Application of attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy to measure sub-lethal effects of potential mutagens.

A thesis submitted for the degree of Doctor of Philosophy in the Faculty of Science and Technology, Lancaster University

Alternative format thesis

November, 2015

Blessing Ebele Obinaju
(Msc)

Lancaster Environment Centre
Abstract

Techniques employed in vibrational spectroscopy monitor the vibrational modes of functional groups within biomolecules and enable a correlation between chemical information and histological structures. Interrogation of biological samples using infrared (IR) techniques generates spectrum with wavenumber-absorbance intensities specific to biomolecules within the sample. Methods are relatively non-destructive, and so samples can subsequently be analyzed by more conventional approaches. Analyses can be carried out ex vivo or in situ in living tissue, where a reference range of a designated normal state can be derived, and anything lying outside this range is potentially atypical. Computational approaches allow one to minimize within-category confounding factors. The application of vibrational spectroscopy in contaminant biomonitoring is a welcome development which has enabled the investigation of real-time contaminant exposure effects in the tissues of sentinels. IR techniques such as attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy, was able to detect changes in various tissue samples exposed to varying levels of polycyclic aromatic hydrocarbons (PAHs). This technique discriminated between spatial and temporal variations in the interrogated tissues. Multivariate analysis was able to relate the alterations at various regions of the fingerprint, to PAH exposure and was able to detect PAH exposure in tissues from sites with no documented knowledge of contamination. ATR-FTIR detected PAH-induced changes in isolated nuclei of cultured cell populations in G₀/G₁ and S-phases of the cell cycle. Findings from the various projects affirm, that techniques involved in IR spectroscopy are highly sensitive to minimal changes in cell molecules. The ability to generate rapid results in real-time is valuable and the wide variety of sample types which can be interrogated using IR techniques makes it a suitable technique for environment biomonitoring.
Acknowledgements

I am extremely grateful to God for grace, health and sustenance during the entire duration of the Ph.D. program. I am thankful to have had the support of my immediate family, friends, and work colleagues both at Lancaster University and the University of Uyo, as well as every individual who has contributed in words and by acts of encouragement. I am grateful and honoured to have been a part of a group of inspiring researchers; Rebecca Strong, Junyi Li, Holly Butler, Kelly Heys, Simon Forgarty, Alana Mitchell, Georgios Theophilou and others who I had the pleasure of interacting briefly with including Dr. Valon Llabjani, Dr. Imran Patel, Dr. Abdullah Ahmadzai, Dr. Matt Riding and Dr. Julio Trevisan. Most importantly, I am extremely grateful to my supervisor Prof. Francis Martin for supporting, guiding and encouraging me throughout the entire Ph.D. I certainly would not have made it without you all.

Declaration

I declare that this thesis is my work and has not been submitted for the award of a higher degree or qualification at this university or elsewhere
Contents

Title page I

Abstract II

Acknowledgements III

Declaration III

Contents IV

Abbreviations V

Chapter 1. General Introduction

Chapter 2. Novel biospectroscopy sensor technologies towards environmental health monitoring in urban environments.

Blessing E. Obinaju and Francis L. Martin

Environmental Pollution 183 (2013) 46-53.

Chapter 3. Distinguishing nuclei-specific benzo[a]pyrene-induced effects in MCF-7 cells from whole-cell alterations using Fourier-transform infrared spectroscopy

Blessing E. Obinaju, Nigel J. Fullwood and Francis L. Martin

Toxicology 335 (2015) 27–34

Blessing E. Obinaju, Alozie Alaoma and Francis L. Martin

Environmental Pollution 192 (2014) 222-231.


Blessing E. Obinaju, Carola Graf, Crispin Halsall and Francis L. Martin

Environmental Pollution 201 (2015) 42-49

Chapter 6. Attenuated total reflection Fourier-transform infrared spectroscopy reveals polycyclic aromatic hydrocarbon contamination despite relatively pristine characteristics of site: results of a field study in the Niger Delta

Blessing E. Obinaju and Francis L. Martin

Submitted Manuscript

Chapter 7. Discussion

Appendix I Using Fourier transform IR spectroscopy to analyze biological materials


Appendix II

*In vitro* protective effects of quercetin in MCF-7 cells despite an underlying toxicity profile

**Blessing E. Obinaju** and Francis L. Martin

Abstract: 35th Annual Meeting of the United Kingdom Environmental Mutagen Society.

*Mutagenesis* 27 (2012) 789-816

Appendix III

Attenuated total reflection Fourier-transform infrared spectroscopy detects real-time polyaromatic hydrocarbon toxicity in fish tissues.

**Blessing E. Obinaju** and Francis L. Martin

Abstract: 36th Annual Meeting of the United Kingdom Environmental Mutagen Society.

*Mutagenesis* 29 (2014) 79-96

Appendix IV

Alterations in infrared spectral signature of *Heterobrachus bidorsalis* reflects polyaromatic hydrocarbon concentrations in Ovia river, Nigeria.

**Blessing E. Obinaju** and Francis L. Martin

Abstract: 43rd Annual Meeting of the European Environmental Mutagen Society.

*Mutagenesis* 29 (2014) 497-559

Bibliography
List of abbreviations

1-OHP: 1-hydroxypyrene

8-oxoGua: 8-oxo-7, 8-dihydroguanine

8-oxoGuo: 8-oxo-7, 8-dihydro-2´ deoxyguanosine

As: Arsenic

ATR: Attenuated total reflectance

B[a]P: Benzo[a]pyrene

Cd: Cadmium

Cu: Copper

Cr: Chromium

DDT: 1,1-dichloro-2,2-bis[4-chlorophenyl]-ethane,

DNA: Deoxyribonucleic acid

EC: Emerging contaminant

EMSC: Extended Multiplicative Signal Correction

FPA: Focal plane array

FTIR: Fourier transform infrared

Ge: Germanium

HCB: Hexachlorobenzene

HCH: Hexachlorocyclohexane,
Hg: Mercury

IR: Infrared

IRE: Internal reflection element

LC₅₀: The concentration of a substance lethal to 50% of the organisms in a toxicity test.

LD₅₀: The individual dose required to kill 50 percent of a population of test animals

LDA: Linear Discriminant analysis

Low-E: Low-emissivity

MCF-7: Human mammary carcinoma cell line

Ni: Nickel

PAH: Polycyclic aromatic hydrocarbon

PC: principal components

PCA: Principal component analysis

PCB: Polychlorinated biphenyl

PCDD/Fs: Polychlorinated dibenzo-p-dioxins and furans

PLS: Partial least squares

PLSR: Partial least squares regression

PM: Particulate matter

RNA: Ribonucleic acid
ROS: Reactive Oxygen species

Se: Selenium

SMR: Standard metabolic rate

SNR: Signal to noise ratio

SOCs: semi-volatile contaminants

SRs: Steroid receptors

UV: Ultra violet

ZnSe: Zinc Selenide
Chapter 1. General Introduction

1.1 Introduction

1.2 Environmental pollution and contamination of ecosystems

1.2.1 Aquatic ecosystems

1.2.2 Atmospheric ecosystems

1.2.3 Terrestrial ecosystems

1.3 Environmental pollution in Africa

1.4 Polycyclic aromatic hydrocarbons

1.4.1 Sources and composition

1.4.2 Environmental and biological fate

1.4.3 Exposure, Biomarkers and Toxicity

1.5 Vibrational Spectroscopy

1.5.1 Fourier Transform Infrared (FTIR) Spectroscopy

1.5.1.1 Background

1.5.1.2 Instrumentation

1.6 Data Handling

1.6.1 Spectra Pre-processing

1.6.2 Feature Extraction, Construction and Selection

1.7 Computational Analysis

1.7.1 Principal Component Analysis

1.7.2 Linear Discriminant Analysis
1.1 Introduction

Environmental contaminants remain a high-risk hazard to human and wildlife populations. Estuaries and other coastal waters within world regions receive a large amount of synthetic chemicals in commercial use. These chemicals enter the various water bodies via a variety of sources i.e. spills, dumping operations, urban runoffs, municipal and industrial waste discharges. These sources contribute materials such as petroleum hydrocarbons, polychlorinated biphenyls (PCBs), other such compounds, pesticides and metals to the water bodies (Malins, 1980; Malins et al., 1984). Most contaminants released into the environment are assimilated by biota with a variety inducing adverse effects at high concentrations at the target organs.

