1,2,3,4,7,8-HxCDD	0.2	0.03	0.1
	1,2,3,6,7,8-HxCDD^{t}	0.06	0.06	0.1
	1,2,3,7,8,9-HxCDD	0.3	**0.002**	0.1
	1,2,3,4,6,7,8-HpCDD^{t}	0.04	**0.2**	0.01
	OCDD	-	**0.005**	0.0003
Chlorinated dibenzofurans	2,3,7,8-TCDF^{t}	0.2	0.1	0.1
	1,2,3,7,8-PeCDF	0.2	**0.6**^{ex}	0.03
	2,3,4,7,8-PeCDF^{t}	0.2	1	0.3
	1,2,3,4,7,8-HxCDF^{t}	0.09	**1**	0.1
	1,2,3,6,7,8-HxCDF	0.07	0.04^{ex}	0.1
	1,2,3,7,8,9-HxCDF	0.3	0.02	0.1
	2,3,4,6,7,8-HxCDF^{t}	0.07	0.06	0.1
	1,2,3,4,6,7,8-HpCDF^{t}	0.01	0.01	0.01
	1,2,3,4,7,8,9-HpCDF^{t}	0.05	**0.3**	0.01
	OCDF	**0.007**^{ex}	**0.2**^{ex}	0.0003
Non-*ortho*–substituted PCBs	3,3',4,4'-tetraCB (PCB 77)^{t}	0.0004	-	0.0001
	3,4,4',5-tetraCB (PCB 81)	0.0002	-	0.0003
	3,3',4,4',5-pentaCB (PCB 126)^{t}	0.09	**0.003**	0.1
	3,3',4,4',5,5'-hexaCB (PCB 169)^{t}	**0.002**	-	0.03
Mono-*ortho*–substituted PCBs	2,3,3',4,4'-pentaCB (PCB 105)^{t}	0.00001	-	0.00003
	2,3,4,4',5-pentaCB (PCB 114)	0.00006	-	0.00003
	2,3',4,4',5-pentaCB (PCB 118)^{t}	0.000009	-	0.00003
	2',3,4,4',5-pentaCB (PCB 123)	0.000009	-	0.00003
	2,3,3',4,4',5-hexaCB (PCB 156)^{t}	0.00008	-	0.00003
	2,3,3',4,4',5'-hexaCB (PCB 157)	0.00003	-	0.00003
	2,3',4,4',5,5'-hexaCB (PCB 167)^{t}	0.000007	-	0.00003
	2,3,3',4,4',5,5'-heptaCB (PCB 189)^{t}	0.000007	-	0.00003

CTF (1) compared to the corresponding WHO-TEFs. For 123478-HxCDF, there are reported *in vitro* REPs below 0.1 as well as above 1. But the *in vivo* REPs have been closer to 0.1 according to the *in vivo* and *in vitro* databases that are based on rat and mouse data [51]. Regarding 123789-HxCDD, there has been, according to Haws et al. [51], only one reported *in vivo* REP (0.029), which was determined in a the study based on mouse fetuses [109]. However, those data were not considered sufficient to decrease the TEF of 123789-HxCDF [123].

At the last reevaluation [123], it was discussed whether or not to increase the TEF of 1234678-HpCDD to 0.03 due to existing *in vivo* REPs. The human CTF of 1234678-HpCDD (0.2) is considerably higher than those REPs and is higher than the *in vitro* REPs reported by Haws et al.[51]. Hence, our results support an increase in the TEF for 1234678-HpCDD. The human CTFs of the hepta-chlorinated DFs (0.01–0.3) are in the same range as the *in vitro* REPs (0.02–0.3) [26, 119] based on rat and mouse cells that were used at the last reevaluation meeting [123]. At that time, the expert panel at the reevaluation meeting concluded that there was too much uncertainty in this limited database to increase the TEF of the hepta-chlorinated DFs. Another reason to keep the TEF at 0.01 was the possibly low absorption of these DLCs in the gastrointestinal tract [123]. The human CTF of OCDD is more than ten times higher than the corresponding WHO-TEF. This is interesting because OCDD has been found at high levels in human blood [1, 50]. The human CTF of OCDF is approximately 600 times higher than the corresponding WHO-TEF and 29 times higher than the rat CTF. However, the reason for the higher CTF of OCDD might, just as for the rat CTF of OCDF, be differences in toxicokinetics, e.g. in bio-availability, between highly chlorinated DLCs and 2378-TCDD.

Among the less chlorinated DLCs, we found that 12378-PeCDF has a more than ten times higher human CTF than the corresponding WHO-TEF. Such an increase in risk factor might have a significant impact on the calculated TEQs in fish, as described in more detail below. Regarding PCBs, most of the tested DLCs were inactive in the applied human assays. Only PCB 126, the most potent PCB according to the WHO-TEFs, was active in a majority of the assays, and achieved a CTF of 0.003, which is 30 times lower than the corresponding WHO-TEF. The human CTF is consistent with earlier reports [125, 140] that also showed human *in vitro* REPs to be lower than the WHO-TEF of 0.1. Recently, based on the currently available *in vitro* data using human cells, it has been suggested that the AhR-mediated risk in humans for PCBs is overestimated in the current TEF concept [126].

6.3 Comparisons of estimated risks associated with DLCs in fish from the Baltic Sea calculated by using human CTFs and WHO-TEFs

In the EU, government agencies use the WHO-TEFs to estimate the risk associated with DLCs in food products by calculating the corresponding TEQ (see Section 3.3). From the TEQ, they conclude whether the risk of DLCs is below the EU limit set for that specific food product. Regarding fish caught in the Baltic Sea, there has long been a problem with TEQs being too high compared to the EU limit [12, 75, 80]. The following case study gives an indication of what impact the different DLCs would have if CTFs were used as weights when calculating TEQs in fish and what would happen to the size of the TEQs, that is, what would happen to the estimated risk of DLCs in fish.

The Swedish National Food Agency (SNFA) continuously measures food products regarding contaminants such as DLCs. In 2015, the SNFA reported that five of the six studied salmonid samples taken from different parts of the Baltic Sea exceeded the EU limit for fish, which is 6.5 pg TEQ/g fresh weight [151]. If one applies the human CTFs in the TEQ calculations instead of the WHO-TEFs, the TEQs would all exceed the EU limit for fish. The estimated risk associated with DLCs of these samples, based on human CTFs, is in the range of 7.1–11.0 pg TEQ/g fresh weight. The decreased impact of PCBs in the calculations is balanced by the increased impact of, in particular, 23478-PeCDF (Table 4). Some compounds that have a high impact in the TEQ calculation, e.g. 12378-PeCDF, have CTFs that are derived from extrapolations, and these values must be treated with caution. Calculations using only the CTFs from the *in vitro* data and from non-extrapolated CTFs, and using the WHO-TEFs for the remaining DLCs, also show that the EU limit would be exceeded for all six samples. The estimated risk of DLCs in the samples would, in that case, be in the range of 6.8–10.4 pg TEQ/g fresh weight.

The greatest contributions to the CTF-based TEQ were from 23478-PeCDF, followed by 12378-PeCDD, 2378-TCDF, 12378-PeCDF, and 2378-TCDD (Table 4). For the WHO-TEFs, the ranking would be 23478-PeCDF, followed by 12378-PeCDD, 2378-TCDF, and 2378-TCDD. That is, it is the same ranking as before except for 12378-PeCDF, which only contributed with 0.5–1% of the WHO-TEQs compared to 5–7% of the CTF-based TEQs. Notably, PCB 126 only contributed with 1–2% of the CTF-based TEQs due to the much lower value of the human CTF compared to the WHO-TEF. In contrast, PCB 126 contributed with approximately 50 % of the reported WHO-TEQs. The hexa-chlorinated DDs and DFs were at very low concentrations, 0–0.36 pg/g fresh weight, in these samples, and thus gave very small contributions to the calculated TEQs. Among the hexa-chlorinated congeners, 123478-HxCDF contributed the most (1–2%) to the CTF-based TEQs due to its high human CTF (Table 3). The hepta- and octa-chlorinated DDs and

Table 4: The contribution of each DLC to the estimated risk of DLCs in fish using the WHO-TEFs and the human CTFs (pg TEQ/g fresh weight). More information regarding the fish studies is given in the SNFA report[151]. This table is continued on page 38.

	Dalälven		Kluntarna		Seskar-Furö	
	WHO-TEF	CTF	WHO-TEF	CTF	WHO-TEF	CTF
2,3,7,8-TCDD	0.26	0.26	0.49	0.49	0.49	0.49
1,2,3,7,8-PeCDD	0.55	0.55	0.91	0.91	0.89	0.89
1,2,3,4,7,8-HxCDD	0.0020	0.00060	0.0050	0.0015	0.0050	0.0015
1,2,3,6,7,8-HxCDD	0.024	0.014	0.036	0.022	0.036	0.022
1,2,3,7,8,9-HxCDD	0	0	0.0020	0.000040	0.0020	0.000040
1,2,3,4,6,7,8-HpCDD	0	0	0	0	0	0
OCDD	0	0	0	0	0	0
2,3,7,8-TCDF	0.44	0.44	0.80	0.80	0.81	0.81
1,2,3,7,8-PeCDF	0.017	0.34	0.033	0.66	0.033	0.66
2,3,4,7,8-PeCDF	1.6	5.4	2.3	7.8	2.3	7.5
1,2,3,4,7,8-HxCDF	0.0060	0.060	0.015	0.15	0.015	0.15
1,2,3,6,7,8-HxCDF	0.011	0.0044	0.022	0.0088	0.025	0.010
1,2,3,7,8,9-HxCDF	0	0	0	0	0	0
2,3,4,6,7,8-HxCDF	0.009	0.0054	0.017	0.010	0.018	0.011
1,2,3,4,6,7,8-HpCDF	0	0	0	0	0	0
1,2,3,4,7,8,9-HpCDF	0	0	0	0	0	0
OCDF	0	0	0	0	0	0
PCB 77	0.0034	0	0.011	0	0.012	0
PCB 81	0.00013	0	0.00033	0	0.00027	0
PCB 105	0.042	0	0.073	0	0.074	0
PCB 114	0.0026	0	0.0048	0	0.0049	0
PCB 118	0.14	0	0.25	0	0.24	0
PCB 123	0.00091	0	0.0013	0	0.0013	0
PCB 126	2.3	0.069	5.4	0.16	5.5	0.16
PCB 156	0.023	0	0.035	0	0.032	0
PCB 157	0.0051	0	0.0079	0	0.0077	0
PCB 167	0.011	0	0.017	0	0.017	0
PCB 169	0.34	0	0.49	0	0.46	0
PCB 189	0.0033	0	0.0043	0	0.0038	0
Total TEQ	5.8	7.1	11.0	11.0	10.9	10.7

Table 4 (continued): The contribution of each DLC to the estimated risk of DLCs in fish using the WHO-TEFs and the human CTFs (pg TEQ/g fresh weight).

