Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects

Magnus Bergknut
Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects

Magnus Bergknut, Environmental Chemistry, Department of Chemistry, Umeå University, Umeå, Sweden

Abstract
The risks associated with a soil contaminated by polycyclic aromatic hydrocarbons (PAHs) are generally assessed by measuring individual PAHs in the soil and correlating the obtained amounts to known adverse biological effects of the PAHs. The validity of such a risk estimation is dependent on the presence of additional compounds, the availability of the compounds (including the PAHs), and the methods used to correlate the measured chemical data and biological effects. In the work underlying this thesis the availability, chemical composition and biological effects of PAHs in samples of soils from PAH-contaminated environments were examined. It can be concluded from the results presented in the included papers that the PAHs in the studied soils from industrial sites were not generally physically trapped in soil material, indicating that the availability of the PAHs was not restricted in this sense. However, the bioavailable fraction of the PAHs, as assessed by bioassays with the earthworm *Eisenia Fetida*, could not be assessed by a number of abiotic techniques (including: solid phase micro extraction, SPME; use of semi-permeable membrane devices, SPMDs; leaching with various solvent mixtures, leaching using additives, and sequential leaching) and it seems to be difficult to find a chemical method that can accurately assess the bioavailability of PAHs. Furthermore, it was shown that PAH-polluted samples may be extensively chemically characterized by GC-TOFMS using peak deconvolution, and over 900 components can be resolved in a single run. The chemical characterization also revealed that samples that appeared to be similar in terms of their PAH composition were heterogeneous in terms of their overall composition. Finally, single compounds from this large set of compounds, which correlated with different biological effects, could be identified using the multivariate technique partial least squares projections to latent structures (PLS). This indicates that PLS may provide a valid alternative to Effect Directed Analysis (EDA), an established method for finding single compounds that correlate to the toxicity of environmental samples. Thus, the instrumentation and data evaluation tools used in this thesis are clearly capable of providing a broad chemical characterization as well as linking the obtained chemical data to results from bioassays. However, the link between the chemical analyses and the biological tests could be improved as an organic solvent that solubilised virtually all of the contaminants was used during the chemical analysis while the biological tests were performed in an aqueous solution with limited solubility for a number of compounds. Consequently the compounds probably have a different impact in the biological tests than their relative abundance in profiles obtained by standard chemical analyses suggests. The availability and bioavailability of contaminants in soil also has to be studied further, and such future studies should focus on the molecular interactions between the contaminants and different compartments of the soil. By doing so, detailed knowledge could be obtained which could be applied to a number of different contaminants and soil types. Such studies would generate the data needed for molecular-based modelling of availability and bioavailability, which would be a big step forward compared to current risk assessment practices.

Keywords: polycyclic aromatic hydrocarbons, PAHs, availability, bioavailability, chemical analysis, characterization, GC-TOFMS, bioassay, toxicity, biological testing, multivariate methods, PCA, PLS.

Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects

Magnus Bergknut

Umeå University
Department of Chemistry, Environmental Chemistry
Umeå 2006
Abstract

The risks associated with a soil contaminated by polycyclic aromatic hydrocarbons (PAHs) are generally assessed by measuring individual PAHs in the soil and correlating the obtained amounts to known adverse biological effects of the PAHs. The validity of such a risk estimation is dependent on the presence of additional compounds, the availability of the compounds (including the PAHs), and the methods used to correlate the measured chemical data and biological effects. In the work underlying this thesis the availability, chemical composition and biological effects of PAHs in samples of soils from PAH-contaminated environments were examined. It can be concluded from the results presented in the included papers that the PAHs in the studied soils from industrial sites were not generally physically trapped in soil material, indicating that the availability of the PAHs was not restricted in this sense. However, the bioavailable fraction of the PAHs, as assessed by bioassays with the earthworm Eisenia Fetida, could not be assessed by a number of abiotic techniques (including: solid phase micro extraction, SPME; use of semi-permeable membrane devices, SPMDs; leaching with various solvent mixtures, leaching using additives, and sequential leaching) and it seems to be difficult to find a chemical method that can accurately assess the bioavailability of PAHs. Furthermore, it was shown that PAH-polluted samples may be extensively chemically characterized by GC-TOFMS using peak deconvolution, and over 900 components can be resolved in a single run. The chemical characterization also revealed that samples that appeared to be similar in terms of their PAH composition were heterogeneous in terms of their overall composition. Finally, single compounds from this large set of compounds, which correlated with different biological effects, could be identified using the multivariate technique partial least squares projections to latent structures (PLS). This indicates that PLS may provide a valid alternative to Effect Directed Analysis (EDA), an established method for finding single compounds that correlate to the toxicity of environmental samples.

Thus, the instrumentation and data evaluation tools used in this thesis are clearly capable of providing a broad chemical characterization as well as linking the obtained chemical data to results from bioassays. However, the link between the chemical analyses and the biological tests could be improved as as an organic solvent that solubilised virtually all of the contaminants was used during the chemical analysis while the biological tests were performed in an aqueous solution with limited solubility for a number of compounds. Consequently the compounds probably have a different impact in the biological tests than their relative abundance in profiles obtained by standard chemical analyses suggests. The availability and bioavailability of contaminants in soil also has to be studied further, and such future studies should focus on the molecular interactions between the contaminants and different compartments of the soil. By doing so, detailed knowledge could be obtained which could be applied to a number of different contaminants and soil types. Such studies would generate the data needed for molecular-based modelling of availability and bioavailability, which would be a big step forward compared to current risk assessment practices.

Keywords: polycyclic aromatic hydrocarbons, PAHs, availability, bioavailability, chemical analysis, characterization, GC-TOFMS, bioassay, toxicity, biological testing, multivariate methods, PCA, PLS.
Sammanfattning

Risken av en jord förorenad med polycykliska aromatiska kolväten (PAHer) uppskattas i regel genom att mäta enskilda PAHer i jorden och sedan korrelera de erhållna mängderna med kända skadliga effekter som orsakas av PAHer. Giltigheten av ett sådant förfarande är beroende av förekomsten av ytterligare ämnen, tillgängligheten av dessa ämnen (inklusive PAHerna) och hur korrelationen mellan kemisk och biologisk data erhölls. Denna avhandling har studerat tillgänglighet, kemiska sammansättning och biologiska effekter hos PAH-förorenade jordar. Resultaten visar att PAHer i en av de studerade tomterna inte var fysisk inkapslad i marken och att PAHernas tillgänglighet således inte var begränsad i det hänseendet. Emellertid kunde den biotillgängliga fraktionen av PAHern, bedömd genom mätningar med daggmasken *Eisenia Fetida*, inte uppskattas genom mätningar med ett flertal icke biologiska tekniker (fastfas mikroextraktion (SPME), semipermeabla membrananordningar (SPMDs), lakning med olika lösningsmedelsblandningar, lakning med tillsatser eller sekventiell lakning) och det verkar som det är svårt att hitta en kemisk metod för att kunna uppskatta biotillgängligheten av PAH. En bred kemisk karakterisering av PAH-förorenade prover kunde erhållas genom analys med GC-TOFMS, vilket i kombination med masspektrometris med topsepuration, resulterade i att över 900 toppar kunde separeras i en enda körning. Den kemiska karakterisering visade att prover som hade liknande sammansättning av PAH var olika om man såg till deras totala sammansättning. Slutligen kunde enskilda ämnen som korrelerade med olika biologiska effekter identifieras genom att använda den multivariata tekniken PLS. Dessa resultat visar att PLS kan vara ett gångbart alternativ till effektstyrden analys (EDA), som är en etablerad metod för att hitta enskilda föroreningar med toxiska egenskaper i miljöprov.

Det är därmed uppenbart att den instrumentering och databehandling som använts i avhandlingen kan användas för att erhålla en bred kemisk analys, såväl som för att koppla de erhållna kemiska resultaten till resultat från olika biologiska testsystem. Den kemiska och biologiska analysen verkad dock kunna förbättras eftersom lösningsmedel med god löslighet för PAHer och liknande ämnen användes under de kemiska analyserna medan de biologiska testerna utfördes i en vattenlösning där lösligheten var begränsad för flera av ämnena. Såldes har ämnene sannolikt en annan verkningsgrad i de biologiska systemena än vad som indikeras av de mängder som uppmätts i den kemiska analysen. Tillgängligheten och biotillgängligheten av organiska föroreningar i mark måste också studeras närmare och helst bör dessa studier fokusera på de molekylära interactionerna mellan föroreningen och olika delar av marken. Genom att göra detta kan detaljerad kunskap erhållas som kan användas på flertalet föroreningar och på olika jordar. Sådana studier skulle resultera i det underlag som krävs för molekylär modellering av tillgänglighet och biotillgänglighet, vilket skulle vara ett stort steg framåt jämfört med den existerade riskuppskattningstekniken.

Table of contents

LIST OF ABBREVIATIONS ........................................................................................................II
LIST OF PAPERS .................................................................................................................. III
1. INTRODUCTION ............................................................................................................... 1
2. POLYCYCLIC AROMATIC HYDROCARBONS ................................................................ 5
   2.1 PROPERTIES AND ENVIRONMENTAL FATE OF PAHS .............................................. 6
   2.2 DETERMINATION OF PAHS ....................................................................................... 8
      2.2.1 Extensive characterization using peak deconvolution ........................................... 11
3. BIOAVAILABILITY OF PAHS IN SOIL ..................................................................... 13
   3.1 EQUILIBRIUM PARTITIONING THEORY ..................................................................... 17
   3.2 ABIOTIC TECHNIQUES FOR ASSESSING BIOAVAILABILITY ........................................ 18
4. BIOLOGICAL TEST METHODS ...................................................................................... 21
   4.2 INHIBITION OF RESPIRATION, GROWTH AND REPRODUCTION ............................... 22
   4.1 MUTAGENICITY AND CARCINOGENICITY ................................................................. 24
5. LINKING CHEMICAL AND BIOLOGICAL DATA .................................................. 27
   5.1 EFFECT DIRECTED ANALYSIS ................................................................................. 28
   5.2 MULTIVARIATE METHODS ......................................................................................... 29
6. ASSESSING THE POTENTIAL RISK OF CHEMICALLY COMPLEX ENVIRONMENTAL SAMPLES .................................................. 35
7. CONCLUDING REMARKS AND FUTURE PERSPECTIVES ..................................... 41
8. ACKNOWLEDGMENTS ................................................................................................... 43
9. REFERENCES .................................................................................................................. 45
TACK ..................................................................................................................................... 57
### List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>EC&lt;sub&gt;50&lt;/sub&gt;</td>
<td>Concentration required for 50% of maximum effect</td>
</tr>
<tr>
<td>EDA</td>
<td>Effect Directed Analysis</td>
</tr>
<tr>
<td>GC</td>
<td>Gas chromatography</td>
</tr>
<tr>
<td>HMW</td>
<td>High molecular weight</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>IS</td>
<td>Internal standard</td>
</tr>
<tr>
<td>K&lt;sub&gt;ow&lt;/sub&gt;</td>
<td>Octanol-water partitioning coefficient</td>
</tr>
<tr>
<td>LMW</td>
<td>Low molecular weight</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>NIST</td>
<td>National Institute of Standards and Technology</td>
</tr>
<tr>
<td>PAC</td>
<td>Polycyclic aromatic carbons</td>
</tr>
<tr>
<td>PAH</td>
<td>Polycyclic aromatic hydrocarbons</td>
</tr>
<tr>
<td>PBT</td>
<td>Persistent, bioaccumulative and toxic</td>
</tr>
<tr>
<td>PC</td>
<td>Principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PCB</td>
<td>Polychlorinated biphenyl</td>
</tr>
<tr>
<td>PLE</td>
<td>Pressurized liquid extraction</td>
</tr>
<tr>
<td>PLS</td>
<td>Partial least squares projections to latent structures</td>
</tr>
<tr>
<td>POP</td>
<td>Persistent organic pollutants</td>
</tr>
<tr>
<td>QSAR</td>
<td>Quantitative structure-activity relationship</td>
</tr>
<tr>
<td>RS</td>
<td>Recovery standard</td>
</tr>
<tr>
<td>RT</td>
<td>Retention time</td>
</tr>
<tr>
<td>SPMD</td>
<td>Semi-permeable membrane devices</td>
</tr>
<tr>
<td>SPME</td>
<td>Solid phase micro extraction</td>
</tr>
<tr>
<td>vPvB</td>
<td>Very persistent, very bioaccumulative</td>
</tr>
<tr>
<td>TOFMS</td>
<td>Time of flight mass spectrometry</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet absorption</td>
</tr>
</tbody>
</table>
List of Papers
This thesis is based on the following papers, which are referred to in the text by the corresponding roman numerals.