Biomonitoring has facilitated the evaluation/risk assessment of several chemical contaminants within the environment using biological organisms (and their biodiversity) as indicators of environmental health. This is particularly useful because these organisms when studied are able to reveal evidence of toxicity ranging from cellular alterations to endocrine disruption as a result of exposure to chemical contaminants (LeBlanc and Bain, 1997; Paoletti, 1999). Similarly, toxicological studies have aided the identification and characterization of the various contaminants based on the chemical dose (LD₅₀, LC₅₀) and biological response (carcinogens, mutagens and genotoxins). These classifications are largely based on in vitro studies which however accurate, are more often performed using high concentrations which neither reflect concentration levels within the environment, nor the response of biological systems to lower or multiple environmental exposures (Martin, 2007).
Biospectroscopy is a sensitive technique which generates a signature (fingerprint) spectrum of cellular constituents present within cells or tissues, using the ability of molecular bonds within each constituent to absorb and vibrate at specific frequencies of the electromagnetic spectrum (Martin et al., 2010). Spectroscopic methods optimised for use with most biological assays have provided more sensitive, less expensive and less time consuming approaches with the possibilities of monitoring cellular alterations as well as tracking these changes over time (Trevisan et al., 2010).

Environmental contamination is a growing concern for most African and developing countries. Within the Niger Delta region of Nigeria, the various activities possibly responsible for the anthropogenic materials deposited within water bodies are quite prevalent. This suggests that human and wildlife populations are exposed, perhaps to varying concentrations of various environmental contaminants. Impacts and consequences on the resident populations remain largely unknown due to a scarcity of studies providing region-specific data (Essien et al., 2011).

Using vibrational spectroscopy, this thesis identifies polycyclic aromatic hydrocarbon (PAH) induced alterations in sentinels *Heterobranchus bidorsalis* and *Ipomoea aquatic* obtained from the Niger Delta region of Nigeria.

### 1.2 Environmental pollution and contamination of ecosystems

Environmental awareness has increased global concerns regarding the effects of pollution and contaminants. Studies are increasingly designed to assess/understand the ecological effects of contaminants *i.e.* direct effects on organisms, populations and communities including the ecological state of endangered species and their habitats (Chapman, 2004). Declining biodiversity populations and the link to environmental
population is equally an important concern. However, declines in biodiversity population could be difficult to detect as various species occupy same habitats and are thus equally vulnerable to the various forms of environmental alterations such as habitat degradation, deforestation, draining of wetland and xenobiotic contamination (Gibbons et al., 2000). Furthermore, the contamination of one component often leads to contamination of one or all other components of the global ecosystem.

1.2.1 Aquatic ecosystems

The aquatic environment provides a sink for several contaminants with varying toxic potentials (Kelly and Giulio, 2000). Contamination of most aquatic bodies occurs as a result of the contamination of other components (air/land) of the environment. Run-off from contaminated land surfaces and atmospheric depositions are likely avenues by which most aquatic environments receive chemical contaminants (Bayen, 2012). Contaminants such as petroleum hydrocarbons, heavy metals, pesticides as well as several other chemical compounds, can cause direct toxic effects when released into aquatic environments and sensitive species may be impaired by sub-lethal effects or decimated by lethality (Fleeger et al., 2003). Synthetic organic compounds mostly produced for industrial, domestic, or agricultural use, unless specifically removed by wastewater treatment processes, may persist as part of the effluent and be released into receiving waters as trace pollutants. Receiving waters for trace pollutants may constitute a direct source of drinking water or indirectly reach a water supply as recharge water (Murray et al., 2010). Advanced technologies involving the use of granular activated carbon, membrane technology, ozonation, and ultraviolet radiation have been used with relative success to remove pharmaceutical and environmental contaminants from water destined for human consumption (Dorne et al., 2007).
Synthetic compounds classed as environmental contaminants (Murray et al., 2010) have been detected in aquatic ecosystems in various parts of the world (Antizar-Ladislao, 2008; Oehme, 1991). In particular, pesticides (carbamates, chloroacetanilides, chlorophenoxy acids, organochlorines, organophosphates, pyrethroids, and triazines) mainly organochlorines and organophosphates have the potential to induce adverse effects in biota and humans (Leong et al., 2007; Murray et al., 2010). A wide range of pesticides including HCH: hexachlorocyclohexane, DDT: 1,1-dichloro-2,2-bis[4-chlorophenyl]-ethane, HCB: hexachlorobenzene and lindane, have been detected in precipitation, fresh and marine water within Europe (Dubus et al., 2000; Graymore et al., 2001; Loos et al., 2009). Most environmental contaminants are detected in low concentrations, normally in nanogram (ng) or microgram per litre (µgL\(^{-1}\)) range and are referred to as emerging contaminants (ECs) due to their potential to adversely alter human/environmental health (Murray et al., 2010).

1.2.2 Atmospheric ecosystems

Industrialization in the various regions of the world has been greatly associated with the emission of various substances which constitute atmospheric contaminants and increase air pollution. Depending on emission sources, these contaminants may contain complex mixtures of chemical and/or biological components, including viable or non-viable microorganisms and fragments of microorganisms which could include toxic components such as endotoxin and mycotoxins (Gangamma, 2012). Atmospheric contaminants \textit{i.e.} particulate matter (PM) and organochlorines have considerable potential to persist in the atmosphere and be transported over long distances. Long range transport is a term widely used to describe the transport of contaminants over a few to many thousands of kilometres and as such, long range atmospheric transport contributes to the dispersion of contaminants over several
distances, along with other forms of long-range transport via sea currents, biota and ice transport in polar regions (Dubus et al., 2000). Other expressions such as ‘regional scale’ and ‘meso-scale’ have also been used to describe the movement of contaminants (Glotfelty et al., 1990).

Atmospheric contaminants are widely associated with several forms of adverse health conditions in humans (Bell and Holloway, 2007). PM fractions of air pollution increase reactive oxygen species (ROS) generation (Knaapen et al., 2004) and other indirect effects mediated by pulmonary oxidative stress and inflammatory responses (Brook et al., 2004).

1.2.3 Terrestrial ecosystems

Within terrestrial ecosystems, pasture vegetation and soil/sediments are the main concentrated sources of most contaminants. Atmospherically deposited semi-volatile contaminants (SOCs) such as polycyclic aromatic hydrocarbons (PAHs), PCBs, polychlorinated dibenzo-p-dioxins and furans (PCDD/Fs) are primarily intercepted by pasture vegetation and the contaminated vegetation becomes the entry point of these chemicals into the agricultural food chain (Smith et al., 2006). Other contaminants include elemental compounds such as copper (Cu), nickel (Ni), arsenic (As), mercury (Hg), selenium (Se), cadmium (Cd) and chromium (Cr). These compounds as well as other such trace and heavy metals arise from anthropogenic activities such as mining, smelting of metals, manufacturing operations and the use of soil fertilizers in agriculture (Llabjani et al., 2014). Exposure to trace and heavy metal contamination has been shown to be endocrine disrupting and capable of altering gene regulation via the closely related glucocorticoid, mineralocorticoid, progesterone, and androgen steroid receptors (SRs) at concentrations as low as 0.01μM (Davey et al., 2008). Exposures have also been shown to significantly decrease acetylcholinesterase
activity in zebrafish (Danio rerio, Cyprinidae) (Richetti et al., 2011), increase ROS
generation in various systems (Jomova and Valko, 2011; Pinto et al., 2003) and alter
standard metabolic rates (SMR) (Rowe et al., 2001).

1.3 Environmental pollution in Africa

Environmental pollution in Africa and developing countries remains a global concern,
particularly due to a scarcity of robust studies providing data regarding exposure
levels and risk estimates for such regions and its resident population (Bruce et al.,
2000), compared to Europe and other developed countries (Beelen et al., 2014; Brauer
et al., 2008; Brunekreef and Holgate, 2002; Dimakopoulou et al., 2014; Dockery and
Pope, 1994; Dockery et al., 1993; Fleisch et al., 2014; Schwartz, 1994). Although
many studies document evidence of environmental contamination within various
regions of Africa (Akanni, 2010; Doherty et al., 2010; Essien et al., 2012; Essien et
al., 2008; Gwaski et al., 2013; Manirakiza et al., 2002), very few epidemiology or
biomonitoring studies have been conducted to assess the risks associated with
exposures, and most of these studies are focused on the contamination levels in soil
and sediments as opposed to water which constitutes a ready exposure route to human
and animal species.

1.4 Polycyclic aromatic hydrocarbons

1.4.1 Sources and composition

PAHs are semi-volatile organic chemicals and tend to persist in the environment
(Moeckel et al., 2013). PAHs are ubiquitous environmental pollutants which occur
more commonly as complex mixtures and derived from three sources: fossil fuel
(petrogenic PAH), combustion of organic matter (pyrogenic PAH) and the
transformation of natural organic precursors in the environment by relatively rapid chemical/biological processes (biogenic PAH) (Lima et al., 2007; Neff et al., 2005). Commonly measured PAH compounds contain two or more fused benzene (aromatic) rings depending on their sources and represent possibly the largest class of environmental carcinogens (Van Metre et al., 2000). PAH input into the various ecosystems originate from a variety of natural (volcanic eruptions, oil seeps and forest fires) and anthropogenic (vehicular emissions, fossil fuel and wood combustion) sources. They are of toxicological importance due to their mutagenicity and carcinogenic potential (Fent and Bätscher, 2000; Malins et al., 1997).