	Nordingrå		Pukavik		Sollefteå	
	WHO-TEF	CTF	WHO-TEF	CTF	WHO-TEF	CTF
2,3,7,8-TCDD	0.45	0.45	0.51	0.51	0.36	0.36
1,2,3,7,8-PeCDD	0.70	0.70	0.82	0.82	0.71	0.71
1,2,3,4,7,8-HxCDD	0.0050	0.0015	0.0050	0.0015	0.0040	0.0012
1,2,3,6,7,8-HxCDD	0.027	0.016	0.033	0.020	0.032	0.019
1,2,3,7,8,9-HxCDD	0.0020	0.000040	0.0020	0.000040	0.0020	0.000040
1,2,3,4,6,7,8-HpCDD	0	0	0	0	0	0
OCDD	0	0	0	0	0	0
2,3,7,8-TCDF	0.75	0.75	0.85	0.85	0.65	0.65
1,2,3,7,8-PeCDF	0.027	0.53	0.033	0.66	0.027	0.55
2,3,4,7,8-PeCDF	1.4	4.8	1.8	6.0	1.7	5.7
1,2,3,4,7,8-HxCDF	0.015	0.15	0.017	0.17	0.016	0.16
1,2,3,6,7,8-HxCDF	0.020	0.0080	0.021	0.0084	0.022	0.0088
1,2,3,7,8,9-HxCDF	0	0	0	0	0	0
2,3,4,6,7,8-HxCDF	0.015	0.0090	0.020	0.012	0.017	0.010
1,2,3,4,6,7,8-HpCDF	0	0	0	0	0	0
1,2,3,4,7,8,9-HpCDF	0	0	0	0	0	0
OCDF	0	0	0	0	0	0
PCB 77	0.014	0	0.014	0	0.0085	0
PCB 81	0.00036	0	0.00030	0	0.00054	0
PCB 105	0.056	0	0.067	0	0.050	0
PCB 114	0.0039	0	0.0041	0	0.0031	0
PCB 118	0.19	0	0.22	0	0.14	0
PCB 123	0.0012	0	0.0014	0	0.0013	0
PCB 126	5.4	0.16	5.8	0.17	4.7	0.14
PCB 156	0.024	0	0.028	0	0.027	0
PCB 157	0.0058	0	0.0068	0	0.0061	0
PCB 167	0.014	0	0.017	0	0.014	0
PCB 169	0.35	0	0.40	0	0.40	0
PCB 189	0.0026	0	0.0032	0	0.0034	0
Total TEQ	9.5	7.6	10.7	9.2	8.9	8.3

DFs did not contribute to the TEQs because all measurements were below the detection limit for these samples.

6.4 Applicability, future adjustments, and remarks of the CTF approach

The CTF approach is applicable to experimental data that have already been translated into the same units and range for comparability, such as REPs (see Section 3.3). In Paper III, we studied AhR activation by *in vitro* responses targeting different steps in the activation chain. We measured transcriptional regulation, mRNA gene expression, and induction of CYP enzymes. That is, we studied *in vitro* responses reflecting the same mechanism of action, which we believe is a prerequisite for the CTF approach. In the future, it would be interesting to apply the approach to related chemical classes—such as PCNs, PBDEs, and brominated or mixed-halogenated analogues of PCDD/Fs and PCBs—with regard to AhR mediated effects. Recently, polybrominated dibenzo-*p*-dioxins (PBDDs) and polybrominated dibenzofurans (PBDFs) have been recommended to be included in the TEF concept [124], and polychlorinated naphthalenes (PCNs) were previously considered for inclusion [41]. The approach could also be applied to other available *in vitro* and *in vivo* databases covering other mechanisms of action, e.g., those mediated through the estrogen receptor.

In Paper III, the input data were log-transformed and scaled to unit variance in the development of the CTFs. In retrospect, this step was not necessary because REPs were used for the PCA and thus the data were normalized (Figure 9). By not scaling the data, the score values remain in the same scale as the initial REPs and hence also the WHO-TEFs. This should simplify the procedure for creating CTFs in the future. The strength of the CTF approach is that it focuses on finding the joint toxicological ranking in the large data set and assigns relative values for the included compounds based on the located ranking. A weaknesses or issue might be what to do if several significant PCs emerge in the PCA. In our case, we could assign the second significant PC to species differences (Figure 10), but the situation might not always be as straightforward.

PCA has previously been used to make classifications or rankings of compounds based on toxicological effects or specific properties of concern, such as persistent, bio-accumulative, and toxic properties [8, 86]. In those studies, the score values from PCAs were used as inputs for linear regression modeling aimed to predict non-ranked compounds. In particular, Andersson et al. [8] created a toxicological ranking of PCBs based on REPs from two CYP1A induction assays in monkey, pig, and chicken hepatocytes. In that case, the species differences were not pronounced and, consequently, all of the data were used to create one PCA and one ranking value for each PCB. In our case, the PCA of the 17 *in vitro* responses showed signif-

icant differences between the human and rodent data (Figure 10), and this led us to split the data and to establish human and rat CTFs separately. All human *in vitro* responses were used, while a subset of the rat responses was selected for the CTFs. The selection was based on how the responses influenced the two PCs in the initial PCA of the 17 *in vitro* responses (Figure 10). The determining factor for including all human responses was that they were closely clustered (Figure 10b).

7. Potential AhR ligands among industrial chemicals by ligand- and structure-based virtual screening (Paper IV)

In addition to DLCs (see Section 3.4), many other structurally diverse compounds have been suggested to activate the AhR pathway [34, 79]. Figure 11 illustrates some structural differences among AhR modulators. Besides DLCs, more flexible and hydrophilic compounds, such as bilirubin and several prostaglandins, have been shown to activate AhR [98, 101]. Moreover, several pesticides have shown *in vitro* AhR-mediated effects [110], and more recently high-throughput screening approaches have been used to identify compounds that activate AhR [150]. In Paper IV, we searched for potential AhR ligands among a large set of industrial chemicals. More specifically, we developed a virtual screening protocol and applied it to a set of 6,445 industrial chemicals. This protocol is both ligand and structure-based. The two ligand-based steps consist of nearest neighbor analysis, which is based on structural and chemical properties in a PCA, and structural fingerprints. The structure-based step is a molecular docking that uses a rat homology model and known AhR ligands. These three steps are run in parallel because multiple scoring and data fusion have proven to be more robust than, and often outperform, a single virtual screening method for various data sets and receptor targets [14, 106, 107, 133]. Applying the results from at least two of the three methods yielded 41 potential AhR ligands (see Section 7.2). Our results were compared to existing data produced by a large high-throughput screening [150] (see Section 7.3).

7.1 Selection of AhR modulators as AhR ligands for the virtual screening protocol

Compounds having shown AhR-mediated effects in rat or mouse *in vitro* studies were collected from the literature and used to analyze the chemical space of AhR modulators (see Section 3.4). The PCA based on 68 2D descriptors and 18 3D descriptors of the 214 collected compounds showed that the chemical space was heterogeneous (Figure 12). A majority of these compounds were environmental pollutants, and moreover, they were (or resembled) halogenated or non-halogenated aromatic compounds, like the DLCs or polycyclic aromatic hydrocarbons (Figure 11; **1–3, 10**). However, approximately 40 compounds differed structurally from the rest of the 214 AhR modulators. This smaller selection included endogenous compounds of different sizes, pharmaceuticals, and natural products (Paper IV). The first two PCs of the PCA reflected a great variation in structural and chemical properties among AhR modulators. For example, the halogenated aromatics

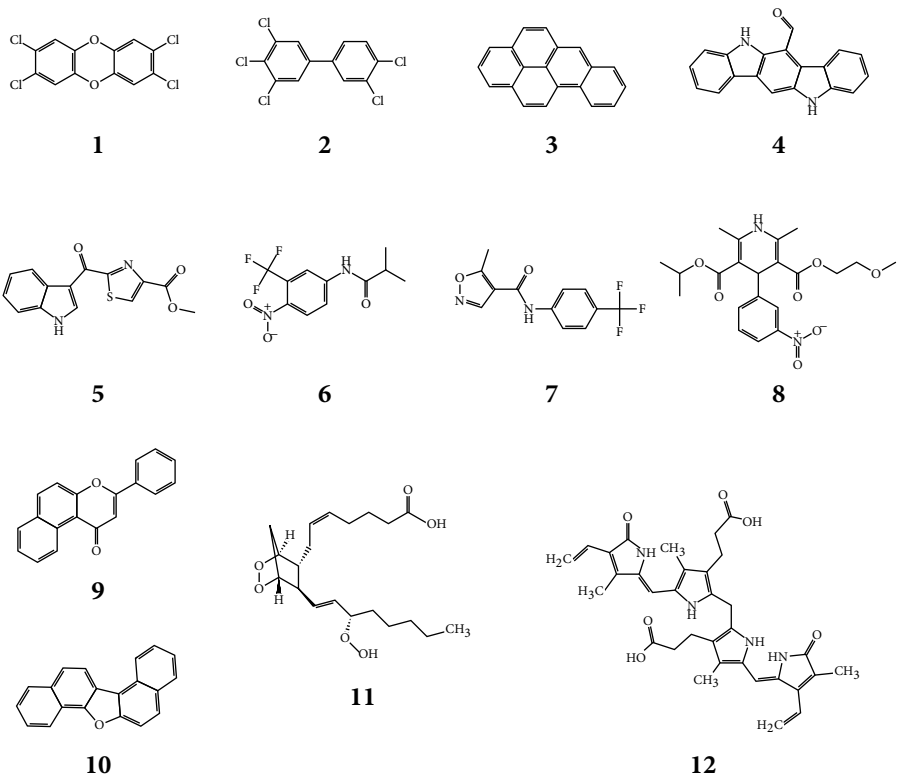


Figure 11: The chemical structures of selected AhR modulators. Nine of these compounds (**1–9**) have also been reported as binders to AhR. **1)** 2,3,7,8-tetrachloro-dibenzo-*p*-dioxin (2378-TCDD), **2)** 3,3',4,4',5-pentachloro-biphenyl (PCB 126), **3)** benzo-a-pyrene, **4)** 6-formylindolo[3,2-b]carbazole, **5)** 2-(1'H-indole-3'-carbonyl)-thiazole-4-carboxylic acid ester, **6)** flutamide, **7)** leflunomide, **8)** nimodipine, **9)** beta-naphthoflavone, **10)** dinaphtho[1,2-b;1'2'-d]furan, **11)** prostaglandin G2, and **12)** bilirubin (Paper IV).

(the large cluster to the left in Figure 12) are less flexible, more hydrophobic, and have a larger hydrophobic surface area in relation to their size compared to the endogenous compounds and pharmaceuticals (Figure 11; **5–8**). However, among the endogenous compounds, some are as rigid as the DLCs (Figure 11; **4**), while other compounds are not aromatic (Figure 11; **11**), or are much larger compared to the other AhR modulators (Figure 11; **12**). Due to the great structural variation of the AhR modulators (Figure 12), one might question whether it really could be that all of these compounds actually directly bind to AhR. Denison et al. [34] proposed that many compounds that have shown AhR-mediated responses might only indirectly activate the receptor. Therefore, only the compounds for which binding experiments have been reported in the literature were used as known

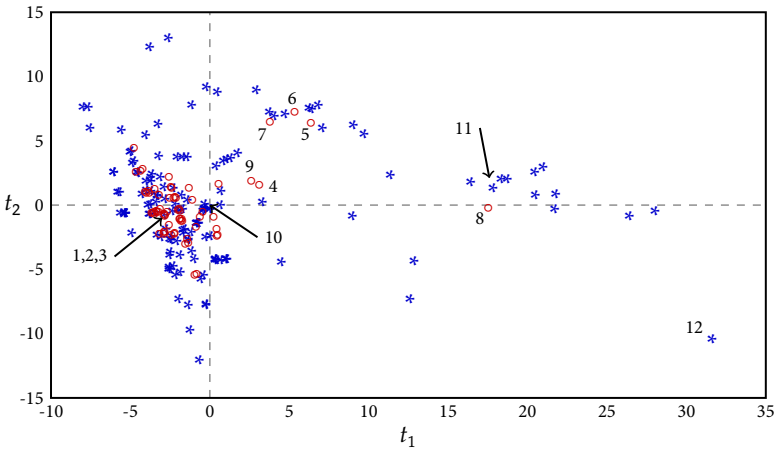


Figure 12: The score plot of the first and second PCs (48 % and 17 % of the explained variation, respectively) of the 214 obtained AhR modulators based on 68 2D and 18 3D descriptors (Paper IV). The 66 AhR binders are marked with red circles, and the remaining AhR modulators are marked with blue stars. The numbers refer to locations of compounds whose structures are given in Figure 11.