II. Magnus Bergknut*, Emma Sehlin, Staffan Lundstedt, Patrik L. Andersson, Peter Haglund, Mats Tysklind. "Comparison of techniques for estimating PAH bioavailability: Uptake in *Eisenia fetida*, passive samplers and leaching using various solvents and additives”. Accepted for publication in *Environmental Pollution*.


IV. Magnus Bergknut*, Adam Kucera, Kristina Frech, Erika Andersson, Magnus Engwall, Ulf Rannug, Vladimir Koci, Patrik L. Andersson, Peter Haglund, Mats Tysklind. “Identification of potential toxic compounds in complex extracts of environmental samples using GC-MS and multivariate data analysis”. *Submitted to Environmental Toxicology and Chemistry*.

The published papers are reproduced with the kind permission of the respective journals.
1. Introduction

According to the Swedish Environmental Protection Agency there are over 50,000 contaminated sites in Sweden, in many of which there are polluted soils. Although widespread environmental concern and strict environmental regulations have helped in reducing levels of pollutants in the environment, modern society continues to release pollutants to the environment, many of which are eventually deposited in soil. Furthermore, due to the persistence of many of the environmental pollutants, levels may remain high enough to pose a threat to the environment and human health long after their release. Consequently, contaminated soils constitute a major environmental issue that needs to be resolved. Soils associated with gasworks, coke and tar production sites, fuel processing facilities, and sites where wood preservatives have been heavily used regularly contain elevated levels of polycyclic aromatic hydrocarbons (PAHs). A number of PAHs have been shown to be acutely toxic, mutagenic, and carcinogenic and are hence of great concern with respect to both the environment and human health.

Once deposited on or released in soil, a compound’s ability to be transported to other compartments of the environment (e.g. air, water, and sediments) is controlled by its availability. High availability indicates that the compound may spread and potentially pose a threat to other parts of the environment. The amount of a compound that is taken up by organisms in the soil (thereby potentially being able to cause an adverse effect in the organism or facilitating its transport to other parts of the food web) is governed by the compounds’ bioavailability. Both the availability and bioavailability of a compound is hence important to determine when assessing the potential risk of a contaminated soil. Due to the nature of past and present industrial activities, as well as the pollutants’ ability to be transformed by metabolisation, degradation, and naturally occurring chemical reactions, the number of possible different compounds related to a polluted may be large (well over 1000 compounds). Developing chemical methods that are capable of analyzing all of these compounds is difficult and highly attenuated and selective methods are often required to detect the compounds. The analytical methods ability to discriminate between pollutants may be a prerequisite for the analysis to be successful, but consequently the chemical analysis will seldom cover all of the compounds present in a sample. An alternative to chemical analysis is to use biological test systems, integrative measurements of the compounds that induce responses in the biological test systems used can be obtained. The range of chemicals covered by a biological test system may hence be larger than that covered by the chemical analysis.

The risks posed by a PAH-contaminated soil are generally assessed by measuring or predicting the concentration of individual compounds or classes of compounds in the soil and correlating the obtained amounts to known adverse biological effects of the compounds. This is routinely done by extracting the total amount of the PAHs from a sample and chemically analysing them. The
concentrations obtained by this procedure are then compared to the concentrations of the compounds that are assumed to cause no adverse effects, as estimated by controlled laboratory experiments. However, if the aim is to evaluate the sample as it occurs in the environment, this procedure will lead to some discrepancies:

- The extraction procedure does not consider availability and bioavailability. Hence, the amount of pollutants included in the chemical analysis and the amounts available in the environment may be different.
- The analysis generally targets specific compounds, so the chemical analysis will not include all of the compounds present in the sample.
- The compounds are usually transferred to an organic solvent that can solubilise virtually all of the contaminants during the chemical analysis, while the biological tests are often performed in an aqueous solution. The biological tests and chemical analyses are consequently performed on different extracts.
- In the environment a large number of compounds coexist in what can be best described as a complex mixture of compounds. These compounds may interact in a number of ways, resulting in a combined adverse effect that is difficult to estimate by performing tests on single compounds.

To address the points listed above, studies related to the availability, chemical analysis, biological test systems and methods that can link availability, chemical and biological effect data are needed. In order to assess the risks posed by a contaminated soil as it occurs in the environment, it is crucial to perform these studies in such a way that the information obtained within each study can be correlated to the results of the other studies and, ultimately, to the original sample (Fig. 1).

**Fig. 1.** A schematic diagram illustrating how the risks posed by a contaminated soil are connected to the pollutants’ availability, the domain of compounds covered by the chemical analysis, the methods used to obtain the biological data, and the link between the chemical and biological data.
Chapter 1. Introduction

If studies on availability, chemical analysis and potential biological effects are to work in tandem they have to be designed to do so from the start. This thesis describes the processes that govern the bioavailability of the PAHs, together with some of the existing methods for assessing the bioavailability of PAHs in soil. It also considers methods for the chemical analysis of PAHs, including the incorporation of peak deconvolution in the analysis, in order to achieve broad chemical characterization. It also includes a section highlighting some of the biological test methods that may be used to assess the adverse effects of PAHs. Two methods for linking chemical and biological data are described: (i) Effect Directed Analysis (EDA), an established method for isolating potent chemicals in a mixture of chemicals, and (ii) partial least squares projections to latent structures (PLS), a multivariate method that may be used for correlating chemical and biological data. Lastly, a section is included that outlines how extractions focusing on bioavailability, extensive chemical analysis, and biological test methods may be performed together with the common goal of assessing the potential risks posed by a chemically complex sample as it occurs in the environment.

Data and conclusions in the abovementioned sections concerning availability, chemical analysis and biological test systems are supported, where appropriate, by the papers underlying this thesis. More specifically, Papers I and II focused on availability and bioavailability by considering different modes of extraction and by comparing uptake in earthworms (Eisenia fetida) with amounts estimated by abiotic techniques. In the study presented in Paper III attempts were made to assess and describe the complexity of the studied samples using GC-TOFMS and peak deconvolution. The data obtained in this study were subsequently linked with results obtained from various biological test systems using multivariate methods, highlighting the potential risks of the total load of compounds, as well as those of individual compounds (Paper IV).
Chapter 2. Polycyclic aromatic hydrocarbons

Polycyclic aromatic hydrocarbons (PAHs) comprise a large group of organic contaminants that are formed as a result of incomplete combustion of organic material. PAHs occur naturally in coal and crude oil and are often associated with the combustion of fossil fuels. PAHs that are emitted into the atmosphere mainly adsorb to particles and may then be transported long distances. PAHs are thus ubiquitous environmental pollutants and elevated levels of site-specific PAHs are generally found near emission sources. The largest source of PAHs in Sweden is domestic heating appliances which contributes about 100 tonnes/year to the total emissions of of slightly more than 150 tonnes/year (Bostrom et al., 2002). However, transport and working machinery will dominate the emission of PAHs in cities and contribute about 50 tonnes/year. The amount of PAHs in soils from remote areas and roadside soils ranges from 0.1 – 5 mg/kg (Paper III, Jones et al., 1989; Benfenati et al., 1992), while much higher levels (> 1000 mg/kg) have been found in contaminated soils from gasworks, coke production sites, and wood treatment and preservation facilities (Paper III, Wilson and Jones, 1993; Lundstedt et al., 2003). The amounts of PAHs in soils connected to wood treatment and preservation processes are generally related to the use of creosote, a coal tar distillate, with water repellent and growth-inhibiting qualities.

PAHs are composed of two or more fused benzene rings and contain only carbon and hydrogen atoms. However, alkyl-substituted PAHs, heterocyclic PAHs containing nitrogen, sulfur and oxygen, and oxidation products of PAHs (oxy-PAHs) – including PAH ketones, PAH quinones, and hydroxylated PAHs – are often grouped together with the unsubstituted PAHs and are then referred to as polycyclic aromatic compounds (PACs).
2.1 Properties and environmental fate of PAHs

PAHs comprise a heterogeneous group of compounds with large differences in physico-chemical properties such as molecular weight, vapour pressure, and water solubility (Table 1).

Table 1. Selected properties of the 16 US-EPA PAHs (Mackay et al., 1992).

<table>
<thead>
<tr>
<th>Number of rings</th>
<th>Molecular weight</th>
<th>Aqueous solubility (mg/l)</th>
<th>Vapour press. (Pa)</th>
<th>Log Kow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naphthalene</td>
<td>2</td>
<td>128</td>
<td>31</td>
<td>1.0x10^2</td>
</tr>
<tr>
<td>Acenaphthylene</td>
<td>3</td>
<td>152</td>
<td>16</td>
<td>9.0x10^1</td>
</tr>
<tr>
<td>Acenaphthene</td>
<td>3</td>
<td>154</td>
<td>3.8</td>
<td>3.0x10^1</td>
</tr>
<tr>
<td>Flourene</td>
<td>3</td>
<td>166</td>
<td>1.9</td>
<td>9.0 x10^-2</td>
</tr>
<tr>
<td>Phenanthrene</td>
<td>3</td>
<td>178</td>
<td>1.1</td>
<td>2.0 x10^-2</td>
</tr>
<tr>
<td>Anthracene</td>
<td>3</td>
<td>178</td>
<td>0.045</td>
<td>1.0 x10^-3</td>
</tr>
<tr>
<td>Pyrene</td>
<td>4</td>
<td>202</td>
<td>0.13</td>
<td>6.0 x10^-4</td>
</tr>
<tr>
<td>Fluoranthenne</td>
<td>4</td>
<td>202</td>
<td>0.26</td>
<td>1.2 x10^-3</td>
</tr>
<tr>
<td>Benzo[an]thracene</td>
<td>4</td>
<td>228</td>
<td>0.011</td>
<td>2.8 x10^-5</td>
</tr>
<tr>
<td>Chrysene</td>
<td>4</td>
<td>228</td>
<td>0.006</td>
<td>5.7 x10^-7</td>
</tr>
<tr>
<td>Benzo[bf]fluoranthenne</td>
<td>5</td>
<td>252</td>
<td>0.0015</td>
<td>-</td>
</tr>
<tr>
<td>Benzo[kf]fluoranthenne</td>
<td>5</td>
<td>252</td>
<td>0.0008</td>
<td>5.2 x10^-8</td>
</tr>
<tr>
<td>Benzo[al]pyrene</td>
<td>5</td>
<td>252</td>
<td>0.0038</td>
<td>7.0 x10^-7</td>
</tr>
<tr>
<td>Dibenzo[a,k]anthracene</td>
<td>5</td>
<td>278</td>
<td>0.0006</td>
<td>3.7 x10^-10</td>
</tr>
<tr>
<td>Indeno[1,2,3-cd]pyrene</td>
<td>6</td>
<td>276</td>
<td>0.00019</td>
<td>-</td>
</tr>
<tr>
<td>Benzo[ghi]perylene</td>
<td>6</td>
<td>276</td>
<td>0.00026</td>
<td>1.4 x10^-8</td>
</tr>
</tbody>
</table>