1.4.2 Environmental and biological fate

PAH deposition in various ecosystems, their sorption to aerosol organic matter and atmospheric long range transport influenced by gas-particle partitioning, are selected reasons for their persistence within the environment (Ma et al., 2013). PAHs emitted into the atmosphere in exhaust gases or by volatization are usually transported over long distances in association with soot particles or PM (Jager et al., 2000; Moeckel et al., 2013; Ruchirawat et al., 2002). PAHs are particularly persistent (under anaerobic conditions) within terrestrial ecosystems e.g. soil, and constitute a potential threat to soil organisms (Jager et al., 2000).

Within contaminated ecosystems, PAHs as nonpolar organic chemicals have low aqueous solubility and high affinity for adsorption to sediments, organic particles as well as absorption by biological organisms (Neff et al., 2005). Organisms may uptake PAHs via dermal or dietary routes of exposure (Fig. 1) and these chemicals tend to accumulate in fatty tissues of aquatic/terrestrial organisms (Mackay et al., 2006; Watanabe et al., 2005). PAHs can be bioconcentrated or bioaccumulated in
lower trophic levels but are rapidly metabolised by organisms e.g. vertebrates, in higher levels.

PAHs are deposited in aquatic ecosystems via a number of sources including spills, urban runoff as well as municipal and industrial waste discharge (Lima et al., 2007; Malins et al., 1984). Generally, contaminant concentrations including PAHs in aquatic organisms vary over time and space because of the influences of natural processes and human activities. For example, within one species such as lake trout, contaminants that bioaccumulate and biomagnify (e.g., methyl Hg and persistent organic pollutants [POPs]) can be 5- to 10-fold higher in one lake than in another neighbouring system because of inherent differences in the species' ecology, the systems' characteristics, or the activities occurring in the watershed (Clements et al., 2012). PAHs released into aquatic systems rapidly become associated with sediments where they may become buried and persist until degraded, resuspended, bioaccumulated or removed by dredging (Cerniglia, 1993). The possible fates of PAHs in the environment include volatization, photooxidation, chemical oxidation, bioaccumulation, adsorption to soil particles, leaching and microbial degradation (Haritash and Kaushik, 2009; Wild and Jones, 1995).

Biodegradation of PAHs has been observed in soils and aquifers (Johnsen and Karlson, 2007), and has been demonstrated to occur under both oxic and anoxic conditions (Bamforth and Singleton, 2005). Studies (Meckenstock et al., 2000; Zhang et al., 2000) have proposed a mechanism for the anaerobic degradation of PAHs (naphthalene) which includes the carboxylation of the aromatic ring to 2-naphthoic acid, activating the aromatic ring prior to hydrolysis. Stepwise reduction of 2-naphthoic acid via a series of hydrogenation reactions results in decaclin-2-carboxylic acid which is subsequently converted to decahydro-2-naphthoic acid. Others include
the bacterial and fungal metabolism of PAHs (Bamforth and Singleton, 2005). There is documentation of a large diversity of bacteria species that are able to oxidise PAHs using dioxygenase enzymes, including organisms from the genus Pseudomonas and Rhodococcus (Cerniglia, 1992; J. G. Mueller et al., 1996). Overall, the low molecular weight PAHs are more volatile, water soluble and less lipophilic than their high molecular weight relatives (Wild and Jones, 1995).

**Fig. 1:** Hypothetical aquatic food web that could occur in contaminated soil/water bodies. Arrows show how nutrient and energy (including contaminants) are transferred from one organism to another in feeding relationships. A more complex food web would include more organisms than shown e.g. decomposers such as bacteria which break down dead organic matter and recycle nutrients within the ecosystem.
1.4.3 Bioavailability versus Bioaccessibility

Bioavailability and bioaccessibility are complex issues that determine whether or not adverse effects are to be expected when organisms or plants are exposed to contaminants. Therefore, the determinants of- and the definitions of bioavailability and bioaccessibility must be understood if one is to monitor or, ultimately predict the effects of potential environmental mutagens (Peijnenburg and Jager, 2003).

Bioavailability is a widely used term and has its origins in pharmacology where it is defined as the fraction of an administered dose of unchanged drug (i.e., parent compound) that reaches the systemic circulation and is one of the principal pharmacokinetic properties of drugs. In ecotoxicology, bioavailability has been defined as the amount of chemical that is actually taken up from the environment and is available to cause a biological response where uptake may include binding to or diffusion through cell membranes including bioaccumulation (McLaughlin and Lanno, 2014).

Within ecotoxicology particularly when applied to contaminants in soil and aquatic systems, it is reckoned that the term bioavailability is used inconsistently and imprecisely (Semple et al., 2004). Thus, Semple et. al., (2004) proposed the following definition- “the concentration of the compound which is freely available to cross an organism’s cellular membrane from the organism’s environment at a given time”. McLaughlin and Lanno, (2014) reckons the fundamental requirement in the definition of bioavailability is that it should be measured in a biological receptor and not simply by chemical or physical analysis of the media, although specific measures can be made which can be correlated to the biological fraction.
Bioaccessibility is a more narrowly used term having its origins in soil science (but rarely used in aquatic toxicology), and is a surrogate measure of bioavailability (McLaughlin and Lanno, 2014). Again, Semple et. al., (2004) define bioaccessibility as “the concentration of the compound which is available to cross an organism’s cellular membrane from the environment, if the organism has access to the compound”. This definition incorporates a time and space dimension where chemicals that are spatially or temporally unavailable to the organism are bioaccessible, but not bioavailable. Thus according to this definition, the magnitude of the bioaccessible pool of the chemical is never less than the magnitude of the bioavailable pool and could also be termed the “potentially bioavailable” pool given no spatial or temporal constraints.

Monitoring bioavailable and bioaccessible fractions in itself may not be considered an endpoint of assessment; rather, the focus is often on assessing adverse risks at species, community, population, or ecosystem levels where bioavailable and bioaccessible fractions can be related to adverse effects in organisms on the basis of the critical body burden concept. The critical internal body burden being the threshold concentration in the body above which physiological functions are irreversibly impaired (Peijnenburg and Jager, 2003). However, properly distinguishing both terms is highly important as it forces practitioners to consider what they actually measure using biological and chemical assays (Semple et al., 2004).
1.4.4 Exposure, Biomarkers and Toxicity

Human and wildlife exposure to PAHs is unavoidable since most PAH compounds occur ubiquitously in the environment and at various workplaces. The earliest examples of occupational cancer among chimney sweeps, workers exposed to coal-tar products, workers in iron foundries, coke ovens and aluminium production plants, is generally agreed to be the result of exposure to PAHs (Phillips, 1999). However, for non-occupationally exposed individuals, diet, ambient air, tobacco smoke and coal-tar-containing medications are the main sources of PAH exposure (Scherer et al., 2000). Mainstream smoke yields of B[a]P for filter cigarettes amount to about 10 ng/cigarette, leading to an intake of about 200 ng/day for a pack-a-day cigarette smoker (Grimmer et al., 1987; Hoffmann, 1997). In the diet, the estimated daily intake of B[a]P ranges from 120-2800 ng/day, with intake from ambient air by inhalation and from water predicted to contribute about 2% and 1% respectively, to the total daily intake in non-smokers (Hattemer-Frey and Travis, 1991).

Biomarkers are physiological or biological responses including variation in cellular or biochemical components, processes, structure or functions measurable in biological systems and/or samples as response to contaminant exposure (Obinaju and Martin, 2013). Biomarkers are useful tools for assessing exposure and induced toxic effects in human and wildlife populations (Table 1). They constitute a biological response that is dose-dependent with toxicant exposure and can be used to monitor exposure and/or effects, with the possibility to intrinsically link toxic compounds to the mechanism by which they cause effects. These responses are most often observed in biological organisms, including their biodiversity (LeBlanc and Bain, 1997; Valavanidis et al., 2006). Biomarkers indicating exposure to PAHs especially in humans include 8-oxo-7,8-dihydroguanine (8-oxoGua), 8-oxo-7,8-dihydro-2'-
deoxyguanosine (8-oxodGuo), 1-hydroxypyrene (1-OHP) and 1-OHP-glucuronide
(Gedik et al., 2002).