AhR ligands for the virtual screening (Paper IV). This final selection consisted of 66 compounds, including some of the more flexible and hydrophilic compounds (Figure 11; **5–8**).

7.2 Potential AhR ligands among the industrial chemicals

The developed virtual screening protocol (Figure 13) included an initial filtration step aiming at identifying chemicals with structural similarities to the 66 known AhR binders (Figure 12). However, the chemical space was defined based solely on the 2D descriptors in order to save computational time. This step was followed by three parallel steps consisting of structural fingerprints, nearest neighbor analysis, and molecular docking. These methods were applied to the 429 industrial chemicals remaining from the initial filtration. The structural fingerprints consisted of 166 predefined structural keys corresponding to various fragments that form the binary fingerprints of the molecules [145], and these strings were compared using the Tanimoto coefficient [132]. The nearest neighbor analysis was based on a PCA of the 66 known AhR ligands and the 429 industrial chemicals using the 2D descriptors and by calculating Euclidean distances in the PCA. The molecular docking of the 429 industrial chemicals and the 66 known AhR ligands was performed using a modified version of a previously reported docking protocol for the used homology model [77]. When the screening protocol was applied to the

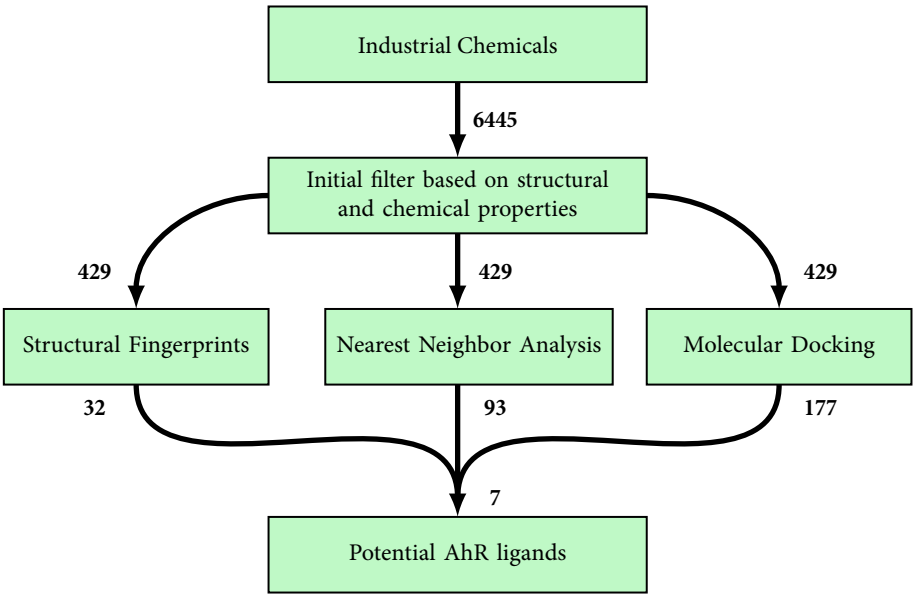


Figure 13: A flowchart illustrating the virtual screening protocol for identifying AhR ligands among industrial chemicals. The numbers next to the boxes correspond to the total number of compounds captured by the previous enrichment step.

set of 6,445 industrial chemicals, the three parallel steps in the virtual screening protocol provided 250 unique industrial chemicals. Furthermore, 41 of these 250 compounds, approximately 16 %, were identified as potential AhR ligands by at least two of the parallel steps. More specifically, 7 compounds were identified by all parallel steps, 12 compounds were identified by the structural fingerprints and nearest neighbor analysis but not the molecular docking, and 22 compounds were identified by the nearest neighbor analysis and the molecular docking but not by the structural fingerprints.

The seven compounds identified by all three enrichment steps included three well-studied environmental contaminants—the pesticide p,p'-dichlorodiphenyl-trichloroethane (pp'-DDT), the brominated flame retardant 2,2',4,4',5-pentabromodiphenylether, and the biocide triclosan (Figure 14; **13–15**). These three compounds have been shown to activate or suppress the AhR pathway *in vitro* [5, 49, 136] and they closely resemble known AhR binders among halogenated aromatic compounds (Figure 11; **1–2**). Among these seven compounds is also a pharmaceutical being used as a calcium blocker, nisoldipine (Figure 14), which has also been shown to activate AhR [150]. Nisoldipine resembles the rather large and globular pharmaceutical nimodipine (Figure 11; **8**).

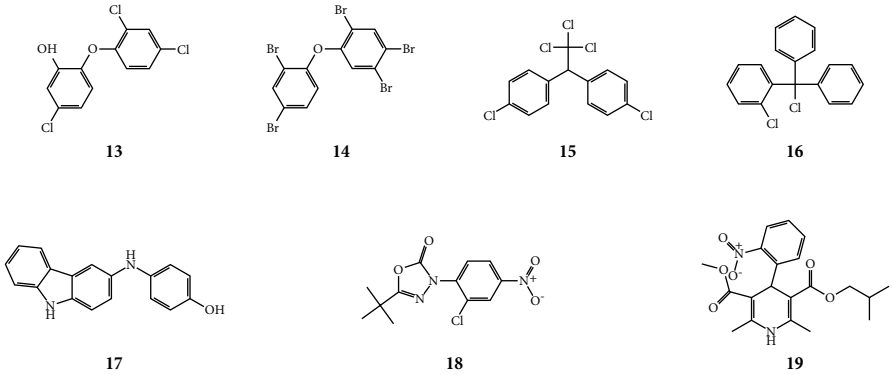


Figure 14: The chemical structures of the seven compounds identified by all three enrichment steps in the virtual screening protocol (Paper IV). **13)** triclosan, **14)** 2,2',4,4',5-pentabromodiphenylether; **15**) p,p'-dichlorodiphenyltrichloroethane (pp'-DDT), **16)** 2-chlorotritylchloride; **17**) 3-p-hydroxy-anilinocarbazole, **18)** 3-(2-chloro-4-nitrophenyl)-5-(1,1-dimethylethyl)-1,3,4-oxadiazol-2(3H)-one, and **19)** nisoldipine.

The remaining three compounds (Figure 14; **16–18**) consisted of a peptide synthesis intermediate (Figure 14; **16**), a dye used for cotton textiles (Figure 14; **17**), and a herbicide synthesis intermediate (Figure 14; **18**). To my knowledge, none of these three compounds have been tested for AhR-mediated effects. 2-Chlorotritylchloride (Figure 14; **16**) is not fully planar but otherwise resembles PCBs and polycyclic aromatic hydrocarbons (Figure 11; **2–3**). The dye 3-p-hydroxy-anilinocarbazole (Figure 14; **17**) resembles the endogenous AhR binders 6-formylindolo[3,2-b]carbazole and 2-(1'H-indole-3'-carbonyl)-thiazole-4-carboxylic acid ester (Figure 11; **4–5**). Finally, the herbicide intermediate 3-(2-chloro-4-nitrophenyl)-5-(1,1-dimethylethyl)1,3,4-oxadiazol-2(3H)-one (Figure 14; **18**) is structurally a mix of the AhR binders flutamide and leflunomide (Figure 11; **6–7**).

Besides pp'-DDT, 2,2',4,4',5-pentabromodiphenyl ether, triclosan, and nisoldipine, another ten compounds among the 41 have been tested for AhR-mediated effects. Seven of these ten additional compounds have been shown to activate the AhR pathway (Table 5). The three remaining compounds, anthracene-9-carbaldehyde (642-31-9), tetradifon (116-29-0), and triphenylphosphine (603-35-0) have been tested for AhR-mediated effects but were found to be inactive [110, 150]. In these tests, the reported purity of anthracene-9-carbaldehyde and tetradifon were above 90 %, while the purity of triphenylphosphine was below 75 %. Hence, based on our findings we suggest triphenylphosphine be retested. Overall, among the 41 industrial chemicals, there were 27 compounds for which, to our knowledge, no testing for AhR-mediated effects has been performed. For future *in vitro* testing regarding AhR-mediated effects, we suggest these 27 compounds and triphenylphosphine be

Table 5: Among the 41 potential AhR ligands, 11 compounds have been reported to activate or suppress the AhR pathway. The number strings show the Chemical Abstracts Service (CAS) numbers of the corresponding compounds.

CAS number	Industrial chemical	Reference
50-29-3	p,p'-dichlorodiphenyltrichloroethane (pp'-DDT)	Wojtowicz et al. [136]
32534-81-9	2,2',4,4',5-pentabromodiphenyl ether	Hamers et al. [49]
3380-34-5	triclosan	Ahn et al. [5]
63675-72-9	nisoldipine	NCBI [150]
115-32-2	dicofol	Chan et al. [29]
120-78-5	di(benzothiazol-2-yl) disulphide	He et al. [52]
25059-80-7	ethyl 4-chloro-2-oxo- 2H-benzothiazole-3-acetate	NCBI [150]
3896-11-5	bumetrizole	Fent et al. [43], Nagayoshi et al. [78]
50471-44-8	vinclozolin	de Oca et al. [33]
5495-84-1	2-isopropylthioxanthone	Peijnenburg et al. [88], NCBI [150]
97-32-5	4-methoxy-3-nitro-N-phenylbenzamide	NCBI [150]

tested (Table 6). In particular, we suggest testing of the three untested compounds, which all three methods identified as potential AhR ligands, be tested. In Table 6, these compounds are given a rank of 1, and the remaining 25 compounds are given a rank of 2.

7.3 Applicability, future adjustments, and remarks on the developed virtual screening protocol

The virtual screening protocol was developed to identify potential AhR ligands among a specific set of industrial chemicals. However, the protocol is not dependent on the tested data set and can thus be applied to other data sets. In this study, we used rat and mouse *in vitro* data to create the initial filter and to identify the known AhR ligands for the ligand-based steps in the protocol. Also, we used a rat homology model for the structure-based step. We focused on rodents because binding energies estimated by this particular rat homology model have been shown to correlate well with experimentally derived competitive binding affinities for PCDDs [77]. Moreover, there is, to our knowledge, very little ligand binding information based on human cell studies reported in the literature compared to rat and mouse data. In Chapter 6, we identified differences in potency (i.e. in REPs) between DLCs depending on whether the experimental data were based on measured *in vitro* responses in human or in rodent cells. Also, not all PCBs were active in the measured *in vitro* responses in human cells (Paper III). However, there were representatives from all three studied chemical classes (PCDDs, PCDFs,

Table 6: The resulting priority list of 28 industrial chemicals from the virtual screening. Chemicals identified by all three enrichments in the virtual screening are given rank 1. Chemicals identified by two, but not all three, enrichment methods are given rank 2. SL stands for combining a ligand-based and structure-based method, and L stands for combining just the ligand-based methods.