The fused rings can be positioned in a linear (e.g. anthracene), angular (e.g. phenanthrene) or globular (e.g. pyrene) arrangement (Fig. 2) and generally the physico-chemical properties are correlated to the number of rings, while minor differences within each ring-homologue can be attributed to the arrangement of the rings. The physico-chemical properties of the PAHs largely determine their environmental behaviour. Low molecular weight (LMW) PAHs, containing two or three fused rings, are more water soluble and volatile, and hence more available, than high molecular weight (HMW) PAHs containing >3 fused rings, which are primarily associated with particles. The generally higher availability of LMW PAHs, as compared to HMW PAHs, makes them more susceptible to various biological, chemical and photochemical degradation processes. PAH pollution in soil differs from air and water pollution in several ways. In air, most organic pollutants are degraded by reaction with OH-radicals, and thus their half-lives tend to shorter in air than in either water or soil. In addition, air and water may mix, resulting in dilution of the PAHs into larger volumes and transport to other parts of the environment. However, in soil, where the PAHs are generally adsorbed to compartments of the soil, degradation, dilution and transport are limited (Madsen, 2003).
In soil, biological degradation is the main process responsible for removal of PAHs, although photochemical reactions, volatilization and leaching also contribute (Wilson and Jones, 1993; Johnston et al., 1993; Kochany and Maguire, 1994; Wild and Jones, 1995). PAH-degrading bacteria and fungi have different metabolic pathways (Sutherland et al., 1995; Cerniglia, 1997; Kanaly and Harayama, 2000; Bamforth and Singleton, 2005; Johnsen et al., 2005). Bacteria often use the PAHs as carbon and energy sources and after transformation of the PAHs into compounds that are able to enter the metabolic pathways of the bacteria, the resulting end products are carbon dioxide and water (Bamforth and Singleton, 2005). Fungi metabolize the PAHs to more water-soluble compounds, thereby facilitating their excretion. The fungal pathway is similar to those found in humans and other mammals, and is mediated by the cytochrome P-450 enzyme system (Bamforth and Singleton, 2005). Due to their availability and the reactions involved in the metabolization pathways, LMW PAHs are generally more easily degraded than the HMW PAHs (Wilson and Jones, 1993; Allard and Neilson, 1997; Haeseler et al., 1999b; Eriksson et al., 2000).

The adverse biological effects of PAHs include acute toxicity, developmental and reproductive toxicity, mutagenicity and carcinogenicity. However, the main cause for concern regarding PAHs is related to their
carcinogenicity (Delistraty, 1997). A minimum of four fused rings seems to be required for carcinogenic activity, but that does not mean that all PAHs with four rings are carcinogenic. One of the most intensively studied PAHs is benzo[a]pyrene (Fig. 2), since it is one of the most carcinogenic. However, like all unsubstituted PAHs, it requires metabolic activation to obtain direct carcinogenic properties (Pickering, 1999; Pickering, 2000). A number of other PAHs and PACs are also potent carcinogens and some PACs (including N- and O- heterocyclic PAHs) do not require metabolic activation in order to be carcinogenic (Fernandez et al., 1992; Casellas et al., 1995; Delistraty, 1997; Pickering, 2000).

2.2 Determination of PAHs

A simplified diagram outlining the different steps required for the analysis of PAHs is given in Fig. 3. The first step is a representative sampling procedure. Once the sample has been collected, an extraction is performed in which the pollutants are separated from the bulk of the sample and transferred to another medium, often an organic solvent. The following clean-up is needed to separate the pollutants from compounds or materials that could interfere with the subsequent analysis. Sometimes fractionation is included, as an intermediate step between clean-up and analysis, in which different classes of pollutants are separated in order to produce extracts that are more suitable for certain instrumentation or to ease interpretation of the analysis. Finally, instrumental analysis is performed and during this step additional separation and selectivity can often be achieved.

![Fig. 3. Schematic diagram of the analytical procedure for analysing PAHs. Once a sample has been collected it is subjected to extraction, clean-up and fractionation prior to instrumental analysis. Parts of the sample may also be applied to toxicity test systems.](image)

Routine analysis of PACs is often limited to the 16 PAHs regarded as priority pollutants by the US EPA (Table 2). However, hundreds of PACs have been identified in various matrixes, including contaminated soil (Mueller et al.,...
Chapter 2. Polycyclic aromatic hydrocarbons

1989; Meyer et al., 1999; Lundstedt et al., 2003), sediments (Fernandez et al., 1992), diesel exhaust (Choudhury, 1982), airborne particulates (Casellas et al., 1995; Allen et al., 1996; Allen et al., 1997), emissions from wood combustion (Hedberg et al., 2002), and ash from municipal waste incinerators (Akimoto et al., 1997a; Akimoto et al., 1997b). It is difficult to identify and quantify all of the PACs in such complex mixtures and most methods for determining PACs in complex samples are thus based on fractionation of the samples using liquid chromatography or similar strategies. Analysis is often performed using high performance liquid chromatography (HPLC) coupled with fluorescence or UV detection systems, or gas chromatography (GC) systems coupled to mass, C-, S- or N- selective detectors (Hale and Aneiro, 1997).

In the work underlying this thesis, the central analytical procedure was based on (i) extraction by Soxhlet apparatus or pressurised liquid extraction (PLE), (ii) clean-up using open column chromatography or an in-cell PLE technique, and (iii) analysis by gas chromatography- mass spectrometry (GC-MS). The Soxhlet extractor was invented in 1879 by Franz von Soxhlet and was originally designed for the extraction of lipids from solid materials. Soxhlet is a continuous solvent extraction technique in which clean solvent is allowed to flow through the sample by the use of a heater and a cooling system, usually over a period of 18-24 h with toluene as solvent. Its extensive use for the analysis of organic pollutants in diverse environmental samples makes it possible to compare results from different studies, which contributes to the continued use of the technique. PLE was introduced by Richter et al. (1996) and combines elevated temperature and pressure with solvent extraction. Compared to Soxhlet extraction, PLE has the advantages of being less time consuming, less labour intensive, and uses smaller amounts of solvent. Although the same solvents can be used in PLE as in Soxhlet extraction, a mixture of hexane:acetone (1:1 v/v) has proven to be effective for the extraction of PAHs from soil (Lundstedt et al., 2000). Soxhlet extraction was used in the studies described in Papers I and II, which only included a single sample, while PLE was utilised in the studies reported in Papers III and IV since they included over 30 samples.

The open column clean-up step used in the studies described in Papers I and II was developed by Staffan Lundstedt as part of his doctoral studies and allows simultaneous clean-up and fractionation of PAHs and oxy-PAHs (Lundstedt et al., 2003; Andersson et al., 2003). The clean-up is achieved by passing samples through 5 g of 10 % water-deactivated (w/w) silica columns (Ø = 16 mm), and since only PAHs were studied in the investigations reported in these papers, each column was eluted with 15 ml n-hexane/dichloromethane (3:1 v/v). For Papers III and IV a combined extraction and clean-up procedure was used. This was achieved by placing an adsorbent (silica gel) inside the extraction chamber of the PLE which
retained unwanted compounds (Ong et al., 2003). This resulted in a faster clean-up procedure that consumed less solvent.

Instrumental analysis of PAHs was performed using gas chromatography-mass spectrometry (GC-MS), a standard technique for the analysis of PAHs. A GC-system consists of a chromatographic column placed inside an oven. Once the sample extract has been injected into the GC, compounds are propelled through the system by an inert carrier gas. The retention times (RTs) of the various compounds in the column depend on their boiling points and interactions with the stationary phase, hence they can be separated by a suitable temperature gradient. In the MS, the compounds are ionized and fragmented. The mass spectra obtained (i.e. the fragments formed) depend on the structures of the analytes and, hence, are unique for each compound. The mass spectra, in compination with the GC-separation and molecular masses obtained, may thus be used for identification of individual compounds. The ions formed are separated according to their mass to charge (m/z) ratios. In a time of flight mass spectrometry (TOFMS) instrument, the fragments are analysed based on the time it takes for them to travel from the point of fragmentation to the detector, measured in such a way that all fragments reach the detector simultaneously. In a quadropole instrument (qMS) the quadropole acts as a filter that only allows ions of a given m/z to pass to the detector at a given time. Hence, qMS-instruments need to scan the selected m/z range in order to collect all of the fragments and a full mass spectrum is only obtained after a complete scan. GC-qMS was used for the analysis of the 16 US EPA PAHs reported in Papers I and II while both GC-qMS and GC-TOFMS was used for Papers III and IV. The inclusion of GC-TOFMS, coupled with peak deconvolution, allowed a broad chemical analysis that better captured the chemical composition, and hence the chemical complexity, of each sample.

The PAHs were identified and quantified by comparing their retention times and peak areas to those of certified reference standards, while the internal standard technique (Poole and Poole, 1991), i.e. addition of labelled compounds, was used to compensate for losses of target compounds during the different steps of the analysis (e.g. during clean-up, fractionation and evaporation). The basic principle underlying the use of internal standards (IS) is that by adding IS to the sample at the beginning of the analysis and measuring the amount of IS left at the end of the procedure, compensation factors describing how much of the target compounds have been lost during the analysis can be obtained. Hence, a suitable IS should have the same chemical and physical properties as the target compounds. Furthermore, the IS should be added to the sample as early as possible, i.e. prior to or immediately following the extraction and in studies I and II they were added after the extraction step. A recovery standard (RS), added to the sample as the last step before analysis, is used in tandem with an IS. The RS is employed as a quality indicator, since it can be used to calculate the amount of IS lost during the analytical procedure. A recovery of 100 % suggests that no IS was lost and hence indicates that the execution of the analysis was perfect. However, a recovery of 60-120 % is often accepted as reasonable, and this was also the range of recovery
obtained within the studies underlying this thesis. In studies I-II, a range of H2-labelled PAHs that covered the variance in chemical and physical properties of the target PAHs were used as IS. In Paper III no IS was included, since any such compound could have affected the biological test systems that the obtained extracts were intended to be applied to in Paper IV. However, the results from an analysis of a certified reference soil in study III indicated that poor recovery was not an issue in this study.