On the other hand, biomonitoring involves the use of organisms often regarded
as sentinels or biomonitors. These organisms (including the various life stages) are
often sensitive to one or more potential toxic compounds within the environment.
Thus, a biomonitor could be described as organisms within a test system which
provide quantitative information on the quality of its immediate environment
(Madejón et al., 2006; Markert et al., 2003). This information is often presented in any
of the following ways: physiological, chemical or behavioural modifications.
Table 1: Biomarkers indicating exposure and toxic responses in various organisms to chemicals, including polycyclic aromatic hydrocarbons.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humans</td>
<td>8-Oxo-7,8-dihydroguanine (8-oxoGua), 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxoG), Catalase activity, glutathione S-transferase (GST), 1-hydroxypyrene (1-OHP) 1-OHP-glucuronide, Monohydroxy-phenanthrenes, Malonaldehyde (MDA), DNA damage/micronucleus formation, 4-Hydroxyalkenals</td>
</tr>
<tr>
<td>Molluscs</td>
<td>Lysosome membrane stability, Lysosomal lipofuscin, Lysosomal neutral lipid, CaATPase activity, Catalase activity, Total oxidant scavenging capacity, Acetylcholinesterase activity, Malondialdehyde, DNA damage/micronucleus formation, Glutathione-S-transferase in haemolymph, *Metallothioneins</td>
</tr>
<tr>
<td>Birds</td>
<td>Ferritin and Haptoglobin,</td>
</tr>
<tr>
<td>Lichens</td>
<td>Photosynthesis, Chlorophyll content/degradation, Endogenous auxin levels, Ethylene production</td>
</tr>
</tbody>
</table>

* biomarkers of exposure to specific classes of chemicals particularly heavy metals such as lead, cadmium, etc. (Brambilla et al., 1986; Gedik et al., 2002; Park et al., 2006; Scherer et al., 2000; Tintos et al., 2007).
Fig. 2 Processes affecting organism uptake of external substances via biological membranes, their internal distribution, and possible effects. Modifying factors often include genetic variability, age, gender and lifestyle of the exposed organism.
Toxicity is controlled by toxicokinetics that govern the bioaccumulation and distribution of chemicals in tissues (based on their physical and chemical properties and facility for biotransformation) and by toxicodynamics, which govern the biochemical and physiological response of the organism (Fig. 2). The closer the relationship between the concentration of the toxicant in whole tissues and the concentration at the site of toxic action, the better the interpretation of the dose-response relationship (McCarty et al., 2011). Exposure concentration and composition of compounds is often modified by biotransformation of the compound within the exposed organism, increasing the parent compound elimination rate and decreasing the equilibrium concentration of the parent chemicals in tissues. However, biotransformation can change the inherent toxicity of the parent compound because metabolites can be more or less toxic.

Enzyme systems (mixed function oxidases e.g. cytochrome p450s) exists in biological organisms which promote the conversion and biotransformation of hydrocarbons to metabolites (Fig. 3) capable of altering biological macromolecules e.g. DNA (Boysen and Hecht, 2003; Malins and Hodgins, 1981). The toxicity of PAHs occurs as a result of this biotransformation. In themselves, PAH parent compounds are rather unreactive and express little toxicity (Walker, 2012). Major research on toxicity particularly human toxicology focuses on the mutagenic and carcinogenic actions of PAHs because, DNA adducts formed by metabolites of carcinogenic PAHs and other compounds predispose organisms to mutation cancers (Hemminki et al., 2000). While acute toxicity of PAHs to mammals is low, the toxicity of PAHs to aquatic organisms may arise from bioactivation, as well as depend on the level of ultraviolet (UV) radiation to which the organism or test system is exposed (Lampi et al., 2006). PAHs possess the ability to undergo photo-oxidation
(McConkey et al., 2002) and have been shown as photo-mutagenic with a possibility of being activated by light irradiation, without requiring metabolizing enzymes (Yan et al., 2004). The vast majority of toxicity studies are carried out using in vitro assays including Bacterial Mutagenicity Assays (Gatehouse, 2012) and the Comet assay (single-cell gel electrophoresis) (Speit and Hartmann, 2006). The various assays used for toxicity testing are quite sensitive and detect various potential toxic compounds. However, they are also prone to false positives and present the challenge of extrapolating results to realistic and environmentally relevant scenarios.
**Fig. 3:** The metabolic activation of a typical polycyclic aromatic hydrocarbon (Benzo[a]pyrene) into DNA reactive metabolites. Following exposure, most chemical carcinogens are subject to biotransformation, catalysed by “xenobiotic metabolizing enzymes” such as cytochrome P450-dependent monooxygenases (CYPs), hydrolases, and transferases. This process could lead to the generation of electrophilic derivatives which are capable of reacting with DNA and inducing mutations of tumour susceptibility genes such as oncogenes.
1.5 Vibrational Spectroscopy

Molecular vibrations can range from the simple coupled motion of the two atoms of a diatomic molecule to the much more complex motion of each atom in a large poly-functional molecule. Molecules with \( N \) atoms will have \( 3N \) degrees of freedom, three of which represent translational motion in mutually perpendicular directions (\( x, y \) and \( z \) axes) and three represent rotational motion about the \( x, y \) and \( z \) axes. The remaining \( 3N - 6 \) degrees of freedom give the number of ways that the atoms in a nonlinear molecule can vibrate (Griffiths and De Haseth, 2007).

A complex molecule exhibits a variety of vibrational modes which involve the whole molecule. However, some of these molecular vibrations are associated with the vibrations of individual bonds or functional groups (localized vibrations). These localized vibrations are stretching, bending, twisting, rocking or wagging (Williams et al., 1995). The study of the interaction between matter and electromagnetic radiation depends on the wavelength or frequency of the radiation, such that regions of the electromagnetic spectrum become associated with various types of spectroscopy, and the frequency ranges named after the most common source of the radiation (e.g. X-rays) or its practical use (e.g. Radio). These modes are dependent on the type of atoms present and their structural arrangements (Painter et al., 1982).

Vibrational spectroscopic techniques have become potential tools for non-invasive optical tissue diagnosis and have been applied to study a wide variety of pathologic states. A wide variety of biological tissues have been studied using various forms of vibrational spectroscopy especially infrared (IR) (Movasaghi et al., 2008) and Raman (Movasaghi et al., 2007) spectroscopy.
1.5.1 Fourier Transform Infrared (FTIR) Spectroscopy

1.5.1.1 Background

IR rays were discovered by William Herschel in 1800 and are absorbed by matter in the form of several bands localized in discrete frequency intervals. The basis for the widespread use of IR spectroscopy is the observation that many chemical groups such as C = O, absorb in a relatively narrow frequency range, irrespective of the nature of the other functional groups present. Within this frequency range, the observed frequency can be correlated to specific chemical structures and the spectral pattern may be likened to a “molecular fingerprint” particularly because similar molecules may have significantly different IR spectra, especially in the region below 1500 cm\(^{-1}\) (Painter et al., 1982). The IR spectra results from transitions between quantized vibrational energy states with the usual range between 4000 cm\(^{-1}\) at the high frequency range and 625 cm\(^{-1}\) at the low frequency end (Griffiths and De Haseth, 2007).

In the initial stages, spectra acquisition was a time consuming process because spectrometers utilized the technology of a moving grating monochromator, to disperse the single wavelength of the spectrum from a broad range of wavelengths and therefore, only a wavelength of single resolution could be detected at a given time (Stuart, 2005). With the introduction of interferometers, light covering the whole frequency range, typically 5000 – 400 cm\(^{-1}\), is split into two beams and either one beam is passed through the sample or both beams are passed, but one beam is made to traverse a much longer path than the other. The recombination of the two beams produces an interference pattern that is the sum of all the interference patterns created by each wavelength in the beam. By systematically changing the difference in the two paths, the interference patterns change to produce a detected signal varying with optical path difference. This pattern is the interferogram and although it looks nothing
like a spectrum, Fourier transform of the interferogram using a computer built into the machine, converts it into a plot of absorption against wavenumber which resembles the usual spectrum obtains by the traditional method (Williams et al., 1995). FTIR method of acquiring spectra is faster and provides a higher signal to noise ratio (SNR).

Fig 3: showing a schematic illustration of the internal components of a Fourier transform infrared spectrometer (FTIR), fitted with a Michelson interferometer.
Common IR light sources are globar (black body) or synchrotron-based radiation, where globar is a silicon carbide thermal mid-IR source, emitting radiation from $\lambda = 2.5\ \mu\text{m} - 25\mu\text{m}$ ($\bar{\nu} = 4000 - 400\ \text{cm}^{-1}$), and synchrotron-based radiation is ~1000 times brighter than globar sources, producing IR spectra with a significantly higher SNR (Kelly et al., 2011).