CAS number	Chemical name	Rank
31399-83-4	3-(2-chloro-4-nitrophenyl)-5-(1,1-dimethylethyl)-1,3,4--oxadiazol-2(3H)-one	1
42074-68-0	2-chlorotrityl chloride	1
86-72-6	3-p-hydroxyanilino-carbazole	1
1041-00-5	2,2'-vinylenebis[5-methylbenzoxazole]	2SL
1047-16-1	5,12-dihydroquino[2,3-b]acridine-7,14-dione	2SL
132-68-3	3-hydroxy-N-1-naphthyl-2-naphthamide	2SL
2392-48-5	4-chloro-1-(2,4-dichlorophenoxy)2-nitrobenzene	2SL
2475-46-9	1-methylamino-4-ethanolaminoanthraquinone	2SL
2832-40-8	N-[4-[(2-hydroxy-5-methylphenyl)azo] phenyl]acetamide	2SL
2879-15-4	8-benzyl-3,7-dihydro-1,3-dimethyl- 1H-purine-2,6-dione	2SL
56966-52-0	5-chloro-2-(2,4-dichlorophenoxy)aniline	2SL
58979-46-7	N-[5-(diethylamino)-2-[(3,5-dinitro-2-thienyl)azo]phenyl]-acetamide	2SL
76-83-5	chlorotriphenylmethane	2SL
7643-08-5	2-methylthiophenothiazine	2SL
81-98-1	3,9-dibromo-7H-benz[de]anthracen-7-one	2SL
94279-09-1	2-hydroxy-3-[[(4-methylphenyl)sulphonyl]oxy]propyl-tert-decanoate	2SL
94279-10-4	3-hydroxy-2-[[(4-methylphenyl)sulphonyl]oxy]propyl-tert-decanoate	2SL
15307-93-4	2,6-dichloro-N-phenylaniline	2L
1775-95-7	2-amino-5-nitrobenzophenone	2L
24112-48-9	N-[5-(diethylamino)-2-[(2,4-dinitro-phenyl)azo]phenyl] acetamide	2L
26140-60-3	terphenyl	2L
2866-43-5	2,2'-thiophene-2,5-diylbis(benzoxazole)	2L
38982-12-6	3-(9-anthryl)acrylaldehyde	2L
41604-19-7	4-bromo-2-fluoro-1,1'-biphenyl	2L
41642-51-7	N-[2-[(2,6-dicyano-4-nitrophenyl)azo]-5-(diethylamino)-phenyl]acetamide	2L
59383-11-8	2-chloroethyl 3-nitro-p-toluate	2L
603-35-0	triphenylphosphine	2L
81-37-8	2,8-dimethylnaphtho[3,2,1-kl]xanthene	2L

and PCBs) that were active in evoking both human and rodent *in vitro* responses. If this trend were to also apply to other chemical classes, this would suggest that using rodent-based data yields an overly cautious outcome and might provide false positives regarding compounds that are hazardous to human health.

The National Center for Biotechnology Information (NCBI) [150] has reported high-throughput screening results for approximately 10,000 chemicals regarding activation of AhR in human cells. Their work covered approximately one fourth of the 429 chemicals that we screened in our parallel steps. To minimize the influence of impurities, we only included chemicals that had the highest purity rate (>90%), and where the compounds consistently have been stated as active or inactive. This resulted in 38 active chemicals and 56 inactive chemicals. Notably, 14 of the 38 compounds reported as active by the NCBI [150] were not identified as potential AhR ligands by any of our parallel steps, despite being screened in those steps. However, AhR activation might not always be dependent on binding [34, 79]. Compounds that activate AhR by, for instance, cross-talking to other nuclear receptors as suggested by Denison et al. [35], might therefore be hard to locate in a virtual screening based on AhR. Another possibility is that the unidentified active chemicals might bind to the human AhR but not to the rat AhR, which was used in our screening. Another 15 compounds of the 38 were only identified as potential AhR ligands by the molecular docking method. Hence, by delimiting the final selection of potential AhR ligands to those identified by at least two of the methods, fewer of these 38 compounds were identified. To take a more cautious approach, one might want to also include compounds that are only identified by one of the individual steps.

The three parallel steps in the virtual screening protocol include approaches that provide different aspects of ligand identification. The structural fingerprints and nearest neighbor analysis both focus on the ligands; the former addresses the specific functional groups present in the structures, while the latter is based on the chemical, physical, and structural properties of the ligands. The molecular docking differs from the other two methods because it takes the protein properties and the binding pocket into account. By ranking the compounds based on their occurrences in the enrichment from each method, we believe that we have identified the most probable AhR ligands. However, our screening did not discriminate between AhR activators and suppressors. From a health perspective, it might be interesting to keep track of AhR suppressors because they might substantially decrease the effect of AhR activators [11].

8. Concluding summary

In this thesis, we have developed computational approaches to further interpret and use information from *in vitro* studies on human and rodent cells. To study congeneric and species-specific differences among DLCs, we developed an alternative method—the consensus toxicity factor (CTF) approach—to analyze continuous data. At the last TEF reevaluation meeting [123], the available toxicological data mainly consisted of information from rodent studies. This meeting resulted in the currently used weights referred to as the WHO-TEFs. Our analysis of 17 *in vitro* responses for 18 DLCs (Paper III) shows clear differences in the responses for human cells compared to rat, mouse, and guinea pig cells. Applying the CTF approach, yielded clear differences for several DLCs in human CTFs compared to rat CTFs and WHO-TEFs.

In our studies, the human CTF for 23478-PeCDF is as high as the CTF of 2378-TCDD. This means that 23478-PeCDF is estimated to have the same potency as 2378-TCDD. Moreover, specific hexa- and hepta-chlorinated DDs and DFs have more than ten times higher human CTFs than the corresponding WHO-TEFs. In particular, 123478-HxCDF and 1234789-HpCDF have very high CTFs (1 and 0.3, respectively) compared to the corresponding WHO-TEFs (0.1 and 0.01, respectively). To investigate what impact the human CTFs might have if they were to be used to estimate the risk of DLCs in food, a case study based on fish from the Baltic Sea region was performed. Overall, our analysis of using CTFs of DLCs in salmonid fish indicated that the risk estimate expressed in TEQ would be in the same range as if using the WHO-TEFs. The decreased impact of PCBs in the CTF-based calculations was balanced by the increased impact of other congeners, in particular 23478-PeCDF.

The work included in this thesis shows that it is possible to design atom-specific descriptors that convey electronic properties in an unadorned manner, even though the compounds belong to different chemical classes of DLCs. The developed descriptors estimate MO densities or partial charges at the lateral carbon positions of the DLCs. As another approach to identify common electronic features among DLCs, the UV spectra of the DLCs were measured. Due to class-specific peaks for the DLCs, we concluded that only the lower range of the UV spectrum (200–250 nm) is class independent. Electronic descriptors, including some of the above-mentioned atom-specific descriptors and parts of the UV-spectra, were used to establish QSARs for AhR-mediated effects. The used atom-specific descriptors indicate that the lateral positions play an important role in explaining the congeneric differences in the potency of DLCs. Furthermore, the lower part of the UV

range, which reflects the resonance over the benzene rings, indicates that π-bond interactions might be important for the DLCs' interaction with AhR.

The DLCs are believed to primarily exert their toxicity by interacting with AhR [92, 123], and in this thesis we have focused on identifying AhR ligands among 6,445 industrial chemicals. For the analysis, we developed a virtual screening procedure, in which ligand-based and structure-based steps are run in parallel. By ranking the compounds based on their occurrences in the enrichment from each parallel method, we believe that we have identified the most probable AhR ligands. Approximately one fourth of the compounds that were screened in the parallel steps had previously been tested in a high-throughput screening for AhR-mediated effects in human cells [150]. Based on the reported identification of active and inactive chemicals in that screening, our virtual screening protocol correctly classified approximately 60 % of the chemicals. Notably, 15 of the reported active chemicals [150] were only identified as potential AhR ligands by the parallel step consisting of molecular docking. This implies that if one wants to take a more cautious approach, one should also include compounds only identified by one of the parallel steps. Moreover, our study shows that it is difficult to separate AhR-activating and AhR-suppressing chemicals by virtual screening. 11 of the 41 chemicals that we identified as potential AhR ligand chemicals have, in earlier studies, been reported as AhR activators or AhR suppressors. However, most of remaining chemicals (among the 41), have to our knowledge, not been tested for AhR-mediated effects, which make them prime candidates for future testing.

In the future, it would be useful to apply the above-mentioned atom-specific descriptors to other chemical classes of small halogenated aromatic compounds with many congeners. Prime candidates would be PBDDs and PBDFs because congeners of these chemical classes have recently been recommended to be included in the TEF concept [124]. PCNs would also be of interest because they contain specific congeners that have been shown to activate AhR and have previously been considered for inclusion in the TEF concept [41]. In the future, the CTF approach could be applied to *in vitro* data regarding these three chemical classes, as well as for the DLCs analyzed in this study. For future *in vitro* testing regarding AhR-mediated effects, it would be interesting to investigate the outcome of the industrial chemicals reported in this thesis as potential AhR ligands. In the virtual screening, we ranked the industrial chemicals based on their performance in three parallel steps—structural fingerprints, nearest neighbor analysis based on 2D descriptors, and molecular docking. It would also be interesting to test representatives of the structurally diverse compounds ranked high solely by the molecular docking because such compounds might identify new AhR ligands that are structurally different compared to today's known ligands.

9. Acknowledgements

Först och främst vill jag tacka mina handledare **Patrik**, **Anna** och **Mats** för deras tålamod, engagemang och gedigna kunskaper. Patrik, du har alltid försökt se nya möjligheter när idéerna sinat och aldrig givit upp när det känts tungt. Anna, du har haft förmågan att lyfta blicken och ställa just de frågor som behövts för att ge mig klarhet. Och Mats, du har varit en lugn och stabil lyssnare utöver att bidra med ditt breda kunnande. Tillsammans har ni varit en oslagbar trio!

Many thanks to all my colleagues in the SYSTEQ-project. It has always been fun and interesting to spend time with you! In particular, I thank my SYSTEQ-coauthors in Utrecht (**Martin**, **Karin**, and **Majorie**), Brno (**Mirek**, **Jan**, **Simona**, **Petra**, and **Katerina**), Kaiserslautern (**Dieter**, **Sylke**, and **Christiane**), Stockholm (**Lorentz** and **Sandra**), and Umeå (**Tine** and **Mehdi**) for their contributions. Much of the work in this thesis would not have been possible without you! Especially, I thank Karin for answering all my peculiar questions regarding toxicology and for being an encouraging friend throughout this process.

I am grateful for the financial support from the European Commission for the SYSTEQ-project (226694-FP7-ENV-2008-1) and Umeå University for my PhD education. Also, I very much appreciated the travel grant from the Wallenberg Foundation, which enabled my trip to Russia.

Jag vill också passa på att tacka **Petra Fohgelberg** på Livsmedelsverket för fiskdatat för denna avhandling och för att hon tog sig tid att besvara alla frågor däromkring. Also, I thank **Matthew Hogg** for refining the language used in this thesis and for patiently answering questions regarding why I did wrong so I'll learn from my mistakes ;)

Någon som verkligen skall tackas är **David**. Tack David, för att du alltid tagit dig tid att besvara mina frågor, stora som små, och för det arbete du lagt ner på vårt manus och läsandet av min kappa. Du är en klippa! Thank you **Laura** and **Domenico** for your work on the manuscript. Also thank you **Brijesh** for your contributions to my first publication and for your nice lectures in quantum mechanics.

Doktorandtiden är så mycket mer än manuskript och artiklar. Jag vill här ta chansen att tacka alla som gjort denna tid minnesvärd. Tack i synnerhet till alla kollegor på miljökemi för att ni gjort det till en trevlig och intressant arbetsplats att komma till varje dag, och för gott sällskap vid fika, julfester, öl/vin/ost-provningar, spexsångskrivande och mycket annat :)

Tack körgänget på EMG för alla trevliga sångövningar under Ange Turells regi (och tack **Ange** förstås!). Och tack till Nationskören för alla skruvade upptåg och sittningar genom åren. Tack **Eva**, för lån av Zelda-dräkt och leopardutstyrsel!