2.2.1 Extensive characterization using peak deconvolution

The large number of possible compounds present in an environmental sample are challenging for instrumental analyses. A suitable technique must be able to provide a unique response for all of the included compounds with high sensitivity and resolution, accuracy and reliability, across a large dynamic range, in order to allow measurements of a wide range of chemicals present at any possible concentration. GC-MS fulfills most of these criteria, allowing compounds to be identified using their retention times in combination with the obtained mass spectra data. However, as the number of compounds increases it becomes difficult to maintain adequate peak separation. One solution to this is to use peak deconvolution, i.e. to use spectral data to resolve co-eluting chromatographic peaks.

In studies III and IV, gas chromatography-time of flight mass spectrometry (GC-TOFMS) was used since it provides a full mass spectrum for each sampling point (thus eliminating spectral skewing) and high sampling rates (up to 500 spectra/second), which make TOFMS data very suitable for peak deconvolution. Using the peak resolution software supplied with the GC-TOFMS instrument (Leco Chromatof, LECO Corporation, St. Joseph, MI, USA), 962 peaks could be identified in the analysis of a pooled sample and 123 – 527 peaks in the individual samples. In general, lightly-contaminated samples contained fewer compound compared to the highly-contaminated samples while relatively few peaks were found in common between samples with the same sources of contamination. The large differences between the numbers of peaks found in the individual samples, as compared to the number of peaks in common between samples from similar sources, indicated that the samples contained many unique components. Notably, several samples that appeared to be similar in terms of their PAH composition were very heterogeneous in terms of their overall composition.

The method of deconvolution utilized in the Leco GC-TOFMS deconvolution software is proprietary but has similarities with the AMDIS method which was presented in a work by Stein (Stein, 1999). However, it should be noted that the Leco and AMDIS methods are distinctly different in their details, implementation, and function. The AMDIS method involves four sequential steps: (i) noise analysis, (ii) component perception, (iii) spectral deconvolution, and (iv) compound identification. A component is in AMDIS perceived when a sufficient magnitude of its ions maximizes together and is by a number of steps converted
into a model peak profile for the component. The spectrum for each component is then derived from its model peak profile using a least-squares procedure. Compound identification is achieved by comparing the obtained spectra with a reference library or a master sample. The AMDIS approach may fail if peak tops are broad and several maxima are present as this may lead to a component being identified more than once. Furthermore, very broad peaks may not be identified at all and two components that maximize at precisely the same time can not be separated.

A number of matrix-based approaches to peak deconvolution have been presented that, in contrast to the AMDIS approach, are capable to determine the number of components in an overlapping chromatographic peak as well as the spectrum and concentration profile without relying on assumptions regarding peak shape, location or identity. These methods handle the GC-MS measurements as a data table (matrix) composed of retention times (rows) and mass spectra (columns). The matrix ($X_{CS}$) can then be decomposed into spectral (S) and chromatographic (C) profiles (Eq. 1).

$$X_{CS} = CS^T + ECS$$  \hspace{1cm} (Eq. 1)

This decomposition of the data is similar to principal component analysis (PCA, Eq 2) which is explained further in the “Multivariate methods” section of this thesis. However, peak deconvolution algorithms have been developed for their specific application (e.g. orthogonallity is not required and peaks are assumed to be positive) and hence differ from the algorithms used during PCA. There are several different approaches to matrix based peak resolution, but a basic assumption is that by knowing S (i.e. the true spectral profile) it is possible to calculate C (the true chromatographic profile) and vice versa. The different approaches that can be applied for deconvolution are generally divided into direct (also called non-iterative) methods and iterative methods. Direct methods often give good results, but may require manual work and are hence difficult to automate. Examples of direct methods are: Heuristic Evolving Latent Projections (Kvalheim and Liang, 1992), Evolving Factor Analysis (Maeder, 1987), and Orthogonal Projection Resolution (Liang and Kvalheim, 1994). Iterative methods attempt to solve Eq. 1 iteratively, i.e. via stepwise modification of S and C until X is found. Interative methods require little or no manual input and are hence easy to automate. Examples of interative methods are: Alternating Regression (Karjalainen, 1989), Iterative Target Factor Analysis (Gemperline, 1986), and Gentle (Grande and Manne, 2000; Manne and Grande, 2000). Recent studies utalizing a matrix based approach for automatic peak deconvolution and rapid evaluation of GC-TOFMS data has been presented by Jonsson et al. (2004 and 2005).
3. Bioavailability of PAHS in soil

The bioavailable PAHs in soil are envisaged as the fraction that can be taken up by organisms in the soil as governed by the three-way interactions between the PAHs, the matrix and the organism(s) in the matrix (Reid et al., 2000a). The availability and bioavailability of PAHs in soil reflect their potential for transport to other compartments of the environment (e.g. atmosphere, water, and sediments) and are direct indications of the degree to which organisms living in the soil are exposed. Availability and bioavailability are generally not included in risk assessment procedures since the available or bioavailable fractions are difficult to determine, especially if long-term (decades) changes in the environment are considered. Nevertheless, studies on bioavailability are crucial in order to link the amount of PAHs, as determined using the method outlined in Fig. 3, with the actual amounts that are available to cause adverse effects in the environment. Studies on bioavailability could hence result in a revised risk assessment procedure that gives a more detailed understanding of the risks associated with different polluted sites.

Aging is a central term concerning availability and refers to the process of organic compounds in soil becoming less susceptible to degradability, extractability and other related processes in a time-dependent manner. Both the chemical properties of the contaminants and the soil characteristics influence aging, which may include several steps and diverse processes, including covalent bonding, sorption, diffusion, and entrapment (Alexander, 1995; Gevao et al., 2000; Alexander, 2000). Consequently, the use of artificial soil or spiked samples should be avoided when studying bioavailability, since extrapolation of the resulting data to aged contaminated soils is problematic. Instead, it is preferable to use aged, weathered soils (Madsen, 2003). For the same reasons, the samples should not be ground or subjected to other pre-treatments that may substantially change the physical characteristics of the matrix. While aging decreases the amount of pollutants available for degradation and other processes responsible for the removal of PAHs, it consequently also decreases the amount of pollutants available for uptake by organisms, and thus decreases the fraction capable of causing adverse biological effects. Hence, aging is often used to support the hypothesis that the risks associated with organic pollutants in soil may be exaggerated (Alexander, 1995; Alexander, 2000). However, it should be noted that it has not been established if aging is an irreversible process or if long term changes in the environment may cause the pollutants to be re-released into the environment.

It should be noted that the concept of bioavailability lacks a formal definition and there is little agreement on what bioavailability means, how it should be measured, and how it should be calculated. Thus, it is difficult to compare findings by different authors or proposed techniques for assessing the bioavailability of PAHs in soil, since different studies are often based on different concepts of bioavailability. Specific considerations are that uptake rates differ between species and that the timescale considered will affect the amounts of accumulated PAHs. A factor that complicates attempts to define appropriate
timescales is that various fractions may have markedly differing kinetics, including a “rapidly desorbing” or “readily available” fraction (Reid et al., 2000a). Although this term also lack a formal definition it is based on observations which show that desorption of organic pollutants from soils and soil-related materials often follows a biphasic curve, with a rapidly changing part and a plateau (putatively due to kinetic boundaries and sequestration) (Fig. 4).

![Fig. 4](image)

**Fig. 4.** An illustration of how the extracted amounts of a compound in soil first increase with increasing extraction time or extraction efficiency, but reach a plateau after a certain amount has been extracted.

Soil is a very complex matrix, and the PAH sorption/desorption mechanisms associated with it are also complex. A schematic diagram of a soil profile and the processes governing the availability of PAHs in soil is given in Fig. 5. As can be seen in Fig. 5, soil can be viewed from a number of perspectives (2 mm – molecule interactions), each offering differences in detail. However, the interactions between PAHs and constituents of the soil occur at sub-particle and molecular levels, and may include: solubilisation in the aqueous phase (A), adsorption to or into dissolved organic matter (B), adsorption to moist organic surfaces (C), adsorption to moist surfaces, e.g. quartz, (D), adsorption to amorphous or dense OM (E), adsorption to anthropogenic carbon including non-aqueous phase liquids (F), adsorption to soot or similar carbon structures (G), and entrapment in micro pores (H) (Luthy *et al.*, 1997). From the number of possible interactions with soil, and the often large differences between different soils, it can be concluded that studies on availability are intrinsically very complex.
Studies on soil and sediment particles have proved that individual particles may be composed of sub-particle-size regions with differing affinities for PAHs (Gillette et al., 1999). Studies on sediment have found that PAHs tend to be more abundant on surfaces than in the core of particles (30-100 times), that coal and wood sub-particles (5%) in sediment contained 62% of the PAHs found in the
sediment, and that less than 10% of the extracted PAHs came from wood/coal and more than 80% from inorganic sub-particles, further indicating large differences in availability between the constituents of sub-particles (Ghosh et al., 2000). Of the carbon compartments shown in Fig. 5, PAHs bind more strongly to soot and non-amorphous OM (by a factor of 10-100) than amorphous OM (Cornelissen et al., 2005). Of the amorphous OM, PAHs bind more strongly to humic acid than fulvic acid and humin (Perminova et al., 1999; Northcott and Jones, 2000). A number of chemical properties of the PAHs influence the interactions between them and the compartments shown in Fig. 5. These include: their molecular size, volatility, water solubility, lipophilicity, and reactivity. In addition, strong intermolecular forces like \( \pi \)-bonding, hydrogen bonding, ligand exchange reactions, and ionic and dipole-dipole interactions are also important (Gevao et al., 2000).

The availability of PAHs at a contaminated site was studied in Paper I. Analyses of the effects of various pre-treatments (grinding in a mortar, ball-mill, or in acidic conditions, and no grinding) and extraction conditions (extraction with different solvents for various durations: 2-128 h) indicated that the PAHs were associated with the surface of the studied soil and that there were no apparent kinetic boundaries related to the extraction of the PAHs (Paper I). Similar results have been reported in a study focused on characterizing the OM of PAH-polluted soil from former gas plants (Haeseler et al., 1999a).

Intensive efforts have been made to understand soil as a matrix and the interactions between the soil and the pollutants. However, as stated earlier, bioavailability depends on the three-way interactions between the compounds, the matrix and the organism(s) in the matrix and all of these interactions need to be considered when studying bioavailability (Reid et al., 2000a). The uptake of PAHs by organisms from soil is highly species-dependent. Hence the choice of model organism for studies on bioavailability will greatly influence the obtained results. Two common test systems for assessing the bioavailability of PAHs in soil are degradation by microorganisms and uptake in earthworms. Since degradation is not synonymous with accumulation and degradation is more pronounced for two- and three-ring PAHs while PAHs with four or more fused rings are the main causes of concern, it is questionable whether measures of microbial degradation have any relevance as indicators of bioavailability when assessing the potentially adverse effects of PAHs in soil. However, as tools for evaluating different bioremediation scenarios they may still be useful. In contrast, earthworms’ extensive use in soil ecotoxicology analyses, known importance in the terrestrial food chain, high degree of pollutant accumulation, and ease of handling make them suitable for attempts to estimate the potential exposure of biota (Lanno et al., 2004; Jager et al., 2005). The earthworm Eisenia fetida was consequently used as a model organism for assessing bioavailability in study II. However, it should be noted that earthworms are affected by pH and various other soil properties, and they can only be used with limited concentration ranges of some substances, and not at all for other substances. Furthermore, the accumulation of pollutants and sensitivity to soil conditions vary between different earthworm species (Jager et al., 2005).
Other systems for assessing the bioavailability of PAHs in soils could be based on uptake in vegetables or grazing animals. However, plants take up relatively minor amounts of PAHs, and the main source of those they do take up is the atmosphere (Beck et al., 1996; Kipopoulou et al., 1999; Fismes et al., 2002). The amount of soil ingested by grazing farm animals has been estimated to amount to 1-18% of their dry matter intake (Beck et al., 1996). Consequently, by estimating the input from other sources (feed, water, and atmosphere) the input, and hence the bioavailable amount, from soil could potentially be estimated. Although some studies on uptake and mass balance in cows have been performed, none of them has focused on bioavailability (McLachlan, 1993; Beck et al., 1996; Thomas et al., 1999).