The potential of a stand-alone FTIR spectrometer may seem great. However, the coupling of an FTIR to a visible light microscope (FTIR microspectrometry) greatly increases its potential as it permits the examination of complex molecules (e.g. biological tissues) and heterogeneous samples. Infrared microscopes are often high quality visible microscopes redesigned for use with IR radiation. Detection by microscopy may be accomplished by laser-scanning a point illuminated on the sample or by using wide-field illumination and focal plane array (FPA) or linear array detectors (Baker et al., 2014)

1.5.1.2 Instrumentation

Transmission, transflection and attenuated total reflection (ATR) are the three major IR-spectroscopic sampling modes by which spectra acquisition may be carried out. In transmission mode, the IR beam is directed through a sample and collected by a condenser whereas, in transflection mode, the beam is directed through the sample, reflects off an IR-reflective surface [such as that found on low-emissivity (Low-E) slides], travels back through the sample to the detector. With both measurements, the sample thickness is an important criterion, as extremely thick samples will attenuate the IR beam beyond the range where absorption is proportional to chemical concentration and very thin samples will result in low absorption where acquired spectra signal is flooded with noise (Kelly et al., 2011)
The ATR has grown into the most widely practiced technique in IR spectrometry especially because, the technique involved requires little or no sample preparation and consistent results can be obtained with relatively little care or expertise. The ATR mode of spectra acquisition involves passing the IR beam through an internal reflection element [(IRE) usually an IR-transparent element] with a high refractive index [e.g. Zinc Selenide (ZnSe), type II diamond or Germanium (Ge)] (Walsh et al., 2007). When the IRE is placed in contact with the sample and the beam passed through it, the beam is totally internally reflected, generating an evanescent wave which penetrates a few µm beyond the element into the sample (Kelly et al., 2011). The depth of penetration varies from a fraction of a wavelength up to several, depending on the index of refraction of the element and the angle of the incident radiation with respect to the interface between sample and element. It is also wavelength-dependent, increasing with increasing wavelength and has the consequence that if the sample selectively absorbs certain wavelength components of the evanescent radiation, then attenuation of the reflected beam occurs preferentially at the wavelength of absorbance bands (Walsh et al., 2007).

Attenuated total reflection can be said to be the most versatile of all IR sampling techniques because, it requires very little sampling preparation and can be used on samples of almost all morphologies, while often maintaining the structural integrity of the sample. ATR is in large part a surface technique and the interrogation of sample is largely limited to the depth of penetration of the measurement (Griffiths and De Haseth, 2007).
Table 2: FTIR spectroscopy modes commonly used for the interrogation of cellular materials

<table>
<thead>
<tr>
<th>Mode</th>
<th>Suitable samples</th>
<th>Substrate</th>
<th>Typical interrogation area (μm)</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATR</td>
<td>Tissues, cells, biofluids</td>
<td>Calcium or barium fluoride, zinc selenide, MIRR IR Low E</td>
<td>250 × 250</td>
<td>• High SNR</td>
<td>• Low resolution&lt;br&gt;• Can be destructive due to pressure&lt;br&gt;• Air between sample and IRE may affect spectra&lt;br&gt;• Minimum sample thickness required (around 2.3 μm)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Reduced scattering&lt;br&gt;• Analysis of large target area&lt;br&gt;• Better for aqueous samples</td>
<td></td>
</tr>
<tr>
<td>Transmission</td>
<td>Tissues, individual cells, cellular components, biofluids</td>
<td>Calcium or barium fluoride, zinc selenide</td>
<td>5× 5 to 150 × 150</td>
<td>• High resolution&lt;br&gt;• Non destructive&lt;br&gt;• Automated stage allows for spectral acquisition at several different locations of choice with little user interaction</td>
<td>• Lower SNR than ATR&lt;br&gt;• Maximum sample thickness required&lt;br&gt;• Longer sample and machine preparation required</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Automated stage allows for spectral acquisition at several different locations of choice with little user interaction</td>
<td></td>
</tr>
<tr>
<td>Transflection</td>
<td>Tissues, individual cells, cellular components, biofluids</td>
<td>Calcium or barium fluoride, zinc selenide</td>
<td>5× 5 to 150 × 150</td>
<td>• High resolution&lt;br&gt;• Non destructive&lt;br&gt;• Automated stage allows for spectral acquisition at several different locations of choice with little user interaction</td>
<td>• May give rise to standing wave artifacts&lt;br&gt;• Lower SNR than ATR&lt;br&gt;• Maximum sample thickness required&lt;br&gt;• Longer sample and machine preparation required</td>
</tr>
</tbody>
</table>

Source: (Baker et al., 2014).
Fig. 4: Schematic illustration of spectra acquisition using attenuated total reflection mode in infrared (IR) spectroscopy.
1.6 Data Handling

Many studies involve the processing of data derived from several samples grouped into a number of classes and possible containing several variables. The interrogation of samples using spectroscopy, generates large and complex datasets which require robust methods to extract specific information of interest e.g., factors responsible for variability between groups or classes in the dataset. Using specific approaches or a combination of two or more approaches, it is possible to reduce the complexity of datasets and extract meaningful underlying variance within variables.

1.6.1 Spectra Pre-processing

Spectra pre-processing, regarded also as the manipulation of spectra is often performed using software packages built into an FTIR spectrometer (Table 3). Following spectra acquisition and prior to analysis, the acquired spectra must be pre-processed in order to account and correct for noise, sloping baseline effects, differences in sample thickness or concentration, and to select the regions of interest. This process can be summarised thus: cutting, baseline correction and normalization. Skewed baselines in acquired spectra could occur as a result of several factors including resonant Mie scattering; occurring when the wavelength of IR light is comparable or smaller than some of the molecular structures through which it passes, causing the passing light to scatter. Other factors may be reflection, temperature, concentration or instrument anomalies. Baseline correction can be achieved using techniques such as the rubberband baseline correction; stretching the spectra down so minimal areas of the spectral region of interest are used to fit a convex polygonal line and then subtracted from the original spectrum. Baseline correction may also be carried out by differentiating the spectra twice, causing the spectra to lose their slope.
Spectral distortions due to Mie scattering can be corrected using the Extended Multiplicative Signal Correction (EMSC) algorithm (Kohler et al., 2008), which is particularly effective where Mie scattering is weak and where the spectra do not show strong distortion of the Amide I band (Bassan et al., 2010).

Normalization is employed to scale the spectra and remove spectral changes accountable by the thickness or concentration of the sample, thus making all spectra in a batch comparable to each other. Normalization approaches include the Min-max normalization; applied when a known peak is stable and consistent across the specimens e.g. Amide I in animal cell and tissue samples, and vector normalization; where each spectrum is divided by its Euclidean norm rather than relying on a specific peak. The vector normalization method is recommended where differentiation has been carried out as a baseline correction method (Kelly et al., 2011). Pre-processing of spectra is an important step in data handling particularly with IR spectroscopy. the outcome of computational analysis is heavily dependent on the effectiveness of the techniques used to deal with unwanted variability (“noise”) in the data (Martin et al., 2010).

Other forms of pre-processing techniques include spectral subtraction; mostly applied to obtain the spectrum of a component in a mixture, De-noising; used to enhance the information content of a spectrum by removing the “noise” in the spectrum, and deconvolution; mathematical enhancement of spectrum resolution and particularly used to distinguish the positions of overlapping bands within a spectrum. It is important to note that transmission and single beam spectra should not be subtracted because their peak heights and areas are not linearly proportional to concentration (Smith, 2011).
**Fig. 5**: An illustration of a spectra pre-processing technique (baseline correction). Acquired spectra (black) was baseline corrected (blue) using the rubberband baseline correction method.
**Table 3:** Existing FTIR spectroscopy data analysis software.

<table>
<thead>
<tr>
<th>Software</th>
<th>Website</th>
<th>Description</th>
<th>License</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytospec</td>
<td><a href="http://www.cytospec.com">www.cytospec.com</a></td>
<td>Software for hyperspectral imaging (IR, Raman)</td>
<td>Commercial; free demo available</td>
</tr>
<tr>
<td>IRootLab</td>
<td>irootlab.googlecode.com</td>
<td>MATLAB toolbox for biospectroscopy data analysis</td>
<td>Open-source</td>
</tr>
<tr>
<td>OPUS</td>
<td><a href="http://www.bruker.com">www.bruker.com</a></td>
<td>Spectral acquisition software with data processing capabilities</td>
<td>Commercial</td>
</tr>
<tr>
<td>Pirouette</td>
<td><a href="http://www.infometrix.com">www.infometrix.com</a></td>
<td>Chemometrics modelling software</td>
<td>Commercial</td>
</tr>
<tr>
<td>Unscrambler X</td>
<td><a href="http://www.camo.com">www.camo.com</a></td>
<td>Multivariate data analysis and design of experiments</td>
<td>Commercial</td>
</tr>
<tr>
<td>PLS, MIA, EMSC</td>
<td><a href="http://www.eigenvector.com">www.eigenvector.com</a></td>
<td>MATLAB toolboxes for spectroscopy data analysis</td>
<td>Commercial</td>
</tr>
<tr>
<td>OMNIC</td>
<td><a href="http://www.thermoscientific.com">www.thermoscientific.com</a></td>
<td>Spectral acquisition software with data processing capabilities</td>
<td>Commercial</td>
</tr>
<tr>
<td>PyChem</td>
<td><a href="http://pychem.sourceforge.net/">http://pychem.sourceforge.net/</a></td>
<td>Package for univariate and multivariate data analysis</td>
<td>Open-source</td>
</tr>
<tr>
<td>ENVI, IDL</td>
<td><a href="http://www.exelisvis.com">www.exelisvis.com</a></td>
<td>Integrated development, data analysis, image processing suite</td>
<td>Commercial</td>
</tr>
</tbody>
</table>

Source: (Baker et al., 2014).
1.6.2 Feature Extraction, Construction and Selection

In statistics, “features” is a term synonymous with “input variables”, i.e., the inputs to the subsequent analysis method. In IR spectroscopy, the wavenumber absorbance intensities can be used as features. However, it is important to reduce the number of variables within the dataset in order to avoid the “curse of dimensionality” which manifests as over-fitting and can lead to the poor performance of classifiers when tested on independent data, or the formulation of incorrect hypotheses drawn from exploratory analysis (Jain et al., 2000).