Tack **Åsa**, **Danke**, **Roger** och **Olle** för alla undervisningsdiskussioner och huvudskakningar över tokigheter som studenter kan ha för sig och skriva ner i sina rapporter. Olle, jag kommer fortfarande ihåg ditt meddelande: ”det går trögt genom isen idag...”.

Tack **Aleks** för de många roliga och givande stunder vi haft tillsammans, både på kontoret och på fritiden. Tack också för fredagsluncherna du arrangerat med fysikgänget: **Johan** & **Daniel Z**, **Narges**, **Pontus**, **Thomas**, **Daniel V** och **Isak**. Tack gänget för att ni förgyllt mina fredagar med allt från udda forskning till träningstips och berättelser om bergsbestigningar.

Thanks **David C**, for filling my spare time with movies (of various quality), strange candy, and excellent Swiss chocolate and company!

Tack **Christian** och **Annett** (och Aron & Simon :) för alla upplyftande ord och trevliga helger tillsammans.

Tack **Lan** och **Marcus** för alla diskussioner kring forskning och livet i stort (och all god vietnamesisk mat!).

Jag vill tacka min mamma **Birgitta** och pappa **Kenneth** för det stöd de givit mig genom åren. Era uppmuntrande ord har varit betydande såhär mot slutet av min doktorsresa när orken emellanåt börjat svikta.

Jag vill slutligen tacka den som betytt allra mest för mig under denna tid, min egen **Eddie**. Du har funnits vid min sida ända sedan kvällen i Kivik -99 och varit mitt stöd genom allt som livet satt i vår väg. Tack för att du finns, och för att du stått ut med mig alla dessa år.

10. Vetenskaplig översikt och sammanfattning

Det finns idag många miljörelaterade frågor att ta sig an, såsom ökande koldioxidutsläpp, växande öknar, försurning av naturen och dalande biologisk mångfald. En viktig miljörelaterad fråga berör det ökande antalet kemikalier som gemene man utsätts för, antingen frivilligt eller ofrivilligt. Vissa kemikalier är tänkta (och ibland inte så väl genomtänkta) tillägg i vardagsprodukter såsom bilsäten och kosmetika, medan andra är oönskade biprodukter vid förbränning av förbrukade material. Till de senare hör en grupp föreningar kallade dioxiner, vilkas egenskaper och effekter har tydligare undersökts i denna avhandling.

Ett antal dioxiner tillsammans med utvalda närbesläktade föreningar (polyklorerade dibenzofuraner och bifenyler) är kända för att ge samma toxiska effekter i animala celler. Dessa föreningar binder till en och samma receptor i cellens cytoplasma, den så kallade arylkolvätereceptorn, eller enklare bara ”dioxinreceptorn”. Vi exponeras alla för dessa ämnen, först via modersmjölken och senare via födan. Troligtvis kommer vi alltid att vara exponerade för dessa ämnen, men frågan är i vilken utsträckning det är behövligt och var gränserna går till att vara skadligt.

Jämfört med tidigare decennier så har utsläppen av dioxiner och de dioxinlika föreningarna minskat till idag men fortfarande finns det geografiska områden som är högt kontaminerade. På grund av dessa föreningars förmåga att anrikas i näringskedjorna tar det tid att märka effekter av våra miljöinsatser. Sammantaget blir det viktigt att se till att riskbedömningen av dessa föreningar är så korrekt som möjligt. För att kunna göra detta krävs en större insikt i vad som gör dessa föreningar skadliga för levande organismer, och hur organismerna reagerar på exponeringen. Det är också viktigt att se framåt och tänka på framtida hot, så att nya problem ses innan de växer sig för stora. I denna avhandling beskrivs de mest kända dioxinerna och dioxinlika föreningarna utifrån sina elektroniska egenskaper för att ge en djupare förståelse för mekanismen bakom deras biologiska verkan. Artskillnader studeras på cellnivå och det undersöks huruvida det finns föreningar (kemikalier) bland de mest producerade i Europa som liknar dioxiner vad gäller kemisk struktur, kemiska och fysikaliska egenskaper samt förmåga att binda till dioxinreceptorn.

Studierna i denna avhandling visar på att det finns stora artskillnader mellan människor och djur beträffande hur känsliga deras celler är för dessa dioxiner och dioxinlika föreningar. För riskbedömning av dessa föreningar i mat, så används inom EU relativa riskfaktorer för att beskriva föreningarnas totala skaderisk. Dessa riskfaktorer kan ses som vikter för att beskriva hur verksam en dioxin eller dioxinlik förening är jämfört med den mest kända och betraktat mest farliga dioxinen, ofta

förkortad TCDD. Våra studier visar att en del dioxiner och polyklorerade furaner är mer verksamma i mänskliga celler än vad som anges i de riskfaktorer som används idag. Samtidigt visar våra studier att dioxinlika polyklorerade bifenyler (PCBer) är mindre verksamma än vad som anges i dagens riskfaktorer. För att göra en helhetsbedömning av vad som skulle hända med den totala skaderisken undersöktes halter av dioxiner och dioxinlika föreningar i fisk. Beräkningar på lax från olika delar av Östersjö-regionen visade att den totala risken i lax skulle bli i samma storleksordning som tidigare, om vi applicerade vårt data. En stor skillnad är dock att bidraget till den totala risken skulle till cirka 98% vara bidrag från dioxiner och polyklorerade furaner. Baserat på de gällande riskfaktorerna så bidrar dioxinlika PCBer, i synnerhet en som kallas PCB 126, med uppemot 50 % av den totala beräknade skaderisken för dessa lax-prov. Även om koncentrationen av PCBer skulle minska i lax, vilket den förhoppningsvis gör i och med den förbjudna användningen av PCBer, så kommer den totala risken i lax att vara hög enligt våra studier.

För att få insikt i varför dioxiner och dioxinlika föreningar är olika verksamma skapades modeller baserade på egenskaper som härstammar från föreningarnas kemiska struktur. Det som visade sig mest gynnsamt för att beskriva humana cellers respons för dessa föreningar var deras elektroniska egenskaper lokaliserade på yttre delar av molekylerna. Närmare bestämt undersöktes de kolatomer som sitter ytterst på föreningarnas aromatiska ringar och som alltid är bundna till kloratomer. Dessa kloratomer kan ha stor betydelse för hur dioxiner och dioxinlika föreningar interagerar med dioxinreceptorn.

Dioxinreceptorn i sig är i stora drag sett fortfarande en gåta; den har ingen klart definerad bindare i organismen utan dess bästa aktivator är just en dioxin, en totalt organismfrämmande förening. Att det fortfarande är oklart hur receptorns 3D-struktur ser ut bidrar också till problematiken. Aktivering av receptorn påverkar bildandet av olika celler och har även visat sig negativt påverka effekten av vaccinering. Denna avhandlings undersökningar—med hjälp av en virtuell version av dioxinreceptorn—visar att väldigt strukturellt olika molekyler kan potentiellt binda till dioxinreceptorn, men att interaktionen då kan tänkas se olika ut. Screeningen av industrikemikalier visade på att det finns kemikalier som både är strukturellt lika kända bindare, delar kemiska och fysikaliska egenskaper med dessa, samt virtuellt sett kan binda till receptorn. En del av dessa kemikalier har visat sig kunna aktivera eller hämma proteinets framfart, men de allra flesta är ännu oskrivna kort vad gäller deras förmåga att påverka dioxinreceptorn.

Bibliography

[1] A. Åberg, M. Tysklind, T. Nilsson, M. MacLeod, A. Hanberg, R. Andersson, S. Bergek, R. Lindberg, and K. Wiberg. Exposure assessment at a PCDD/F contaminated site in Sweden—field measurements of exposure media and blood serum analysis. *Environmental Science and Pollution Research*, 17(1):26–39, 2010.

[2] U.G. Ahlborg. Nordisk dioxinriskbedömning: rapport från en nordisk expertgrupp, 1988.

[3] U.G. Ahlborg, G.C. Becking, L.S. Birnbaum, A. Brouwer, H.J.G.M. Derks, M. Feeley, G. Golor, A. Hanberg, J.C. Larsen, A.K.D. Liem, S.H. Safe, C. Schlatter, F. Waern, M. Younes, and E. Yrjänheikki. Toxic equivalency factors for dioxin-like PCBs: Report on WHO-ECEH and IPCS consultation, December 1993. *Chemosphere*, 28(6):1049–1067, 1994.

[4] U.G. Ahlborg, A. Hanberg, and K. Kenne. Risk assessment of polychlorinated biphenyls (PCBs). Available at https://kiedit.ki.se/sites/default/files/nordpcb-92.pdf [2017-06-15], 1992.

[5] K.C. Ahn, B. Zhao, J. Chen, G. Cherednichenko, E. Sanmarti, M.S. Denison, B. Lasley, I.N. Pessah, D. Kultz, D.P.Y. Chang, S.J. Gee, and B.D. Hammock. In vitro biologic activities of the antimicrobials triclocarban, its analogs, and triclosan in bioassay screens: Receptor-based bioassay screens. *Environmental Health Perspectives*, 116(9):1203–1210, 2008.

[6] C.D. Andersson, T. Karlberg, T. Ekblad, A.E.G. Lindgren, A. Thorsell, S. Spjut, U. Uciechowska, M.S. Niemiec, P. Wittung-Stafshede, J. Weigelt, M. Elofsson, H. Schuler, and A. Linusson. Discovery of ligands for ADP-ribosyltransferases via docking-based virtual screening. *Journal of Medicinal Chemistry*, 55(17):7706–7718, 2012.

[7] P.L. Andersson, P. Haglund, and M. Tysklind. Ultraviolet absorption spectra of all 209 polychlorinated biphenyls evaluated by principal component analysis. *Fresenius' Journal of Analytical Chemistry*, 357(8):1088–1092, 1997.

[8] P.L. Andersson, A. van der Burght, M. van den Berg, and M. Tysklind. Multivariate modeling of polychlorinated biphenyl-induced cyp1a activity in hepatocytes from three different species: Ranking scales and species differences. *Environmental Toxicology and Chemistry*, 19(5):1454–1463, 2000.

[9] S. Arulmozhiraja and M. Morita. Electron affinities and reductive dechlorination of toxic polychlorinated dibenzofurans: A density functional theory study. *Journal of Physical Chemistry A*, 108(16):3499–3508, 2004.

[10] A. Ashek, C. Lee, H. Park, and S.J. Cho. 3D QSAR studies of dioxins and dioxin-like compounds using CoMFA and CoMSIA. *Chemosphere*, 65(3):521–529, 2006.

[11] H. Ashida. Suppressive effects of flavonoids on dioxin toxicity. *Biofactors*, 12(1-4):201–206, 2000.

[12] A.T. Assefa, A. Sobek, K.L. Sundqvist, I. Cato, P. Jonsson, M. Tysklind, and K. Wiberg. Temporal trends of PCDD/Fs in baltic sea sediment cores covering the 20th century. *Environmental Science and Technology*, 48(2):947–953, 2013.

[13] L.L. Aylward, J.J. Collins, K.M. Bodner, M. Wilken, and C.M. Bodnar. Elimination rates of dioxin congeners in former chlorophenol workers from Midland, Michigan. *Environmental Health Perspectives*, 121(1):39–45, 2013.

[14] J.C. Baber, W.A. Shirley, Y. Gao, and M. Feher. The use of consensus scoring in ligand-based virtual screening. *Journal of Chemical Information and Modeling*, 46(1):277–288, 2006.

[15] T. Backhaus and M. Faust. Predictive environmental risk assessment of chemical mixtures: A conceptual framework. *Environmental Science and Technology*, 46(5):2564–2573, 2012.