3.1 Equilibrium partitioning theory

A widely accepted theory concerning the uptake of chemicals by organisms in soil (or sediments) is the equilibrium partitioning (EP) theory, i.e. the hypothesis that the bioavailability of a compound is controlled by equilibrium partitioning between the soil, water and the organisms (Fig. 6) (Shea, 1988; Ditoro et al., 1991; Sijm et al., 2000).

![Fig. 6. Schematic diagram of equilibrium partitioning (EP) theory. The uptake of an organic compound from soil by an organism is putatively controlled by the compound’s equilibrium partitioning between the soil, the water phase and the organisms.](image)

According to EP, organisms do not take up compounds directly from the soil, but from the freely dissolved fraction in pore water, since, according to the laws of thermodynamics, a chemical will be distributed between the soil, water and the organisms. This implies that the concentration in the organism may be calculated if the partitioning coefficients between the soil and water (sorption coefficient of the chemical) and between the water and organisms (the bioconcentration factor (BCF) of the chemical) are known. Some deviations from expected EP results have been
observed, which are usually explained by sequestration of pollutants in the soil and the effects of feeding and biotransformation (Belfroid et al., 1995). However, it has not been concluded if these processes lead to deviations from EP, or if the discrepancies could be explained by extreme sorption coefficients and BCFs.

Based on experimentally determined levels of PAHs in the water phase and the soil, BCFs derived from EP theory and concentrations of individual PAHs in the earthworm Eisenia fetida were compared to actual values for a number of PAHs in Paper II. Calculations were performed according to the procedure described by van der Wal et al. (2004a) as they gave good correlations between observed and predicted values in a study based on the worm species Eisenia andrei and Aporrectodea caliginosa, covering the compounds hexachlorobenzene, telodrin, dieldrin and seven polychlorinated biphenyls. However, in Paper II poor correlations were obtained, especially for the PAHs that were most abundant in the soil, namely phenanthrene, fluoranthene, and pyrene. The reasons for the poor correlations were not explored further, but it was suggested that (i) more accurate kinetic release data and partitioning coefficients between the soil and water and between the water and earthworms; (ii) measurements of both accumulation and elimination of PAHs in earthworm; and (iii) estimation of the degradation of PAHs by microorganisms during the experiment could improve the obtained results.

3.2 Abiotic techniques for assessing bioavailability

It would be beneficial if the bioavailability of PAHs could be reliably estimated by a relatively cheap and fast chemical method instead of using living organisms. In addition to the ethical arguments for using a chemical method, the experimental results would not be influenced by the activity of living organisms or restricted by the organisms’ tolerance to pH, amount of foodstuff and other soil properties, suitable concentration ranges, and the potential toxicity of some compounds. As stated earlier, the concept of bioavailability lacks a formal definition and there is little agreement on what bioavailability means, how it should be measured, and how it should be calculated. Consequently, many attempts have been made to develop methods that, depending on the assumed definition of bioavailability, can assess the fraction of available compounds in the soil. These attempts can generally be divided into three different classes, namely studies focused on: (i) assessing the readily available fraction of compounds, (ii) the available fraction of compounds as defined by EP theory, or (iii) finding a chemical method that gives 1:1 correlations between the extracted amounts and the amounts that are degraded by microorganisms or accumulated by soil-dwelling organisms.
The readily available fraction is generally assessed by different extraction strategies or by estimating uptake using a high capacity passive sampler. Recent extraction-based studies have applied supercritical fluid extraction (SFE) (Hawthorne and Grabanski, 2000; Hawthorne et al., 2002; Cajthaml and Sasek, 2005; Hawthorne et al., 2005; Nilsson and Bjorklund, 2005) while uptake studies have applied semi-permeable membrane devices (SPMDs), or the resins Tenax® and XAD (Macrae and Hall, 1998; Cornelissen et al., 1998; Kraaij et al., 2002; Lei et al., 2004). However, the results from these studies are difficult to compare since they have been performed on different soils, using different reference systems for bioavailability, and sometimes only incorporated an end-step comparison (i.e. did not evaluate the whole uptake curve). The bioavailable fraction of compounds as defined by EP theory is usually studied using solid phase micro extraction (SPME) samplers, which have low capacities and are hence assumed to be capable of sampling pore water concentrations without disturbing the soil-water partitioning of the compounds. SPMEs have been used in a number of recent studies and been shown to give results that correlate with both the predicted and observed uptake of a number of compounds (Mayer et al., 2000; Parkerton et al., 2000; van der Wal et al., 2004b). Experiments on the amounts of sampled compounds in comparison to the amounts degraded by microorganisms or taken up by soil-dwelling organisms have generally been performed by leaching or extraction using solvents, solvent mixtures, or additives including detergents and complex-forming chemicals. (Volkering et al., 1995; Kelsey et al., 1997; Kelsey and Alexander, 1997; Reid et al., 2000b; Liste and Alexander, 2002). Some of these studies have obtained close to 1:1 correlations between the extracted amounts and the amounts assessed by the biological model system. However, the studies were generally limited to single compounds or spiked soils and the results are hence difficult to extrapolate to real samples.

The diversity of available techniques for assessing the bioavailability of PAHs in soils poses problems when trying to compare the data obtained, since these data can be normalized to a number of different variables, e.g. the amount of soil, additive, or lipophilic material used. Furthermore, the range of PAHs assessed must be considered since different PAHs pose different risks. The carcinogenic and mutagenic PAHs (generally containing four, five or six fused rings) are of particular concern in this respect, since they pose greater risks to humans and the environment than smaller PAHs (Delistraty, 1997). These issues were the focus of Paper II, in which the amounts (total and relative) taken up by the earthworm *Eisenia fetida* were compared to the amounts extracted by a number of abiotic techniques (solid phase micro extraction, SPME; use of semi-permeable membrane devices, SPMDs; leaching with various solvent mixtures; leaching using additives, and sequential leaching) for assessing the bioavailable fractions. Using an aged soil, distinct differences were observed which could be explained by differences in the included techniques’ proposed working principles. In general, the PAH profiles yielded by all of the tested bioavailability assessment techniques contained smaller proportions of carcinogenic PAHs and larger proportions of small non-carcinogenic PAHs than the reference system (*Eisenia fetida*). The cause of the
high ratio of carcinogenic/non-carcinogenic PAHs in the earthworms was not established, but uptake via the gut, elimination, and the earthworms’ promotion of microbial degradation of LMW PAHs seemed to be contributing factors. The results suggest that it may be difficult to develop a chemical method that is capable of mimicking biological uptake, and thus the availability of the PAHs.
4. Biological test methods

The adverse effects of PAHs may be classified into two categories: baseline toxicity, and reactive/specific modes of toxic action (Escher and Hermens, 2002). Baseline toxicity (narcosis) is believed to be the result of the partitioning of pollutants into biological membranes, which leads to disturbances in membrane integrity and functions. Reactive compounds and specifically acting compounds exert their toxic effects through binding to receptors or enzymes, and the toxicity of these compounds is dependent on both their affinity and the nature of their interaction with the target site (Escher and Hermens, 2002).

The effects of some compounds in some test systems may however involve both modes of action. Taking PACs and algae growth test systems as an example, several processes may simultaneously contribute to the observed toxicity. The decrease in growth caused by the PACs will be partly due to baseline toxicity (van Wezel and Opperhuizen, 1995). However, since algae have cytochrome P450-type enzymes, the PAHs may also act via P450-mediated modes of action (Safe, 1993; Delistraty, 1997). In addition, some PACs are reactive and may hence have specific toxicities (Warshawsky et al., 1995). PAH are also photolabile (phototoxic) and the photomodification of PAHs often results in products or metabolites that are more toxic to the algae than the original PAHs (Warshawsky et al., 1995; Mallakin et al., 1999; Mallakin et al., 2000). Furthermore, since the PAHs may be metabolized, the intermediate compounds and the resulting metabolites may also influence the algae (Semple et al., 1999; Mallakin et al., 2000). In addition to all this, some of the observed effects may be caused by factors other than the PAHs, e.g. other pollutants, solvents, pH, and matrix effects. Consequently, if a general test is used to assess the effects of a mixture, the number of confounding factors will be greater than if a selection of extraction, clean-up and fractionation techniques is used in combination with more specific test systems (e.g. systems that target a specific and well-characterized mechanism like mutagenicity).

When assessing the adverse effects of mixtures, like PACs, it is possible that the compounds in the mixture may show additive, antagonistic, or synergistic interactions, or no interactions at all (Carpenter et al., 1998; Groten, 2000; Groten et al., 2001; Escher and Hermens, 2002; Altenburger et al., 2004). The additive (and presumably also antagonistic or synergistic) effects may act through different mechanisms (Carpenter et al., 1998). However, PAHs generally cause additive rather than antagonistic or synergistic effects (Erickson et al., 1999; Fent and Batscher, 2000; Escher and Hermens, 2002). This indicates that the contribution from individual PAHs which are present below a generally accepted no observed effect level may still contribute to the overall effect (Escher and Hermens, 2002; Walter et al., 2002).

When applying PAHs to biological test systems, the solubility and availability of the PAHs must be considered. As discussed earlier, the availability of PAHs is correlated to their chemical properties and hence HMW PAHs
generally have lower availability than LMW PAHs. Consequently HMW PAHs must also be assumed to be less available in any biological test system compared to LMW PAHs. In most in vitro test systems a co-solvent is used to enhance the PAHs’ solubility in order to overcome some of these restrictions. However, it should be noted that the amount of the compounds, as determined by chemical analysis, and the amounts available in the test system may still differ (Gulden and Seibert, 1997; Brown et al., 2001; Gulden et al., 2001). The most commonly used co-solvent is dimethyl sulfoxide (DMSO), which was also used in study IV.

4.2 Inhibition of respiration, growth and reproduction

In study IV several different test systems were used in order to obtain a broad characterization of the possible adverse effects of exposure to PAHs, including dehydrogenase activity (DHA), root growth (Hordeum vulgare), reproduction of springtails (Folsomia candida), algal growth (Desmodesmus subspicatus), germinability (Sinapis alba), and the bacterium Vibrio fischeri (Table 2). The studies also included the DR-CALUX and Ames Salmonella assays, which target specific mechanisms: mutagenicity and the aryl hydrocarbon receptor (AhR) mechanism, respectively (Table 2).