Feature construction and selection serves as an approach to reduce the number of variables within a data set and include common techniques such as principal component analysis (PCA), linear discriminant analysis (LDA) or partial least squares (PLS), where the constructed variables are linear combinations of the wavenumbers (Kelly et al., 2011). Although it is practically impossible to find the best subset from all the existing $2^n$ possibilities, where $n$ is the number of wavenumbers originally present in the dataset, several suboptimal feature selection strategies exist which could rank the relevance of the wavenumbers individually based on an evaluation criterion e.g. Pearson correlation or $t$-test, in order to determine which wavenumbers are retained or, the application of a selection algorithm “wrapped around” a classifier used to rank the subsets (Guyon and Elisseeff, 2003).

1.7 Computational Analysis

The complexity of the datasets generated from spectrometry presents the challenge of extracting meaningful underlying variances within variables. Thus, computational analysis employs mathematical algorithms (tools) to extract the variance within variables. Most of the approaches are linear techniques of feature extraction and
multivariate data analysis tools such as PCA, LDA, PLS or a combination of PCA and LDA (PCA-LDA) which generates loading vectors with the ability to identify the contribution of each wavenumber-variable to generate the new variable (Trevisan et al., 2012)

1.7.1 Principal Component Analysis

PCA is an unsupervised exploratory data analysis tool which reveals relationships in data that might have otherwise been ignored or not observed. PCA is most often the first analysis performed on a new dataset, because being unsupervised, it is unbiased and reveals the most prominent variation patterns in data, whether these variations are correlated to classes or not (Trevisan et al., 2012). Within IR spectroscopy, PCA is employed to reduce dimensionality and generate a visualization of data. It is a linear transformation of the wavenumber dataset operated by the PCA loadings matrix. The loadings vectors (principal components [PCs]) within this matrix are eigenvectors of the covariance matrix of the data and each loadings vector contains the coefficients of a linear combination that generates one new variable called a PCA factor. PCA factors are uncorrelated and each PC has a corresponding eigenvalue which exactly matches the variance of its corresponding PCA factors, enabling these factors to be ranked according to the magnitude of variance captured by each one. Thus, the first 3 PCs are most commonly used as they contain the most variance, often up to 99%, ensuring optimum visualization of the data (Kelly et al., 2011). Using PCA, each spectrum is viewed as a single point or score in n-dimensional space and selected PCs are used as Cartesian coordinates to reveal clusters which inform the formulation of hypotheses regarding similarities or differences in dataset by exploiting proximity or segregation levels between clusters. This approach allows the intrinsic dimensionality of large and
complex data to be interrogated and analysed for clustering when viewed in a particular direction (Davies and Fearn, 2004).

1.7.2 Linear Discriminant Analysis

LDA is a supervised technique which is used to achieve class segregation. It is more likely to over-fit, if the number of spectra is insufficient. Thus, it is generally recommended that the number of spectra in the dataset be 5-10 times bigger than the number of variable (Trevisan et al., 2012). LDA forms linear combinations of variables dependent on the differences between the classes in the dataset and the LDA loading vectors are successive orthogonal solutions to the problem to “maximize the between-class variance over the within-class variance”. Following the application of LDA on a dataset, the dataset will only have $c - 1$ variable, where $c$ is the number of data classes (Kelly et al., 2011). Applied to IR spectroscopy especially with regards to answering biological questions, LDA is used to reduce confounding factors of within-category heterogeneity whilst maximizing between-category discriminating biomarkers (Martin et al., 2007).

1.7.3 Partial Least Squares Regression

PLSR is a supervised multivariate analysis method which addresses the problem of making good predictions in multivariate datasets (Mehmood et al., 2012). PLSR constructs a set of linear combinations of the wavenumbers the same way as PCA, but uses the data classes in the construction. Where PCA ranks the PCs according to variance within the dataset, PLSR employs a different approach by finding a sequence of new variables that are maximally correlated with a numerical representation of the data classes while being independent to each other. PLSR has a tendency to over-fit and requires more validation than PCA (Kelly et al., 2011).
1.7.4 Combined Multivariate Analysis: PCA-LDA

The cascade application of LDA on the factors resulting from PCA is a popular multivariate analysis performed on IR spectral datasets particularly with regards to biomarker extraction (Trevisan et al., 2012). PCA-LDA gives loading vectors which identify the contribution of each wavenumber-variable to generate the new variables (factors). The weights for each factor are represented by a vector called a “loadings vector”. PCA-LDA presents each scalar value of each factor as a “score” which may be visualized through 1-, 2-, or 3-dimensional scatter plots also known as “scores plots.” LDA allows for data visualization in the form of cluster vector plots which may be used to identify biomarkers (i.e. wavenumbers) associated with specific treatment conditions. Each cluster vector is a linear combination of the loadings vectors and it can be plotted as y-values having the wavenumbers as x-values. It is also possible to apply a peak detection algorithm to identify prominent peaks within each cluster vector (Llabjani et al., 2011).

1.7.5 Visualization of Processed Data

Following the application of multivariate analysis, results could be visualized in multiple fashions. Most commonly used are the scores and loadings plots as well as the cluster vector approach. Scores plots are scatter charts drawn using the data values obtained after multivariate analysis as Cartesian coordinates. Scores and loadings plots provide a visual representation and interpretation of variables responsible for any segregation, following the construction of factors from any one of the multivariate techniques earlier discussed. The loadings vectors of the afore mentioned techniques have the same resolution as the original spectra and their coefficients can be plotted against the wavenumber axis to reveal the contributions of each wavenumber to form each corresponding factor (Kelly et al., 2011).
The cluster vector approach is a geometric construction applied to linear multivariate techniques although it is most commonly employed following the application of PCA-LDA (German et al., 2006). The idea of cluster vectors follows from the fact that loading vectors are found to be more informative when they “pass through” data points rather than pointing towards void space. There is therefore one cluster vector for each data class where each cluster vector is a vector that points from the origin to the centre of its corresponding data class in the vector space spanned by the vectors (Kelly et al., 2011).
Fig. 6 Flow diagram of data processing and analyses used in the various projects contained in this thesis. Red boxes represent pre-processing methods, Green box represents computational analysis using multivariate approach and the purple boxes represent the output and visualization.
Fig. 7 An example of cluster vector derived following the application of multivariate analysis. In this case, cluster vectors were used to visualize results after the application of principal component analysis combined with linear discriminant analysis (PCA-LDA). A peak detection algorithm was applied to detect prominent peaks which distinguished each class in dataset (green and red lines) from the reference class (black line).
Fig. 8 showing the potential biomarkers extractable from the IR biochemical fingerprint of cells and tissues. Alterations to any one of the identified regions of the fingerprint could provide an insight as to the mechanistic action (e.g. genomic damage) of the compound under observation.
1.8 Summary of publications

1.8.1 Novel biospectroscopy sensor technologies towards environmental health monitoring in urban environments

Obinaju and Martin (2013) presents a brief introduction to the use of vibrational spectroscopy techniques for real-time detection of sub-lethal effects of compounds in the environment. It highlights previous successful applications of vibrational techniques particularly in clinical diagnosis and most importantly, it presents a hypothetical scenario using land snails (*Helix aspersa*) as potential biomonitors. (Chapter 2).

1.8.2 Distinguishing nuclei-specific benzo[a]pyrene-induced effects in MCF-7 cells from whole-cell alterations using Fourier-transform infrared spectroscopy

Several chemical compounds possess the ability to alter metabolic processes within various parts of the cell, alterations which are key to the onset or progression of pathologic conditions. Obinaju et. al., (2015a) observed the changes occurring in the nucleus the human mammary carcinoma (MCF-7) cell line as a result of exposure to concentrations of a typical carcinogen polycyclic aromatic hydrocarbon: benzo[a]pyrene. Observations were carried out using ATR-FTIR and the study observed that ATR-FTIR was able to detect changes to nucleus and highlight wavenumbers responsible for these changes. Importantly it was able to highlight wavenumbers which could indicate the possibility of apoptotic induction at high concentrations. (Chapter 3).
1.8.3 PAH biomonitoring in Niger Delta, Nigeria

1.8.3.1 Novel sensor technologies towards environmental health monitoring in urban environments: a case study in the Niger Delta (Nigeria).

Environmental pollution in the Niger delta is a persistent concern. Using ATR-FTIR, Obinaju et. al., (2014) interrogated various tissues of the African catfish (*Heterobranchus bidorsalis*) and leaves of the water spinach (*Ipomea aquatica*) sampled from the Ovia River; recipient of petroleum hydrocarbons. The study observed that ATR-FTIR was able to discriminate between samples from various sites. It was also able to discriminate between samples obtained at different seasons and able to detect alterations in tissues relative to chosen controls (Chapter 4).