[16] S. Bandiera, Stephen Safe, and A.B. Okey. Binding of polychlorinated biphenyls classified as either phenobarbitone-, 3-methylcholanthrene- or mixed-type inducers to cytosolic Ah receptor. *Chemico-Biological Interactions*, 39(3):259–277, 1982.

[17] S. Bandiera, T. Sawyer, M. Romkes, B. Zmudzka, L. Safe, G. Mason, B. Keys, and S. Safe. Polychlorinated dibenzofurans (PCDFs) - effects of structure on binding to the 2,3,7,8-TCDD cytosolic receptor protein, AHH induction and toxicity. *Toxicology*, 32(2):131–144, 1984.

[18] R.J. Bartlett and M. Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291–352, 2007.

[19] A.D. Becke. Density-functional thermochemistry. III. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[20] R.D. Beger, D.A. Buzatu, and J.G. Wilkes. Combining NMR spectral and structural data to form models of polychlorinated dibenzodioxins, dibenzofurans, and biphenyls binding to the AhR. *Journal of Computer-Aided Molecular Design*, 16(10):727–740, 2002.

[21] P.A. Behnisch, K. Hosoe, and S. Sakai. Brominated dioxin-like compounds: in vitro assessment in comparison to classical dioxin-like compounds and other polyaromatic compounds. *Environment International*, 29(6):861–877, 2003.

[22] A. Bessede, M. Gargaro, M.T. Pallotta, D. Matino, G. Servillo, C. Brunacci, S. Bicciato, E.M.C. Mazza, A. Macchiarulo, C. Vacca, R. Iannitti, L. Tissi, C. Volpi, M.L. Belladonna, C. Orabona, R. Bianchi, T.V. Lanz, M. Platten, M.A. Della Fazia, D. Piobbico, T. Zelante, H. Funakoshi, T. Nakamura, D. Gilot, M.S. Denison, G.J. Guillemin, J.B. DuHadaway, G.C. Prendergast, R. Metz, M. Geffard, L. Boon, M. Pirro, A. Iorio, B. Veyret, L. Romani, U. Grohmann, F. Fallarino, and P. Puccetti. Aryl hydrocarbon receptor control of a disease tolerance defence pathway. *Nature*, 511(7508):184–190, 2014.

[23] C. Bissantz, B. Kuhn, and M. Stahl. A medicinal chemist's guide to molecular interactions. *Journal of Medicinal Chemistry Perspectives*, 53(5):5061–5084, 2010.

[24] W.H. Bisson, D.C. Koch, E.F. O'Donnell, S.M. Khalil, N.I. Kerkvliet, R.L. Tanguay, R. Abagyan, and S.K. Kolluri. Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new ahr ligands. *Journal of Medicinal Chemistry*, 52(18):5635–5641, 2009.

[25] L.A. Boule, B. Winans, and B.P. Lawrence. Effects of developmental activation of the AhR on CD4+ T-Cell responses to influenza virus infection in adult mice. *Environmental Health Perspectives*, 122(11):1201–1208, 2014.

[26] D.J. Brown, M. Chu, I. Van Overmeire, A. Chu, and G.C. Clark. Determination of REP values for the CALUX® bioassay and comparison to the WHO TEF values. *Organohalogen Compounds*, 53:211–214, 2001.

[27] J.M.B. Carlstedt-Duke. Tissue distribution of the receptor for 2,3,7,8-tetrachlorodibenzo-p-dioxin in the rat. *Cancer Research*, 39(8):3172–3176, 1979.

[28] L.A. Carver, J.B. Hogenesch, and C.A. Bradfield. Tissue-specific expression of the rat Ah-receptor ARNT messenger-RNAs. *Nucleic Acids Research*, 22(15):3038–3044, 1994.

[29] W. Chan, J. Liao, C. Chou, P. Chan, C. Wei, and T. Ueng. Induction of CYP1A1, 2B, 2E1 and 3A in rat liver by organochlorine pesticide dicofol. *Toxicology Letters*, 190(2):150–155, 2009.

[30] F.P. Colonna, G. Distefano, V. Galasso, K.J. Irgolic, C.E. King, and G.C. Pappalardo. The conformation, UV-absorption spectra and photoelectron spectra of phenoxachalcogenins. *Journal of Organometallic Chemistry*, 146(3):235–244, 1978.

[31] M.T.D. Cronin and T.W. Schultz. Pitfalls in QSAR. *Journal of Molecular Structure: THEOCHEM*, 622(1):39–51, 2003.

[32] M.D. Curtis, R.K. Lee, and A.L. Allred. Ultraviolet spectra of biphenyl derivatives of group 4B. *Journal of the American Chemical Society*, 89(20):5150–5152, 1967.

[33] F.G.G. de Oca, M. de Lourdes López-González, D.C. Escobar-Wilches, R. Chavira-Ramírez, and A. Sierra-Santoyo. Vinclozolin modulates hepatic cytochrome P450 isoforms during pregnancy. *Reproductive Toxicology*, 53:119–126, 2015.

[34] M.S. Denison and S.R. Nagy. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annual Review of Pharmacology and Toxicology*, 43:309–334, 2003.

[35] M.S. Denison, A.A. Soshilov, G. He, D.E. DeGroot, and B. Zhao. Exactly the same but different: Promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicological Sciences*, 124(1):1–22, 2011.

[36] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, and J.J.P. Stewart. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society*, 107(13):3902–3909, 1985.

[37] J. Diao, Y. Li, S. Shi, Y. Sun, and Y. Sun. QSAR models for predicting toxicity of polychlorinated dibenzo-p-dioxins and dibenzofurans using quantum chemical descriptors. *Bulletin of Environmental Contamination and Toxicology*, 85(2):109–115, 2010.

[38] K.M. Dolwick, J.V. Schmidt, L.A. Carver, H.I. Swanson, and C.A. Bradfield. Cloning and expression of a human receptor cDNA. *Molecular Pharmacology*, 44(5):911–917, 1993.

[39] L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström, and S. Wold. *Multi- and Megavariate Data Analysis. Part I : Basic Principles and Applications*. Umetrics AB, Box 7960, S-907 19 Umeå, Sweden, second edition, 2006.

[40] L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström, and S. Wold. *Multi- and Megavariate Data Analysis. Part II : Advanced Applications and Methods Extensions*. Umetrics AB, Box 7960, S-907 19 Umeå, Sweden, second edition, 2006.

[41] J. Falandysz, A. Fernandes, E. Gregoraszczuk, and M. Rose. The toxicological effects of halogenated naphthalenes: A review of aryl hydrocarbon receptor-mediated (dioxinlike) relative potency factors. *Journal of Environmental Science and Health, Part C: Environmental Carcinogenesis and Ecotoxicology Reviews*, 32(3):239–272, 2014.

[42] E. Fattore, C. Trossvik, and H. Hakansson. Relative potency values derived from hepatic vitamin A reduction in male and female Sprague-Dawley rats following subchronic dietary exposure to individual polychlorinated dibenzo-p-dioxin and dibenzofuran congeners and a mixture thereof. *Toxicology and Applied Pharmacology*, 165(3):184–194, 2000.

[43] K. Fent, G. Chew, J. Li, and E. Gomez. Benzotriazole UV-stabilizers and benzotriazole: antiandrogenic activity in vitro and activation of aryl hydrocarbon receptor pathway in zebrafish eleuthero-embryos. *Science of the Total Environment*, 482:125–136, 2014.

[44] D.J. Funk, R.C. Oldenborg, D.P. Dayton, J.P. Lacosse, J.A. Draves, and T.J. Logan. Gas-phase absorption and laser-induced fluorescence measurements of representative polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and a polycyclic aromatic hydrocarbon. *Applied Spectroscopy*, 49(1):105–114, 1995.

[45] H.J. Geise, A.T.H. Lenstra, C. Deborst, and G.W.H. Moes. 2,2',4,4',5,5'-Hexachlorobiphenyl (PCB 153). *Acta Crystallographica Section C: Crystal Structure Communications*, 42(9):1176–1178, 1986.

[46] H.J. Geyer, K.W. Schramm, E.A. Feicht, A. Behechti, C. Steinberg, R. Bruggemann, H. Poiger, B. Henkelmann, and A. Kettrup. Half-lives of tetra-, penta-, hexa-, hepta-, and octachlorodibenzo-p-dioxin in rats, monkeys, and humans - a critical review. *Chemosphere*, 48(6):631–644, 2002.

[47] H. Gohlke and G. Klebe. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angewandte Chemie International Edition*, 41(5):2644–2676, 2002.

[48] P. Gramatica and E. Papa. Screening and ranking of POPs for global half-life: QSAR approaches for prioritization based on molecular structure. *Environmental Science and Technology*, 41(8):2833–2839, 2007.

[49] T. Hamers, J.H. Kamstra, E. Sonneveld, A.J. Murk, M.H.A. Kester, P.L. Andersson, J. Legler, and A. Brouwer. In vitro profiling of the endocrine-disrupting potency of brominated flame retardants. *Toxicological Sciences*, 92:157–173, 2006.

[50] M. Hansson, T. Grimstad, and C. Rappe. Occupational exposure to polychlorinated dibenzo-p-dioxins and dibenzofurans in a magnesium production plant. *Occupational and Environmental Medicine*, 52(12):823–826, 1995.

[51] L.C. Haws, S.H. Su, M. Harris, M.J. DeVito, N.J. Walker, W.H. Farland, B. Finley, and L.S. Birnbaum. Development of a refined database of mammalian relative

potency estimates for dioxin-like compounds. *Toxicological Sciences*, 89(1):4–30, 2006.

[52] G. He, B. Zhao, and M.S. Denison. Identification of benzothiazole derivatives and polycyclic aromatic hydrocarbons as aryl hydrocarbon receptor agonists present in tire extracts. *Environmental Toxicology and Chemistry*, 30(8):1915–1925, 2011.

[53] C. Heilmann, E. Budtz-Jørgensen, F. Nielsen, B. Heinzow, P. Weihe, and P. Grandjean. Serum concentrations of antibodies against vaccine toxoids in children exposed perinatally to immunotoxicants. *Environmental Health Perspectives*, 118(10):1434–1438, 2010.

[54] E.S. Heimstad, P.M. Bastos, J. Eriksson, A. Bergman, and M. Harju. Quantitative structure - photodegradation relationships of polybrominated diphenyl ethers, phenoxyphenols and selected organochlorines. *Chemosphere*, 77(7):914–921, 2009.

[55] J. Hu, M. Zheng, W. Liu, C. Li, Z. Nie, G. Liu, K. Xiao, and S. Dong. Occupational exposure to polychlorinated dibenzo-p-dioxins and dibenzofurans, dioxin-like polychlorinated biphenyls, and polychlorinated naphthalenes in workplaces of secondary nonferrous metallurgical facilities in China. *Environmental Science and Technology*, 47(14):7773–7779, 2013.

[56] O. Humblet, L. Birnbaum, E. Rimm, M.A. Mittleman, and R. Hauser. Dioxins and cardiovascular disease mortality. *Environmental Health Perspectives*, 116(11):1443–1448, 2008.

[57] J.E. Jackson. *A User's Guide to Principal Components*. John Wiley, New York, 1991.

[58] J.H. Jensen. *Molecular Modeling Basics*. CRC Press, 2010.

[59] C.W. Johnson, W.C. Williams, C.B. Copeland, M.J. DeVito, and R.J. Smialowicz. Sensitivity of the SRBC PFC assay versus ELISA for detection of immunosuppression by TCDD and TCDD-like congeners. *Toxicology*, 156(1):1–11, 2000.

[60] I. Kania-Korwel, S. Parkin, L.W. Robertson, and H.J. Lehmler. 2,4,5-Trichlorobiphenyl. *Acta Crystallographica Section E-Structure Reports Online*, 60:O1652–O1653, 2004.