<table>
<thead>
<tr>
<th>Bioassay</th>
<th>Type</th>
<th>Compartment tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>Folsomia candida</td>
<td>Insect</td>
<td>Soil</td>
</tr>
<tr>
<td>(Reproduction)</td>
<td>(Springtail)</td>
<td></td>
</tr>
<tr>
<td>Hordeum vulgare</td>
<td>Plant</td>
<td>Soil</td>
</tr>
<tr>
<td>(Root growth)</td>
<td>(Barley)</td>
<td></td>
</tr>
<tr>
<td>Sinapis alba</td>
<td>Plant</td>
<td>Extract</td>
</tr>
<tr>
<td>(Germinability)</td>
<td>(Mustard)</td>
<td></td>
</tr>
<tr>
<td>Dehydrogenase activity (DHA)</td>
<td>Microorganisms</td>
<td>Soil</td>
</tr>
<tr>
<td>Vibrio fischer</td>
<td>Microorganism</td>
<td>Extract</td>
</tr>
<tr>
<td>(Luminescence)</td>
<td>(Bacteria)</td>
<td></td>
</tr>
<tr>
<td>Desmodesmus subspicatus</td>
<td>Microorganism</td>
<td>Extract</td>
</tr>
<tr>
<td>(Growth)</td>
<td>(Algae)</td>
<td></td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Cell</td>
<td>Extract</td>
</tr>
<tr>
<td>(AhR agonists)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ames</td>
<td>Bacteria</td>
<td>Extract</td>
</tr>
<tr>
<td>(Mutagenicity)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The DHA and *Vibrio fischeri* assays are both microorganism-based, but whole soil samples were used in the DHA assays, and hence any of the microorganisms in the soil could have contributed to the observed responses, while the *Vibrio fischeri* test, which involves use of a luminescent bacterial strain, was performed on soil extracts. The root growth (*Hordeum vulgare*) and germinability (*Sinapis alba*) assays both measure the effects of pollutants on the first stages of ontogenesis, but the root growth assay was performed with soil samples and the germinability assay was performed with extracts from the soils. The reproduction of springtails (*Folsomia candida*), algal growth (*Desmodesmus subspicatus*), Ames *Salmonella*, and DR-CALUX assays were all performed with extracts from the soils.

The results of the DHA and algal growth tests displayed good correlations to the total amount of PAHs, indicating that baseline modes of action were involved. The results of the *Vibrio fischeri* assay, a microorganism-based test system, displayed weaker correlations to the PAHs than those of the DHA and algae growth (which is also a microorganism-based test system) assays. The reasons for these differences were not investigated, but they could be due to differences in availability or mode of action between the different systems. The reproduction of springtails (*Folsomia candida*) assay showed a clear dose-dependent response, which was verified to be due to the amounts of the samples added, using a complementary toxicity test (*Daphnia Magna*, ISO 6341:1996). However, the compounds responsible for the toxicity could not be identified, highlighting the fact that biological test systems often cover a wider range of chemicals than the chemical analysis. The root growth (*Hordeum vulgare*) and germinability (*Sinapis alba*) tests did not appear to be good options for assessing the presence of PAHs or PAH-related compounds, since their responses showed limited dose-dependence and there appeared to be no correlation between the responses and the measured compounds (Fig. 7).

![Fig. 7](image-url)  
**Fig. 7.** Normalized response from the germinability (*Sinapis alba*) and the algal growth (*Desmodesmus subspicatus*) bioassays. The algal response displays a dose-dependent curve while the *Sinapis* response does not.
4.1 Mutagenicity and carcinogenicity

Several PAHs have been shown to bind to DNA and cause mutations, which may result in cancer. The Ames assay, involving the use of *Salmonella thyphimurium*, is one of the most widely used *in vitro* genotoxicity assays (Maron and Ames, 1983). It is simple, quick and sensitive to a wide range of mutagenic and carcinogenic chemicals. The bacterial strain used in the assay has a mutation that prevents histidine biosynthesis. Because of this mutation, the bacteria require exogenous histidine to survive and will starve to death if grown without this nutrient. However, mutagenic compounds may cause a reverse mutation which restores the function of the gene, permitting the cell to grow in the absence of histidine. In order to assess if metabolic activation and transformation is required for mutagenicity, the assay is usually performed in the presence or absence of a metabolizing system. Different systems may be needed for different compounds and in study IV a liver preparation (S9) from Aroclor-treated male Spague-Dawley rats was used. During testing, the bacteria are mixed with the extract containing the target compounds, the metabolizing system, a small amount of histidine, and molten agar. The added histidine will only allow the bacteria to undergo limited numbers of cell divisions. After 48 hours, only those cells that have undergone a reverse mutation at the histidine-synthesizing gene site are still proliferating and producing mutant (revertant) colonies. The number of revertant colonies is compared to those in a control assay (performed in the absence of extract) and the results are reported as revertants/amount of extract.

Dioxin-like compounds exert their toxicity via their effects on the aryl hydrocarbon receptor (AhR) (van den Berg et al., 1998), which initiates transcription of several genes that are part of the biotransformation system, most notably the cytochrome P450 (CYP) genes, resulting in a release of oxidative enzymes and free radicals. Induction of the AhR can be determined in a number of assays. A common method is the 7-ethoxyresorufin-0-deethylase (EROD) assay. In this assay, 7-ethoxyresorufin is deethylated by the cytochrome P450 1A1 (CYP1A) enzyme, resulting in a highly fluorescent substance (resorufin), which can be detected at low concentrations. One drawback of the EROD assay is that a bell-shaped dose-response curve is often obtained (Verhallen et al., 1997). The reason for the response declining instead of levelling out above a certain level is not fully understood, but may lead to an underestimation of the AhR-mediated activity (Petrulis and Bunce, 1999; Petrulis et al., 2001). In study IV the chemical-activated luciferase expression (CALUX) assay was used as an indicator for AhR agonists (Garrison et al., 1996; Murk et al., 1996). In this assay, cells are used that have been transfected with the firefly luciferase reporter gene, resulting in the ability to produce light when the Ah receptor is activated. The relationship between PAHs and the AhR activity is well established and is also, in part the basis for the risk assessment of PAHs (Safe, 1993). However, it should be noted that these assays are mechanism-based (i.e. they only indicate binding to the AhR) and hence do not provide measurements of toxicity *per se*. Over a hundred chemicals, including a variety of naturally occurring compounds, are able to bind to the AhR and many of
these compound are not toxic, but are readily metabolized and excreted (Denison et al., 2002; Baumgart et al., 2005).

The Ames Salmonella and the CALUX assays were used in study IV. It was shown that the 5- and 6-ring PAHs in a set of 63 compounds identified using GC-TOFMS were correlated to the response observed in the Ames Salmonella assay. The correlations between specific compounds and the CALUX response were less clear, although some weak correlation to 5-ring PAHs could be observed. However, it would be interesting to investigate if these correlations could be improved by increasing the availability of compounds with low water solubility, such as 5- and 6-ring PAHs.
Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects.
5. Linking chemical and biological data

From a risk assessment perspective, it is important to consider all of the compounds and biological effects of a complex mixture found in environmental matrixes, such as soils contaminated with PAHs and related compounds. However, linking complex chemical data with results from biological tests is difficult for a number of reasons: (i) the chemical analysis may not include all the pollutants (due either to exclusion or poor limits of detection) involved in the biological response, (ii) different compounds may be present in different samples but still have a similar mode of action, (iii) the biological variation is often large compared to the variance in the chemical data, and (iv) the experimental set-up may not allow all the factors (matrix effects, pH, solvents, availability) contributing to the observed effect to be controlled. These factors may result in little or no correlation between the chemical and biological data or lead to false (positive or negative) correlations.

Two techniques for linking data from chemical instrumental analyses to data obtained from biological test systems will be discussed in this thesis. The first, Effect Directed Analysis (EDA), is based on an analytical procedure designed to isolate the compound or group of compounds responsible for observed effects. The second technique, partial least squares projections to latent structures (PLS), is a mathematical approach aimed at finding correlations between two sets of data, e.g. data from chemical instrumental analyses and data from biological test systems. In analogy with EDA, a PLS-based method can be regarded as a virtual fractionation in which compounds with the strongest correlation to the biological tests are singled out from a large number of compounds.
5.1 Effect Directed Analysis

Effect Directed Analysis (EDA) is an established method for linking toxicants in complex mixtures to biological effects, which has been used to identify active compounds in a number of recent studies (Chen and White, 2004; Claxton et al., 2004; Biselli et al., 2005; Grote et al., 2005a; Grote et al., 2005b) and has been reviewed by Schuetzle and Lewtas (1986) and Brack (2003). EDA is based on a combination of fractionation, biological effect screening, and chemical analysis (Fig. 8).

![Figure 8](image-url)

**Fig. 8.** A graphical representation of Effect Directed Analysis (EDA). The sample is divided into several fractions. The fractions are then tested in one or several bioassays, and the active fractions (grey) are fractionated further. Active fractions are also subjected to chemical analysis in order to identify compounds that contribute to the biological activity.

Repeated fractionation of active fractions, followed by screening and subsequent chemical analysis, allows the isolation and identification of the most potent contaminants. The fractionation is usually related to physico-chemical properties of the compounds including polarity, hydrophobicity, molecular size, planarity or the presence of functional groups. Hence the fractionation procedure itself may provide information on the properties of the fractionated compounds which can later be used as starting points for identification of the isolated compounds. Successful application of EDA relies on (i) obtaining samples close to the source, since dilution (which increases the level of background toxicity) makes identification of specific compounds difficult, (ii) test systems that incorporate specific and non-specific effects, and (iii) sophisticated fractionation procedures with respect to the biological endpoint, kind of sample and expected target compounds (Brack, 2003). Analysis is often performed using GC-MS, although liquid-chromatography (LC-MS) may be required for the identification of polar compounds. Finally, a confirmation step using individual reference standards is required in order to establish a reliable cause-effect relationship (Brack, 2003). Since EDA requires the screening of every sample/fraction by one or several biological test methods, followed by chemical analysis of active fractions, the method is most suitable for the evaluation of small numbers of samples.
5.2 Multivariate methods

Multivariate methods, as applied in chemical analyses (Chemometrics) have been used for a number of different objectives, including: experimental design, pattern recognition/cluster analysis, calibration/predictions, modelling, and curve resolution/peak deconvolution. Multivariate methods are suitable for evaluating large amounts of data and may hence also be used when linking extensive chemical datasets with results from biological test systems. Since the results from a chemical analysis may include data on over 1000 compounds, data evaluation methods that are capable of handling large sets of data and data with more variables than objects are needed. Missing data also need to be handled since it is unlikely that all of the measured compounds will be present in all of the samples. The multivariate methods principal component analysis (PCA) and partial least squares projections to latent structures (PLS) both fulfil these criteria with the added benefit of yielding graphical plots that are easy to interpret.

Fig. 9. Geometrical representation of PCA using three variables (m=3). However, PCA can handle any number of variables. A1: Representation of a single object in a three dimensional space. A2: The principal components (PC1 and PC2) provide the best linear summary of all objects (the complete data matrix). The result of the PCA is displayed using score (B) and loading (C) plots. The score plot shows the relation between objects while the loading plot shows the relation between variables. By superimposing the score plot on the loading plot, it is possible to identify the variables that are responsible for the groupings observed in the score plot.