1.8.3.2 Linking biochemical perturbations in tissues of the African catfish to the presence of polycyclic aromatic hydrocarbons in Ovia River, Niger Delta region.

Obinaju et. al., (2015b) measured the concentration of PAH compounds in the dissolved phase of the Ovia River in the Niger Delta, and attempted to relate the concentrations of PAHs detected to the observed biochemical changes in tissues of the African catfish. It explored the potential impact of seasonal variations on the observed changes. These changes were documented as shifts in centroid positions of absorption bands as well as increased or reduced intensity to bands at certain wavenumbers (Chapter 5).
1.8.3.3 Attenuated total reflection Fourier-transform infrared spectroscopy reveals polycyclic aromatic hydrocarbon contamination despite relatively pristine characteristics of site: results of a field study in the Niger Delta

Biomarkers for disease can be identified by comparing the IR spectra of malignant tissue samples to the IR spectra signature of a reference ‘normal’ tissue. Similarly, it is possible to distinguish between chemical exposures in tissues based on the IR spectra of the exposed tissues. This study aimed to identify the biomarkers of PAH exposure in the tissues of the African catfish by comparing tissues obtained from sites with known contamination sources and potential contaminant compounds, to tissues obtained from a relatively pristine site. It found that using spectra of samples exposed to known compounds, ATR-FTIR was able to identify potential exposure to similar compounds in samples from sites with undocumented contaminant history (Chapter 6).

1.9 Aims and objectives

This thesis is composed of four primary author research projects which investigate the application of FTIR spectroscopy to detect and measure sub-lethal effects of potential mutagens in the environment using sentinel organisms.

The study hypothesizes that 1) the technique ATR-FTIR is sensitive to and able to detect minimal cellular changes occurring in tissues exposed to potential mutagens. 2) ATR-FTIR can extract potential biomarkers to signature chemical induced changes in tissues.

Also included in the appendix is a co-author project which explores the potential standardization of methods and procedures that could optimise the application of IR
spectroscopy to an even wider variety of biological questions including disease screening and diagnosis.

- To investigate the effects of low dose PAH exposure in intact cells and isolated nuclei as a baseline for real-time environmental exposure scenarios, using ATR-FTIR spectroscopy (Chapter 3).
- To detect seasonal variations in exposure and real-time exposure effect in fish tissue and plant leaves, using ATR-FTIR spectroscopy (Chapter 4).
- To investigate the correlation of PAH concentrations detected in dissolved phase of water column to the alterations in fish tissues (Chapter 5)
- To explore the potential identification of biomarkers of PAH exposure in fish tissues using ATR-FTIR spectroscopy (Chapter 6)
References


Glotfelty, D., Williams, G., Freeman, H., Leech, M., 1990. Regional atmospheric transport and deposition of pesticides in Maryland. Lewis Publishers: Chelsea, MI.


Mackay, D., Shiu, W.-Y., Ma, K.-C., Lee, S.C., 2006. Handbook of physical-chemical properties and environmental fate for organic chemicals. CRC/Taylor and Francis, Boca Raton, FL.


Martin, F.L., 2007. Complex mixtures that may contain mutagenic and/or genotoxic components: a need to assess in vivo target-site effect (s) associated with in vitro-positive (s). Chemosphere 69, 841-848.


Chapter 2

Novel biospectroscopy sensor technologies towards environmental health monitoring in urban environments.

Blessing E. Obinaju and Francis L. Martin

Environmental Pollution 183 (2013) 46-53.

Contribution:

- I prepared the first draft of the manuscript

-------------------------------
Blessing E. Obinaju           Prof Francis L. Martin
Chapter 3

Distinguishing nuclei-specific benzo[a]pyrene-induced effects in MCF-7 cells from whole-cell alterations using Fourier-transform infrared spectroscopy

**Blessing E. Obinaju,** Nigel J. Fullwood and Francis L. Martin

Toxicology 335 (2015) 27–34

Contribution:

- I conducted all experiments for the study
- Scanning Electron Microscopy was performed by Nigel Fullwood
- I acquired the data and carried out the computational analysis
- I prepared the first draft of the manuscript

Blessing E. Obinaju
Prof Francis L. Martin
Chapter 4

Novel sensor technologies towards environmental health monitoring in urban environments: a case study in the Niger Delta (Nigeria).

Blessing E. Obinaju, Alozie Alaoma and Francis L. Martin

Environmental Pollution 192 (2014) 222-231.

Contribution:

- I acquired the samples required for the project
- I prepared processed and acquired 80% of the data and carried out computational analysis.
- I prepared the first draft of the manuscript

Blessing E. Obinaju

Prof Francis L. Martin
Chapter 5

Linking biochemical perturbations in tissues of the African catfish to the presence of polycyclic aromatic hydrocarbons in Ovia River, Niger Delta region.

Blessing E. Obinaju, Carola Graf, Crispin Halsall and Francis L. Martin

Environmental Pollution 201 (2015) 42-49

Contribution:

- I acquired the samples required for the project

- I prepared processed and acquired data for African catfish including carrying out computational analysis.

- Water analysis for polycyclic aromatic hydrocarbon was acquired by Carola Graf, Crispin Halsall.

- I prepared the first draft of the manuscript

................................. ..........................................................

Blessing E. Obinaju                      Prof Francis L. Martin
Chapter 6

Attenuated total reflection Fourier-transform infrared spectroscopy reveals polycyclic aromatic hydrocarbon contamination despite relatively pristine characteristics of site: results of a field study in the Niger Delta

Blessing E. Obinaju and Francis L. Martin

Contribution:

- I acquired the samples required for the project
- I prepared processed and acquired data for African catfish and carried out computational analysis.
- I prepared the first draft of the manuscript

Blessing E. Obinaju               Prof Francis L. Martin
Attenuated total reflection Fourier-transform infrared spectroscopy reveals polycyclic aromatic hydrocarbon contamination despite relatively pristine characteristics of site: results of a field study in the Niger Delta

Blessing E. Obinaju and Francis L. Martin*

_Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YQ, UK_

**Corresponding author:** Prof Francis L. Martin PhD, Centre for Biophotonics, LEC, Lancaster University, Lancaster LA1 4YQ, UK; Tel.: +44(0)1524 510206; Email: [f.martin@lancaster.ac.uk](mailto:f.martin@lancaster.ac.uk)
Abstract

Fourier-transform infrared (FTIR) spectroscopy is becoming a technique to detect biochemical alterations in biological tissues, particularly changes due to sub-lethal exposure to environmental contaminants. We have previously shown the potential of attenuated total reflection FTIR (ATR-FTIR) spectroscopy to detect real-time exposure to contaminants in sentinel organisms as well as the potential to relate spectral alterations to the presence of specific environmental agents. In this study based in the Niger Delta (Nigeria), changes occurring in fish tissues as a result of polycyclic aromatic hydrocarbon (PAH) exposure at contaminated sites are compared to the infrared (IR) spectra of the tissues obtained from a relatively pristine site. Multivariate analysis revealed that PAH contamination could be occurring at a pristine site, based on the IR spectra and significant ($P <0.0001$) differences between sites. The study provides evidence of IR spectroscopy techniques’ sensitivity and supports their potential application in environmental biomonitoring.

**Keywords:** African catfish; Environmental pollution in Nigeria; Fourier-transform infrared spectroscopy; *Heterobrachus bidorsalis*; Niger Delta pollution; Polycyclic aromatic hydrocarbon
1. Introduction

Human activities generate potentially toxic compounds, some with unusual characteristics. Most of these compounds end up in various parts of the ecosystem and constitute a degree of hazard to biological populations including humans. Synthetic chemicals and materials such as petroleum hydrocarbons, persistent organic pollutants (POPs), pesticides and metals which mostly contaminate aquatic systems, are more often linked to advancements in industrialization (Li et al., 2001; Zhang et al., 2005). Exposures to these compounds have been linked to a variety of adverse effects including neurodevelopmental effects following *in utero* exposure (Perera and Herbstman, 2011; Perera et al., 2006; Wormley et al., 2004).

Studies continue to show that chemical contaminants are capable of inducing toxicity in organisms, even at very low concentrations (Kalantzi et al., 2004; Pang et al., 2012; Ukpebor et al., 2011), and their accumulation in the tissues of organisms following exposure, particularly aquatic and wildlife species, generates concern for the possibility of contaminant transfer through the food chain (Gwaski et al., 2013; Lozano et al., 2012). Several analytical techniques exist to biomonitor contaminants in the ecosystem and within organisms. In recent years, the field of biospectroscopy, a technique which employs the use of infrared (IR) spectrometry or the coupling of IR spectrometry to other techniques (*e.g.*, microscopy [IR microspectroscopy]) to understand changes in cells and tissues, especially those which occur as a result of exposure to environmental chemicals, has gained immense attention. The application of biospectroscopy to observe these changes, is based on the knowledge of the vibrational modes of biomolecules which generates spectral information [often known as the “biochemical-cell fingerprint” (biofingerprint)] when exposed to IR radiation (Martin et al., 2010). Based on changes to the biofingerprint, it is possible to
distinguish between cell/tissue types (German et al., 2006) with potential cell characterization (Grude et al., 2007).