[61] A.S. Kende and J.J. Wade. Synthesis of new steric and electronic analogs of 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Environmental Health Perspectives*, 5:49–57, 1973.

[62] C.D. Klaassen and J.B Watkins III. *Casarett and Doull's Essentials of Toxicology*. The McGraw-Hill Companies, Inc., 2003.

[63] C.M. Klinge, J.L. Bowers, P.C. Kulakosky, K.K. Kamboj, and H.I. Swanson. The aryl hydrocarbon receptor (AHR):AHR nuclear translocator (ARNT) heterodimer interacts with naturally occurring estrogen response elements. *Molecular and Cellular Endocrinology*, 157(1-2):105–119, 1999.

[64] M.H. Kutner, C.J. Nachtsheim, J. Neter, and W. Li. *Applied Linear Statistical Models*. McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY, 10020., fifth edition, 2005.

[65] A.R. Leach. *Molecular Modelling. Principles and Applications*. Pearson Education Limited, Harlow, England, second edition, 2001.

[66] C. Lee, W. Yang, and R.G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785–789, 1988.

[67] D.H. Lee, I.K. Lee, M. Porta, M. Steffes, and D.R. Jr. Jacobs. Relationship between serum concentrations of persistent organic pollutants and the prevalence of metabolic syndrome among non-diabetic adults: results from the National Health and Nutrition Examination Survey 1999-2002. *Diabetologia*, 50(9):1841–1851, 2007.

[68] H.J. Lehmler, S. Parkin, and L.W. Robertson. 2,3,4 '-Trichlorobiphenyl. *Acta Crystallographica Section E-Structure Reports Online*, 57:o111–o112, 2001.

[69] D.J. Livingstone. Building QSAR models: A practical guide. In M.T.D. Cronin and D.J. Livingstone, editors, *Predicting Chemical Toxicity and Fate*, page 151–170. CRC Press, Boca Raton, FL, USA, 2004.

[70] E. Lo Piparo, K. Koehler, A. Chana, and E. Benfenati. Virtual screening for aryl hydrocarbon receptor binding prediction. *Journal of Medicinal Chemistry*, 49(19):5702–5709, 2006.

[71] F. Luan, W.P. Ma, X.Y. Zhang, H.X. Zhang, M.C. Liu, Z.D. Hu, and B.T. Fan. QSAR study of polychlorinated dibenzodioxins, dibenzofurans, and biphenyls using the heuristic method and support vector machine. *QSAR and Combinatorial Science*, 25(1):46–55, 2006.

[72] M. MacLeod, M. Breitholtz, I.T. Cousins, C.A. de Wit, L.M. Persson, C. Ruden, and M.S. McLachlan. Identifying chemicals that are planetary boundary threats. *Environmental Science and Technology*, 48(19):11057–11063, 2014.

[73] N. Marinkovic, D. Pasalic, G. Ferencak, B. Grskovic, and A.S. Rukavina. Dioxins and human toxicity. *Arhiv Za Higijenu Rada I Toksikologiju-Archives of Industrial Hygiene and Toxicology*, 61(4):445–453, 2010.

[74] G. Mason, K. Farrell, B. Keys, J. Piskorskapliszczynska, L. Safe, and S. Safe. Polychlorinated dibenzo-p-dioxins: Quantitative in vitro and in vivo structure-activity relationships. *Toxicology*, 41(1):21–31, 1986.

[75] A. Miller, J.E. Hedman, E. Nyberg, P. Haglund, I.T. Cousins, K. Wiberg, and A. Bignert. Temporal trends in dioxins (polychlorinated dibenzo-p-dioxin and dibenzofurans) and dioxin-like polychlorinated biphenyls in baltic herring (clupea harengus). *Marine Pollution Bulletin*, 73(1):220–230, 2013.

[76] P. Miramontes. DNA and RNA physicochemical constraints, cellular automata and molecular evolution. Report at the workshop "Cellular Automata: Theory and Applications" in Los Alamos, New Mexico, 1989.

[77] I. Motto, A. Bordogna, A.A. Soshilov, M.S. Denison, and L. Bonati. New aryl hydrocarbon receptor homology model targeted to improve docking reliability. *Journal of Chemical Information and Modeling*, 51(11):2868–2881, 2011.

[78] H. Nagayoshi, K. Kakimoto, S. Takagi, Y. Konishi, K. Kajimura, and T. Matsuda. Benzotriazole ultraviolet stabilizers show potent activities as human aryl hydrocarbon receptor ligands. *Environmental Science and Technology*, 49(1):578–587, 2014.

[79] L.P. Nguyen and C.A. Bradfield. The search for endogenous activators of the aryl hydrocarbon receptor. *Chemical Research in Toxicology*, 21(1):102–116, 2008.

[80] E. Nyberg, S. Faxneld, S. Danielsson, U. Eriksson, A. Miller, and A. Bignert. Temporal and spatial trends of PCBs, DDTs, HCHs, and HCB in swedish marine biota 1969–2012. *AMBIO*, 44(3):484–497, 2015.

[81] T. Oberg. Virtual screening for environmental pollutants: Structure-activity relationships applied to a database of industrial chemicals. *Environmental Toxicology and Chemistry*, 25(4):1178–1183, 2006.

[82] F. Ohtake, Y. Fujii-Kuriyama, K. Kawajiri, and S. Kato. Cross-talk of dioxin and estrogen receptor signals through the ubiquitin system. *Journal of Steroid Biochemistry and Molecular Biology*, 127(1-2):102–107, 2011.

[83] J. Olivero-Verbel, R. Vivas-Reyes, L. Pacheco-Londono, B. Johnson-Restrepo, and K. Kannan. Discriminant analysis for activation of the aryl hydrocarbon receptor by polychlorinated naphthalenes. *Journal of Molecular Structure-Theochem*, 678(1–3):157–161, 2004.

[84] T.I. Oprea, L. Kurunczi, M. Olah, and Z. Simon. MTD-PLS: A PLS-based variant of the MTD method. A 3D-QSAR analysis of receptor affinities for a series of halogenated dibenzoxin and biphenyl derivatives. *SAR and QSAR in Environmental Research*, 12(1-2):75–92, 2001.

[85] E. Papa and P. Gramatica. Screening of persistent organic pollutants by QSPR classification models: A comparative study. *Journal of Molecular Graphics and Modelling*, 27(1):59–65, 2008.

[86] E. Papa and P. Gramatica. QSPR as a support for the EU REACH regulation and rational design of environmentally safer chemicals: PBT identification from molecular structure. *Green Chemistry*, 12(5):836–843, 2010.

[87] H. Park, E. Park, and Y. Chang. Ten-year time trend of dioxins in human serum obtained from metropolitan populations in Seoul, Korea. *Science of The Total Environment*, 470-471:1338–1345, 2014.

[88] A. Peijnenburg, J. Riethof-Poortman, H. Baykus, L. Portier, T. Bovee, and R. Hoogenboom. AhR-agonistic, anti-androgenic, and anti-estrogenic potencies of 2-isopropylthioxanthone (ITX) as determined by in vitro bioassays and gene expression profiling. *Toxicology in Vitro*, 24(6):1619–1628, 2010.

[89] R. Pohjanvirta. *The AH Receptor in Biology and Toxicology*. John Wiley & Sons, 2011.

[90] A. Poland and J.C. Knutson. 2,3,7,8-tetrachlorodibenzo-para-dioxin and related halogenated aromatic-hydrocarbons - examination of the mechanism of toxicity. *Annual Review of Pharmacology and Toxicology*, 22(1):517–554, 1982.

[91] T. Puzyn, J. Leszczynski, and M.T.D. Cronin. *Challenges and Advances in Computational Chemistry and Physics. Volume 8 : Recent Advances in QSAR Studies*. Springer, Dordrecht Heidelberg London New York, 2010.

[92] S. Safe. Polychlorinated biphenyls (PCBs), dibenzo-para-dioxins, dibenzofurans (PCDFs), and related compounds - environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). *Critical Reviews in Toxicology*, 21(1):51–88, 1990.

[93] S.H. Safe. Comparative toxicology and mechanism of action of polychlorinated dibenzo-p-dioxins and dibenzofurans. *Annual Review of Pharmacology and Toxicology*, 26(1):371–399, 1986.

[94] S.H. Safe. Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment. *Critical Reviews in Toxicology*, 24(2):87–149, 1994.

[95] T. Sawyer and S. Safe. PCB isomers and congeners: induction of aryl hydrocarbon hydroxylase and ethoxyresorufin o-deethylase enzyme activities in rat hepatoma cells. *Toxicological Letters*, 13:87–93, 1982.

[96] M. Scholze, E. Silva, and A. Kortenkamp. Extending the applicability of the dose addition model to the assessment of chemical mixtures of partial agonists by using a novel toxic unit extrapolation method. *PLOS One*, 9(2), 2014.

[97] G.A.F. Seber and A.J. Lee. *Linear Regression Analysis*. Wiley & Sons, Inc., Hoboken, New Jersey, second edition, 2003.

[98] S.D. Seidel, G.M. Winters, W.J. Rogers, M.H. Ziccardi, V. Li, B. Keser, and M.S Denison. Activation of the Ah receptor signaling pathway by prostaglandins. *Journal of Biochemical and Molecular Toxicology*, 15(4):187–196, 2001.

[99] L. Shade, R.D. Beger, and J.G. Wilkes. The use of carbon thirteen nuclear magnetic resonance spectra to predict dioxin and furan binding affinities to the aryl hydrocarbon receptor. *Environmental Toxicology and Chemistry*, 22(3):501–509, 2003.

[100] N.S. Shaikh, S. Parkin, G. Luthe, and H. Lehmler. The three-dimensional structure of 3,3 ',4,4 '-tetrachlorobiphenyl, a dioxin-like polychlorinated biphenyl (PCB). *Chemosphere*, 70(9):1694–1698, 2008.

[101] C.J. Sinal and J.R. Bend. Aryl hydrocarbon receptor-dependent induction of cyp1a1 by bilirubin in mouse hepatoma hepa 1c1c7 cells. *Molecular Pharmacology*, 52(4):590–599, 1997.

[102] S.S. So and M. Karplus. Three-dimensional quantitative structure-activity relationships from molecular similarity matrices and genetic neural networks. 2. Applications. *Journal of Medicinal Chemistry*, 40(26):4360–4371, 1997.

[103] S. Sorg, M. Zennegg, P. Schmid, R. Fedosyuk, R. Valikhnovskyi, O. Gaide, V. Kniazevych, and J.H. Saurat. 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) poisoning in Victor Yushchenko: identification and measurement of TCDD metabolites. *Lancet*, 374(9696):1179–1185, 2009.

[104] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, and M.J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[105] S.B. Stølevik, U.C. Nygaard, E. Namork, M. Haugen, H.M Meltzer, J. Alexander, H.K. Knutsen, I. Aaberge, K. Vainio, H. van Loveren, M. Løvik, and B. Granum. Prenatal exposure to polychlorinated biphenyls and dioxins from the maternal diet may be associated with immunosuppressive effects that persist into early childhood. *Food and Chemical Toxicology*, 51(0):165–172, 2013.

[106] F. Svensson, A. Karlén, and C. Sköld. Virtual screening data fusion using both structure- and ligand-based methods. *Journal of Chemical Information and Modeling*, 52:225–232, 2012.

[107] S.L Swann, S.P Brown, S.W Muchmore, H. Patel, P. Merta, J. Locklear, and P.J. Hajduk. A unified, probabilistic framework for structure-and ligand-based virtual screening. *Journal of Medicinal Chemistry*, 54(5):1223–1232, 2011.