PCA is a method whereby dominant patterns, i.e. systematic variation, in a data table can be extracted and described by a few parameters (Wold et al., 1987; Höskuldsson, 1995) and is used for evaluating single datasets. In analogy with finding correlations by simple regression analysis between two variables (e.g. correlation between people’s weight and height), PCA involves finding vector lines (principal components) of closest fit to any number of variables. A geometrical representation of PCA, using three variables, is outlined in Fig. 9. For a matrix (data table) with n objects and m variables, each of the m variables can be treated
as an axis in an \( m \) dimensional space. Each object can then be plotted as a point in this space and all the \( n \) objects form a swarm of points in the space (Fig. 9 A1 and A2). PCA will place a straight line (principal component) through the centre of the points in space so that the deviation is as small as possible in the least square sense (i.e. according to the largest variation in the data set). A second principal component, orthogonal to the first, is then calculated according to the remaining largest variation. This may then be repeated and often several principal components are required to describe the dataset satisfactorily. Theoretically PCA corresponds to a mathematical decomposition of the data matrix (\( X \)) into score (\( T \)) and loading (\( P \)) vectors and a residual (\( E \)) according to Eq. 2.

\[
X_{TP} = TP^T + E_{TP} \quad \text{(Eq. 2)}
\]

Score values are identical to the values obtained if each object is orthogonally projected down on the principal component, and the distance between the projected point and the principal component’s centre is measured. Loading values describe the direction coefficients of the principal components, i.e. the influence that each variable has on the direction of each component. The result of the PCA is displayed using score and loading plots (Fig. 9 B and C). The score plot (B) shows the relation between samples; samples with similar variable values will be grouped together while samples with dissimilar values will be located further apart. Sample groupings are thus easily identified. The loadings show the relation between variables and variables that are correlated will be grouped together. By superimposing the score plot on the loading plot, it is possible to identify which variables are responsible for the groupings observed in the score plot.

PLS is similar to PCA, but is used when exploring the correlation between two sets of data (often denoted the descriptor matrix, \( X \), and the response matrix, \( Y \), e.g. chemical and biological data, respectively) (Wold et al., 1984; Geladi and Kowalski, 1986; Wold et al., 2001a; Wold et al., 2001b). This is done by calculating principal components that describe \( X \) and \( Y \), respectively, but at the same time provide a good correlation between the \( X \) and \( Y \) components. PLS has been used for linking physico-chemical properties and chemical data with toxicity in a number of studies, ranging from modelling the toxicity of PCB and dioxins to the identification of groups of compounds that correlate strongly with the responses of a given set of biological test systems (Tysklind et al., 1992; Tysklind et al., 1993; Eriksson et al., 1995; Andren et al., 1998; Eriksson et al., 2002; McDonald et al., 2004).

Multiple samples are required to generate reliable results from PCA and PLS, and around 10-20 samples could be considered as a minimum. Unlike EDA, for which a chemical analysis and biological test has to be performed for each active fraction of a sample, a PLS-based method for linking chemical data with results from biological tests only needs one chemical and one biological effect analysis for each sample. Consequently, large increases in the numbers of samples lead to much smaller increases in the workload of a PLS-based method than in an
EDA-based approach. Another advantage of a PLS-based method is that several endpoints may be included in the same model and, by doing so, the endpoints can be correlated to each other, to the chemicals being evaluated, and to the contents of specific chemicals. However, for optimal performance of PLS, the samples should contain the same pollutants, but at different relative proportions. If the composition is very uniform, no correlations can be established, although if the composition is too diverse, i.e. if the samples contain very different pollutants, it is often better to construct local PLS models using only samples with a similar composition.

One key constraint of both PCA and PLS is that data have to be organized into a matrix with the different samples as rows and the variables (observations) describing the samples as columns. This may be problematic when dealing with chromatographic data. For this kind of data it may be tempting to use the retention times as column indicators and the corresponding intensities as the recorded variables. However, it is essential to match individual peaks correctly and to place them in corresponding columns, despite possible shifts in retention time and composition. A common strategy for doing this is to align the chromatogram from different runs using distinctive peaks, then partition the chromatograms into smaller retention time-windows in which further alignment is performed (Eide et al., 2002; Jonsson et al., 2004). A similar approach was used in study III. During the GC-TOFMS analysis 10 µl from each sample was pooled to obtain a composite reference standard (CRS). The CRS was included in the sample run sequence and PAHs (one each for 3-, 4-, 5- and 6-ring PAHs) in the CRS were used as markers for retention time (R.T.) shifts that may have occurred during the sequence. By applying spectrum matching and compensating for R.T. shifts, the use of a CRS ensured that the peaks in the samples were not miss-assigned, i.e. that chemical A was correctly assigned even if it was peak number 3 at R.T. = X in one sample and peak number 8 at R.T. = Y in another sample.

PCA was used in study III to show that samples could be grouped according to their source of contamination (Fig. 10). Based on 31 samples in which the 16 US EPA PAHs were measured, the resulting score and loading plots showed that samples from contaminated soils (Fig. 10: cluster A) did not form a separate group, while samples from municipal incinerators and samples related to traffic formed well-defined clusters (Fig. 10: clusters B and C, respectively). However, further analysis of the samples by GC-TOFMS revealed that individual samples contained up to 527 peaks and that several samples that appeared to be similar in terms of their PAH composition were very heterogeneous in terms of their overall composition. Many of the compounds that were in common for the soils samples were tentatively identified as methylated PAHs. This was consistent with expectations since unsubstituted PAHs are generally more easily degraded than alkylated PAHs (Douglas et al., 1996).
Fig. 10. Score plot from principal component analysis (PCA) of the relative levels of polycyclic aromatic hydrocarbons (PAHs) in samples from wood preservation sites (closed triangles), gasworks (open triangles), a coke production facility (open squares), traffic (crosses), municipal solid waste incinerators (diamonds), and a certified reference material (CRM) and laboratory blank (closed squares). Clusters A, B, and C represent different sample sources and sample matrices. Clusters A1 and A2 represent soils from wood treatment or preservation sites with high (cluster A1) and intermediate (cluster A2) levels of PAHs. Cluster A3 represents a mixed group containing soils from wood treatment or preservation sites with low levels of PAHs, soils from the gasworks and coke production facility, and an anti-skid sand sample. Clusters B and C represent traffic-contaminated and solid waste municipal incineration samples, respectively.
PLS was used in study IV, and by correlating data from chemical analysis with data from biological tests, single compounds that correlated strongly to the observed effects could be identified within a large dataset. In Paper IV this was highlighted by results from the algal (*Desmodesmus subspicatus*) and Ames *Salmonella* assays. The algal assay yielded similar coefficients of correlation for all of the PAHs, indicating an unspecific response, i.e. any PAH (and possibly also related derivatives) appear to be toxic to the algae while the Ames assay displayed a strong correlation with 5- and 6-ring PAHs (Fig. 11). The results presented in Paper IV, together with results presented in studies in which a similar approach was adopted (Eide et al., 2002) (McDonald et al., 2004) imply that PLS may be a viable alternative to EDA for finding individual compounds that cause the observed adverse effects. However, as with EDA, the effects of the identified compounds still has to be validated using reference standards (Eide et al., 2002).

![Fig. 11.](image)

**Fig. 11.** PLS correlation coefficients for selected US EPA PAHs derived from models between a large set of PACs and the total amount of PAHs (Tot), the unicellular alga (*Desmodesmus subspicatus*) growth inhibition assay, and mutagenicity as measured by the Ames *Salmonella typhimurium* (TA 98 +S9) assay.
6. Assessing the potential risk of chemically complex environmental samples

The current risk assessment procedure includes (van Leeuwen and Hermens, 1995):

**Hazard identification**, consisting of identification of the adverse effects that a substance has an inherent capacity to cause (i.e. effects on behavior and reproduction, birth defects, neurological defects, cancer, and mortality). It also involves establishing the conditions under which these effects may occur and characterization of the behavior of a chemical within a organism.

**Effects assessment**, i.e. estimation of the relationship between dose or level of exposure to a substance, and the incidence and severity of an effect. For most chemicals, laboratory-derived no effects concentrations are converted to predicted no effects concentration (PNEC) for ecological risk assessment by applying uncertainty factors (usually in the range of 10-10 000). The uncertainty factors are numbers reflecting the amount of uncertainty that has to be considered when extrapolating laboratory results to larger populations and a variety of species.

**Exposure assessment** involves determining the emissions, pathways and rates of movement of a substance released into the environment and its transformation or degradation. It also involves describing the nature and size of the population or compartments exposed to a substance, and the magnitude and duration of their exposure. The exposure assessment ultimately results in predicted environmental concentrations (PEC) for a specific compartment (water, sediment, soil, air) in the environment.

**Risk characterization** generally integrates the previous three steps and is often expressed as PEC/PNEC ratios. A PEC/PNEC ratio above one indicates risk and also warrants further actions in order to reduce the risk. The PEC/PNEC ratio should not be viewed as an absolute measurement of risk, but instead allows relative risks to be ranked and, thus, the risks posed by individual compounds or groups of substances to be compared.

After the risk assessment has been completed, risk classification, risk-benefit analysis, risk reduction, and monitoring have to be performed to complete the risk management procedure for a chemical or group of chemicals (van Leeuwen and Hermens, 1995). The European Commission has developed a new system for controlling chemicals, REACH (Registration, Evaluation, Authorization, and restriction of Chemicals). As part of the REACH procedure, companies are required to perform a preliminary risk assessment of chemicals produced at over 1 tonne/year, called a chemical safety assessment, in order to improve the efficiency of the risk assessment process. The chemical safety assessment includes the following steps (Hansson and Rudén, 2004): (i) human health hazard assessment, (ii) human health assessment of physico-chemical properties, (iii) environmental hazard assessment, and (iv) persistent, bioaccumulative and toxic (PBT) and very persistent and very bioaccumulating (vPvB) assessment. If a chemical is determined to be dangerous, or assessed to be a PBT or vPvB, the chemical safety
assessment should also include the following steps: (v) exposure assessment and (vi) risk characterization.

As a complement to the current risk assessment procedure, this thesis proposes the approach outlined in Fig. 12 as a way of assessing the potential risk posed by chemically complex environmental contamination as it occurs in the environment. The following text includes a critical review of the steps included in the analytical procedure for analysing PAHs (namely extraction, clean-up, fractionation, chemical analysis and biological tests) but also, where possible, indicates suggestions for improving the individual steps for assessing bioavailability, PAH contamination, biological tests and associated interactions.

The first step of the process outlined in Fig. 12 is to decide if organic or inorganic pollutants are to be analyzed and here the organic pollutants will be in focus. The second step is to perform an extraction in which the pollutants are separated from the bulk of the sample. However, in contrast to the standard procedure, the extraction step must be highly attuned to the concept of bioavailability. As discussed in the bioavailability section of this thesis, bioavailability lacks a formal definition, and thus it is difficult to suggest exactly how the extraction should be performed. Here an approach including estimations of the readily available pool of pollutants is suggested. This approach could be complemented by estimating the amount taken up by organisms in the soil, i.e. a bioavailability measurement that is related to EP theory. However, the soil concentrations used as input in the EP calculations should be based on estimates of the readily available pool of pollutants, not the total amount of pollutants. As part of the suggested risk assessment procedure the soil matrix is characterized and quantified, again focusing on bioavailability. This characterization should ultimately quantify all of the compartments and interactions displayed in Fig. 5, i.e. solubilisation in the aqueous phase, adsorption onto or into dissolved organic matter, adsorption to moist organic surfaces, adsorption to moist surfaces, e.g. quartz, adsorption into amorphous or dense OM, adsorption to anthropogenic carbon including non-aqueous phase liquids, adsorption to soot or similar carbon structures, and entrapment in micro pores. Obtaining useful descriptors of the soil compartments and soil-organic pollutant interactions may be difficult as soils are generally very heterogeneous. It is therefore suggested that, as a first step, the matrix characterization should focus on compartments and interactions that are assumed to play a major part in limiting the readily available pool of contaminants, e.g. soot (and similar structures) and amorphous OM (Cornelissen et al., 2005). By doing so, a rough estimate of the degree to which the soil would restrict the availability of the pollutants could be obtained.
Chapter 6. Assessing the potential risk of chemically complex environmental samples

QSAR Models capable of predicting bioavailability and (mixture)toxicity based on molecular descriptors.