[108] E. Swedenborg and I. Pongratz. AhR and ARNT modulate ER signaling. *Toxicology*, 268(3):132–138, 2010.

[109] A. Takagi, A. Hirose, Y. Hirubayushi, T. Kaneko, M. Ema, and J. Kanno. Assessment of the cleft palate induction by seven PCDD/F congeners in the mouse fetus. *Organohalogen Compounds*, 64:336–338, 2003.

[110] S. Takeushi, M. Iida, H. Yabushita, T. Matsuda, and H. Kojima. In vitro screening for aryl hydrocarbon receptor agonistic activity in 200 pesticides using a highly sensitive reporter cell line, DR-EcoScreen cells, and in vivo mouse liver cytochrome P450-1A induction by propanil, diuron and linuron. *Chemosphere*, 74(1):155–165, 2008.

[111] N. Tang, J. Liu, P.J. Coenraads, L. Dong, L. Zhao, S. Ma, X. Chen, C. Zhang, X. Ma, W. Wei, P. Zhang, and Z. Bai. Expression of AhR, CYP1A1, GSTA1, c-fos and TGF-α in skin lesions from dioxin-exposed humans with chloracne. *Toxicological Letters*, 177(3):182–187, 2008.

[112] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic Press, 2003.

[113] R. Todeschini and V. Consonni. *Molecular descriptors for chemoinformatics*. Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim, 2009.

[114] J. Trygg and S. Wold. Orthogonal projections to latent structures. *Journal of Chemometrics*, 16(1):119–128, 2002.

[115] K. Tuppurainen and J. Ruuskanen. Electronic eigenvalue (EEVA): a new QSAR/QSPR descriptor for electronic substituent effects based on molecular orbital energies. a QSAR approach to the Ah receptor binding affinity of polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs). *Chemosphere*, 41(6):843–848, 2000.

[116] D.B. Turner, P. Willett, A.M. Ferguson, and T. Heritage. Evaluation of a novel infrared range vibration-based descriptor (EVA) for QSAR studies .1. general application. *Journal of Computer-Aided Molecular Design*, 11(4):409–422, 1997.

[117] M. Tysklind, K. Lundgren, and C. Rappe. Ultraviolet-absorption characteristics of all tetrachlorinated to octachlorinated dibenzofurans. *Chemosphere*, 27(4):535–546, 1993.

[118] M. Tysklind, K. Lundgren, C. Rappe, L. Eriksson, J. Jonsson, and M. Sjostrom. Multivariate quantitative structure-activity-relationships for polychlorinated dibenzo-p-dioxins and dibenzofurans. *Environmental Toxicology and Chemistry*, 12(4):659–672, 1993.

[119] M. Tysklind, D. Tillitt, L. Eriksson, K. Lundgren, and C. Rappe. A toxic equivalency factor scale for polychlorinated dibenzofurans. *Fundamental and Applied Toxicology*, 22(2):277–285, 1994.

[120] H. Uemura, K. Arisawa, M. Hiyoshi, A. Kitayama, H. Takami, F. Sawachika, S. Dakeshita, K. Nii, H. Satoh, Y. Sumiyoshi, K. Morinaga, K. Kodama, T. Suzuki, M. Nagai, and T. Suzuki. Prevalence of metabolic syndrome associated with body burden levels of dioxin and related compounds among Japan's general population. *Environmental Health Perspectives*, 117(4):568–573, 2009.

[121] A.P.J.M. van Birgelen, M.J. DeVito, J.M. Akins, D.G. Ross, J.J. Diliberto, and L.S. Birnbaum. Relative potencies of polychlorinated dibenzo-p-dioxins,

dibenzofurans, and biphenyls derived from hepatic porphyrin accumulation in mice. *Toxicology and Applied Pharmacology*, 138(1):98–109, 1996.

[122] M. van den Berg, L. Birnbaum, A.T.C. Bosveld, B. Brunstrom, P. Cook, M. Feeley, J.P. Giesy, A. Hanberg, R. Hasegawa, S.W. Kennedy, T. Kubiak, J.C. Larsen, F.X.R. Van Leeuwen, A.K.D. Liem, C. Nolt, R.E. Peterson, L. Poellinger, S. Safe, D. Schrenk, D. Tillitt, M. Tysklind, M. Younes, F. Waern, and T. Zacharewskr. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. *Environmental Health Perspectives*, 106(12):775–792, 1998.

[123] M. van den Berg, L.S. Birnbaum, M. Denison, M. De Vito, W. Farland, M. Feeley, H. Fiedler, H. Hakansson, A. Hanberg, L. Haws, M. Rose, S. Safe, D. Schrenk, C. Tohyama, A. Tritscher, J. Tuomisto, M. Tysklind, N. Walker, and R.E. Peterson. The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicological Sciences*, 93(2):223–241, 2006.

[124] M. van den Berg, M.S. Denison, L.S. Birnbaum, M.J. DeVito, H. Fiedler, J. Falandysz, M. Rose, D. Schrenk, S. Safe, C. Tohyama, A. Tritscher, M. Tysklind, and R.E. Peterson. Polybrominated dibenzo-p-dioxins, dibenzofurans, and biphenyls: Inclusion in the toxicity equivalency factor concept for dioxin-like compounds. *Toxicological Sciences*, 133(2):197–208, 2013.

[125] M.B.M. van Duursen, J.T. Sanderson, M. van der Bruggen, J. van der Linden, and M. van den Berg. Effects of several dioxin-like compounds on estrogen metabolism in the malignant MCF-7 and nontumorigenic MCF-10A human mammary epithelial cell lines. *Toxicology and Applied Pharmacology*, 190(3):241–250, 2003.

[126] M.B.M. van Duursen, K.I. van Ede, and M. van den Berg. One TEF concept does not fit all: the case for human risk assessment of Polychlorinated Biphenyls. *Current Opinion in Toxicology*, 2017.

[127] K.I. van Ede, P.L. Andersson, K.P.J. Gaisch, M. van den Berg, and M.B.M. van Duursen. Comparison of intake and systemic relative effect potencies of dioxin-like compounds in female mice after a single oral dose. *Environmental Health Perspectives*, 121(7):847–853, 2013.

[128] C.J. van Leeuwen and T.G. Vermeire, editors. *Risk assessment of Chemicals: An Introduction*. Springer, second edition, 2007.

[129] S.H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[130] N.J. Walker, P.W. Crockett, A. Nyska, A.E. Brix, M.P. Jokinen, D.M. Sells, J.R. Hailey, M. Easterling, J.K. Haseman, M. Yin, M.E. Wyde, J.R. Bucher, and C.J. Portier. Dose-additive carcinogenicity of a defined mixture of "dioxin-like compounds". *Environmental Health Perspectives*, 113(1):43–48, 2005.

[131] E.B. Wedebye, M. Dybdahl, N.G. Nikolov, S.O. Jonsdottir, and J.R. Niemela. QSAR screening of 70,983 REACH substances for genotoxic carcinogenicity, mutagenicity and developmental toxicity in the ChemScreen project. *Reproductive Toxicology*, 55:64–72, 2015.

[132] P. Willett. Chemical similarity searching. *Journal of Chemical Information and Modeling*, 38(6):983–996, 1998.

[133] P. Willett. Enhancing the effectiveness of ligand-based virtual screening using data fusion. *QSAR and Combinatorial Science*, 25(12):1143–1152, 2006.

[134] P. Willett. Similarity searching using 2D structural fingerprints. *Chemoinformatics and Computational Chemical Biology*, page 133–158, 2011.

[135] B. Winans, M.C. Humble, and B.P. Lawrence. Environmental toxicants and the developing immune system: A missing link in the global battle against infectious disease? *Reproductive Toxicology*, 31(3):327–336, 2011.

[136] A.K. Wojtowicz, E. Honkisz, Zieba-Przybylska, T. Milewicz, and M. Kajta. Effects of two isomers of DDT and their metabolite DDE on CYP1A1 and AhR function in human placental cells. *Pharmacological Reports*, 63(6):1460–1468, 2011.

[137] S. Wold. Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems*, 30(1):109–115, 1995.

[138] S. Wold, M. Sjostrom, and L. Eriksson. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2):109–130, 2001.

[139] M. Yanagisawa, K. Hayamizu, and O. Yamamoto. C-13 chemical-shifts of chlorinated biphenyls. *Magnetic Resonance in Chemistry*, 24(11):1013–1014, 1986.

[140] M. Zeiger, R. Haag, J. Höckel, D. Schrenk, and H.-J. Schmitz. Inducing effects of dioxin-like polychlorinated biphenyls on CYP1A in the human hepatoblastoma cell line HepG2, the rat hepatoma cell line H4IIE, and rat primary hepatocytes: Comparison of relative potencies. *Toxicological Sciences*, 63(1):65–73, 2001.

[141] J.W. Zou, G.X. Hu, Y.J. Jiang, M. Zeng, S.L. Zhuang, and Q.S. Yu. Surface electrostatic potentials of polychlorinated dibenzofurans and their applications in quantitative structure-toxicity relationships. *Chinese Journal of Structural Chemistry*, 24(4):407–412, 2005.

[142] Managing the dioxin problem in the Baltic region with focus on sources to air and fish. Final report from the research project BalticPOPs. Available at http://www.naturvardsverket.se/documents/publikationer6400/978-91-620-6566-9.pdf?pid=8228 [2017-06-15].

[143] European Chemicals Agency (ECHA). Guidance on information requirements and chemical safety assessment, 2011. Available at http://echa.europa.eu/documents/10162/13643/information_requirements_r4_en.pdf [2017-06-15].

[144] COMMISSION REGULATION (EC) No 1881/2006 of 19 December 2006 setting maximum levels for certain contaminants in foodstuffs. Available at http://eur-lex.europa.eu/legal-content/en/txt/pdf/?uri=celex:32006r1881&rid=2 [2017-06-15].

[145] MAACS keys, MDL Information Systems, Inc., 14600 Catalina Street, San Leandro, CA 94577.

[146] Molecular operating environment (MOE) 2006.08, Chemical Computing group, Quebec, Canada, 2008.

[147] North Atlantic Treaty Organization/Committee on the Challenges of Modern Society. International Toxicity Equivalency Factor (I-TEF) method of risk assessment for complex mixtures of dioxins and related compounds. Report number 176, 1988.

[148] OECD principles for the validation, for the regulatory purposes, of (quantitative) structure-activity relationship models, 2004. Available at http://www.oecd.org/chemicalsafety/risk-assessment/37849783.pdf [2017-06-15].

[149] OECD guidance on grouping of chemicals, 2007. Available at http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2007)28&doclanguage=en [2017-06-15].

[150] National Center for Biotechnology Information. PubChem BioAssay Database; AID=743122, https://pubchem.ncbi.nlm.nih.gov/bioassay/743122 [2017-01-26].

[151] Kontroll av främmande ämnen i livsmedel 2012-2013. Livsmedelsverket Rapport 18, 2015. Available at http://www.livsmedelsverket.se/globalassets/rapporter/2015/kontroll-av-frammande-amnen-i-livsmedel-2012-2013-rapport-18---2015.pdf [2017-06-15].

[152] Stockholm Convention on persistent organic pollutants (POPs). Available at http://chm.pops.int/ [2017-06-15], 2004.

[153] National implementation plan for the Stockholm Convention on persistent organic pollutants for Sweden 2012. Available at http://www.naturvardsverket.se/Om-Naturvardsverket/Publikationer/ISBN/6400/978-91-620-6498-3/ [2017-06-15].

[154] SYSTEQ: "the development, validation and implementation of human systemic Toxic Equivalencies (TEQs) as biomarkers for dioxin-like compounds". Available at http://www.systeqproject.eu/ [2017-06-15], 2009-2013.