Metal Extraction (Bioavailability) → Sample → Matrix characterization → Clean-up → Fractionation

Analysis (Hyphenated tech.) → Metals

Identification (Physico-chemical prop.) → Multivariate Analysis

Model describing toxicity as a function of chemical composition

Sample Matrix characterization

Fig. 12. A conceptual representation of an approach for assessing the potential risk posed by chemically complex contamination as it occurs in the environment. The final step of the process is the construction of QSAR models using data generated during the procedure.
After the extraction step in Fig. 12, the sample has to be cleaned-up and fractionated in preparation for the chemical analysis. However, the clean-up may result in the discrimination of toxic compounds, while fractionation may separate compounds that influence each other’s toxic response, e.g. by antagonistic or synergistic effects. Hence minimal clean-up and fractionation should be performed, and if performed at all only mild, non-destructive techniques (for example liquid chromatography) should be used. When possible, all fractions should ultimately also be included in the risk assessment procedure.

Following fractionation, the extracts may be subjected to biological tests (not included in Fig. 12) and chemical analysis. However, in preparation for the chemical analysis (and often during the whole analytical procedure) the compounds are contained in a solvent that may almost totally solubilise the compounds. Most biological tests are, in contrast, performed in an aqueous solution where solubility may be limited for some of the tested compounds. In order to correlate the results from the chemical analysis with the results from the biological tests, as intended in Fig. 12, it is essential for the same extract to be used during the chemical analysis as during the biological tests. Hence the extracts used during the chemical analysis may have to be adjusted, while the biological tests may have to be performed in such a way as to avoid bias with respect to the compounds’ solubility or availability. Ultimately this procedure would require that the target concentration (i.e. the concentration of the compound inside the biological test system) is measured. However, for well defined biological test systems (like the DR-CALUX or the Ames Salmonella assays) the target concentration could probably be predicted for a wide range of compounds based on observed target concentrations for a selection of compounds (Verbruggen et al., 2000).

By using hyphenated techniques during the chemical analysis, the chemical data (e.g. molecular properties such as GC retention times, UV spectra and mass spectra) collected for each sample, as well as individual compounds, are maximized. The use of hyphenated techniques also facilitates the use of peak deconvolution, thereby increasing the possible range of chemicals included in the chemical analysis. This would also, in some cases, reduce the need for fractionation prior to the analysis.

By including multivariate methods in the process outlined in Fig. 12, it would be possible to link the obtained chemical and biological data, and consequently obtain models that are capable of describing toxicity as a function of chemical composition. Furthermore, as shown in this thesis, multivariate methods can also be used to single out the compounds that correlate strongly with the measured biological effects. Once these compounds have been identified, they may be studied further, for example by calculating theoretical molecular properties. These molecular properties may also be complemented by empirical observations, such as GC retention times, UV spectra or mass spectra recorded during the chemical analysis. These properties could be used in the last step shown in Fig. 12, the creation of quantitative structure activity relationship (QSAR) models, i.e.
models that are capable of predicting specific activities (reactivity, toxicity, environmental fate, etc) based on molecular structure. Assuming that the matrix characterization, bioavailability assessment, chemical analysis, and biological test systems are working in tandem, models that are capable of predicting bioavailability or (mixture) toxicity based on molecular descriptors could be constructed. The process outlined is Fig. 12 is, of course, more elaborate and complex than can be displayed in a simple diagram. The extraction and matrix characterization steps have to be investigated further but by limiting the initial studies to the readily available fraction and well defined compartments like soot (and similar structures) preliminary results should be within reach. From paper III and IV it is however evident that the instrumentation and data evaluation tools used in this thesis are capable of providing a broad chemical characterization as well as linking the obtained chemical data to results from bioassays.
Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects
7. Concluding remarks and future perspectives

The studies underlying this thesis investigated the availability, chemical composition, and biological effects of PAH-contaminated samples. From the results in Papers I and II it could be concluded that the PAHs in the studied soils were generally not physically trapped in soil material while the bioavailability of PAHs was governed by a number of processes that are not easily mimicked by chemical methods. Paper III showed that PAH-polluted samples may be extensively chemically characterized by GC-TOFMS using peak deconvolution and over 900 peaks could be resolved in a single run. The chemical characterization also revealed that samples that appeared to be similar in terms of their PAH composition were much more heterogeneous in terms of their overall composition. This may have consequences when evaluating the samples using biological test systems and, hence, shows the importance of a broad chemical analysis. Finally, Paper IV presented an example of how single compounds from a large set of compounds, which correlate with different adverse biological effects, could be identified using the multivariate method PLS. This, together with other studies (McDonald et al., 2004) implies that PLS may provide a valid alternative to EDA. However, the two techniques could very well complement each other, using PLS to screen large numbers of samples and EDA in detailed studies of single samples.

A major objective of the studies underlying this thesis was to evaluate the attributes required for a method that incorporates current knowledge more thoroughly and generate better estimates of the risks associated with environmental contamination than current methods (Fig. 12). A big step towards this objective would be to incorporate understanding from different fields of environmental science in order to decrease the gaps between the bases of chemical and biological effect analyses. One possibility would be to increase the solubility of the compounds in the biological tests. This approach was explored using liposomes in order to enhance the availability of organic pollutants in an in vitro cell-based toxicity test (unpublished data). Liposomes are frequently used for delivering drugs into the body (Lian and Ho, 2001). They have also been used to increase the availability of compounds with poor water solubility during exposure to *Daphnia magna* (Fliedner, 1997). Using a mixture of cell growth medium, liposomes, and a range of organic pollutants (crystals), a clear decrease of the solid phase was observed in comparison with using DMSO. These preliminary observations are promising and warrant further more detailed studies using various biological test systems.

The matrix and the bioavailability of pollutants in soil also have to be further characterised and, ultimately, future studies should focus on the molecular interactions between the PAHs and different compartments of the soil. One of the disadvantages with the methods for assessing the bioavailable fraction included in Paper II is that they do not offer any specific knowledge about the factors affecting the results. Instead, the observations are system-dependent, resulting in the need for a new evaluation for every new soil. However, by focusing on studies that evaluate
molecular based PAH-soil interactions, detailed knowledge could be obtained which could be applied to a number of different pollutants and soils. Such studies could also provide the data needed for molecular based modelling of availability and bioavailability, which would be a big step forward compared to current risk assessment practices.

In comparison to the current risk assessment procedure, the extraction step (bioavailability), chemical analysis and biological testing outlined in Fig. 12 could generate more environmentally relevant PECs and PNECs. However, this comparison is not completely valid, since PNECs are traditionally obtained through tightly controlled experiments on single substances performed in a laboratory setting. In contrast, the adverse effects determined using a method similar to that outlined in Fig. 12 would be performed on real world sample/extracts where a number of factors may influence the observed response. It is hard to tell how assessing the risks associated with an environmental sample using the procedure outlined in this thesis rather than current practice would affect the final judgement. The availability and bioavailability of pollutant are, without doubt, lower than the amounts measured using extensive extraction and hence the risk, as assessed by current practice, is probably exaggerated (Kelsey and Alexander, 1997; Alexander, 2000). However, the complete range of chemicals and the possible toxic interactions of these chemicals are seldom considered. Consequently the risk, in this respect, is probably understated. However, by considering some of the steps included in Fig. 12, knowledge about the potential risks associated with chemically complex environmental samples would certainly increase, and with increasing knowledge, more accurate and relevant methods for assessing the risk of polluted soil could be developed.
8. Acknowledgments

This study was supported by MCN (North Sweden Soil Remediation Center), funded by EU Structural Funds and New Objective 1, contract nos. 133 12534 00 and 304-12738-04. Study I was also carried out as part of the Swedish national research program COLDREM (Soil Remediation in a Cold Climate), funded by MISTRA (the Foundation for Strategic Environmental Research). I have also received research and travel grants from Civilingenjörsförbundets Miljöfond, Stiftelsen J C Kempes Minnes Stipendiefond, and from Wallenbergstiftelsens resebidrag. I am very grateful for all of this financial support.

My thanks go to all of my fellow co-authors of the papers, all of the people involved in obtaining the samples included in the studies, and John Blackwell for an excellent language editing service.

To all past and present work-mates at Environmental Chemistry in Umeå: you are a bunch of wonderful people who all contribute to making Miljökemi a fantastic working place. I wish you all the best of luck in life, both inside and outside the office.

I would finally like to express my deepest gratitude to my supervisors, Mats Tysklind, Peter Haglund, and Patrik Andersson. Without your knowledge and support none of this would have been possible.
Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects
9. References


Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects


Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects


Höskuldsson, A. A combined theory for PCA and PLS. *J Chemometr.* 1995, 9, 91-123.


Kenny, D.V. and Olesik, S.V. Extraction of lignite coal fly ash for polynuclear aromatic hydrocarbons: Modified and unmodified supercritical fluid extraction,


Nilsson, T. and Björklund, E. Selective supercritical fluid extraction as a tool for determining the PCB fraction accessible for uptake by chironomid larve in a limnic sediment. *Chemosphere* 2005, 60, 141-146.


Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects


van der Wal, L., van Gestel, C.A.M., and Hermens, J.L.M. Solid phase microextraction as a tool to predict internal concentrations of soil contaminants in


Tack

Livet består av flera kommunicerande kärl, och en anledning till att kemi- och doktorandstudierna gått så bra är att livet utanför varit så bra. Så ett stort tack till alla ni som finns eller har funnits i min närhet under den här tiden.

Stort tack till alla grabbar och grabbgäng (Anders, Andreas, Erik, Joar, Johan, Kidde, Lars, Matts, Nils, Pontus, Pär, Tomas). Flera av er träffade jag samma sekund som jag satte min fot i Umeå och jag fick genast en plats som en riddare runt det runda? rollspelsbordet. Sedan dess har det fortsatt med dataspel, data, frisbeegolf, innebandy, spriiit och gemensamma luncher. Tack vare er är både denna, och andra, världar betydligt roligare.

Stor tack även till alla andra (men mest till Anders & Linda och J-P & Ulrika) som alltid finns nära, både i stort och smått. Är det inte dags att återigen bestiga de brantaste bergen och snorkla i de djupaste haven?

Åke, Britta, Per, Tomas och alla andra på Sandåsen: Tack för all hjälp, stöd och sällskap.

Knut, Margareta, Matias och Amanda: Tack för att ni finns, utan er hade det aldrig gått. Med föräldrar och syskon som er är inget omöjligt.

Characterization of PAH-contaminated soils focusing on availability, chemical composition and biological effects