Biospectroscopy techniques are non-destructive to samples, relatively reagent-free and can generate rapid, high-throughput and robust results in real-time with high sensitivity to minimal changes within biomolecules (Martin et al., 2010). Thus, biospectroscopy can be employed to study contaminant-induced responses in organisms, using a wide variety of sample types and particularly, it has the potential to biomonitor environmental contaminants in most sentinels, in real-time (Ibrahim et al., 2012; Llabjani et al., 2012; Malins and Gunselman, 1994; Obinaju et al., 2014). These techniques can be optimised for even more applications (Baker et al., 2014). Biospectroscopy techniques involving the use of attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy require minimal sample preparation (Martin et al., 2010; Obinaju and Martin, 2013) and have been shown to detect the slightest chemical-induced variation in samples at very low ($< 10^{-9}$ M) concentration ranges (Ahmad et al., 2008; Llabjani et al., 2014; Llabjani et al., 2011; Ukpebor et al., 2011).

We have previously shown that ATR-FTIR spectroscopy is able to differentiate between real-time exposure effects both animal and plant tissues from sites with varying degrees of environmental contamination (Obinaju et al., 2014). Herein, we compare tissues of the African Catfish (*Heterobrachus bidorsalis*) from sites with a known history of polycyclic aromatic hydrocarbon (PAH) contamination, to samples from a relatively pristine site with no documented history of contamination and no industrial activity. Our aim was to determine if we could signature PAH-induced toxicity in fish tissues using ATR-FTIR spectroscopy.
2. Materials and Methods

Samples of *Heterobrachus bidorsalis* were collected in March 2013 by local fishermen at Gelegele, Ikoro and Ifiayong in Edo and Akwa Ibom States, within the Niger Delta region. Site descriptions as well as sample handling, tissue pre-processing methods and spectral measurements have been previously detailed (Obinaju et al., 2014). Briefly, each site was selected based on the documented knowledge of industrial activities, which yield possible environmental contaminating compounds. Gelegele and Ikoro are located in close proximity to petroleum exploration activities and Ifiayong is a rural fishing community with no documented history of petroleum exploration or similar industrial type activities. Each excised fish tissue was thinly sliced (≤ 1-mm thick/slice) by hand using a Stadie-Riggs handheld microtome and Thomas blade (Taylor et al., 2011). Each slice was rehydrated by washing twice in dH₂O. Sample slices were mounted on Low-E reflective glass slides (Kevley Technologies, Chesterland, OH), allowed to air-dry and desiccated for a minimum of 24 h. prior to interrogation using ATR-FTIR spectroscopy.

2.1 Spectral acquisition and pre-processing

IR spectra were obtained using a Bruker Vector 27 FTIR spectrometer with Helios ATR attachment containing a diamond crystal (Bruker Optics Ltd., Coventry, U.K.). Data acquired for each experimental condition (i.e., each sample slide) consisted of 10 spectra, each from a random area of the tissue slice under interrogation, using an FTIR imaging system coded for 32 scans per spectra and 3.84 cm⁻¹ spectral resolution. The ATR crystal was cleaned with dH₂O, dried thoroughly and a new background spectrum taken prior to analysis of a new sample. Raw spectra were acquired in the 4000 cm⁻¹ - 400 cm⁻¹ range. Spectra in the region of interest (1800 - 900 cm⁻¹) were selected and pre-processed (baseline corrected and normalized to Amide I peak) to
account and correct for noise, sloping baseline effects, differences in sample thickness or concentration.

2.2 Computational analysis

Multivariate analysis [principal component analysis-linear discriminant analysis (PCA-LDA)] were performed in MATLAB R2011b using an in-house developed IRTools toolbox (Trevisan et al., 2013). Results were visualized either as scores plots or cluster vectors plots, and the toolbox was set to identify the top six wavenumbers responsible for site differences. The mean band/peak area of the absorbance at specific regions was measured by calculating the integrated absorbance between the two wavenumbers (max-min) of the given region.

2.3 Statistical analysis

Variation in the tissues within the dataset was tested for statistical significance using Mann Whitney U-test, one-way analysis of variance (ANOVA) and Dunnett’s multiple comparison tests, where the $P$-value of less than 0.05 ($P <0.05$) was considered statistically significant.
3. Results

Figs. 1 and 5A show the mean spectral absorbance for the brain, kidney, heart, liver and gill tissues of African catfish. Tissues showed relatively marked differences within the lipid (~1740 cm\(^{-1}\)), protein (~1700 cm\(^{-1}\) - 1400 cm\(^{-1}\)) and DNA/RNA (~1399 cm\(^{-1}\) - 900 cm\(^{-1}\)) regions of the biofingerprint for all tissues, excluding gills where very subtle alterations to the DNA/RNA region were observed (Fig 5). Of note, most tissue (brain, liver, and gills) spectra sampled from Ifiayong seemed most congruent with tissues sampled from Gelegele.

Using the first LDA factor (LD1) in a one-dimensional (1-D) scores plot, the degree of variation in tissues between the sampling sites was visualized, following the application of PCA-LDA (Figs. 2 and 5B). The tissues obtained from Ikoro seemed most different from the corresponding tissues from other sites and produced a positive index along the LD1 space in most tissues [brain, liver (Fig 2) and gill (Fig 5)]. The variation between sites was tested using one-way analysis of variance and Dunnett’s multiple comparison test, comparing each site against the chosen reference site (Ifiayong). The variations between sites were significant with \(P < 0.0001\) in ANOVA and \(P < 0.01\) in Dunnett’s multiple comparison test, for all tissues.

Cluster vectors plots (Figs. 3 and 4) show wavenumbers responsible for the segregation in scores plot, and each distinguishing wavenumber corresponding to specific biochemical assignment (Table 1 and 2). Mean band areas were calculated for wavenumbers responsible for the differences in mean spectral absorbance. Peak centroids were observed to shift to higher or lower wavenumbers, with significant increase/decrease in the mean band areas (Table 3).
Fig 1. Mean spectra acquired from brain (A); kidney (B); heart (C); and, liver (D) tissues of the African catfish (*Heterobranchus bidorsalis*) obtained in March 2013 from sampling points Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) in the Niger Delta region. Ify (solid black lines), Iko (broken black lines) and Geg (grey lines). Spectra were cut between 1800 and 900 cm$^{-1}$, baseline corrected and normalized to the Amide I peak (1650 cm$^{-1}$).
Fig 2. Principal component analysis coupled with linear discriminant analysis (PCA-LDA) values in dataset acquired from Brain (A), Heart (B), Kidney (C) and Liver (D) tissues of the African Catfish (Heterobranchus bidorsalis) obtained in March 2013 from sampling points Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) within the Niger Delta region. Spectra were cut between 1800 and 900 cm⁻¹, baseline corrected and normalized to the Amide I peak (~1750 cm⁻¹). The Normalized spectra were mean centred before the application of PCA-LDA. As determined using one-way ANOVA, the PCA-LDA values in each class were statistically significant ($P < 0.0001$). Test classes (Geg, Iko) were significant ($P < 0.01$) when compared to reference class (Ify) using Dunnett's Multiple Comparison Test.
Fig. 3. Cluster vector plots acquired from Brain (A), Heart (B), Kidney (C) and Liver (D) tissues of the African Catfish (Heterobranchus bidorsalis) obtained in March 2013 from sampling points Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) within the Niger Delta region. Cluster vector plots were derived using Ifiayong as reference site. Ify (Solid black lines), Iko (Broken black lines) and Geg (Grey lines). Spectra were cut between 1800 and 900 cm⁻¹, baseline corrected, Vector normalized and mean centred before the application of multivariate analysis (PCA-LDA).
**Fig. 4.** Cluster vector plots acquired from Brain (A), Heart (B), Kidney (C) and Liver (D) tissues of the African Catfish (*Heterobranchus bidorsalis*) obtained in March 2013 from sampling points Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) within the Niger Delta region. Cluster vector plots were derived using Ikoro as reference site. Ify (Solid black lines), Iko (Broken black lines) and Geg (Grey lines). Spectra were cut between 1800 and 900 cm$^{-1}$, baseline corrected, Vector normalized and mean centred before the application of multivariate analysis (PCA-LDA).
**Fig. 5.** Mean spectra (A), Principal component coupled with Linear discriminant analysis (PCA-LDA) scores (B) and cluster vector (C & D) plots acquired from Gill tissues of the African Catfish (*Heterobranchus bidorsalis*) obtained in March 2013 from sampling points Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) within the Niger Delta region. Cluster vector plots were derived using Ifiayong (C) and Ikoro (D) as reference sites. Ify (Solid black lines), Iko (Broken lines) and Geg (Grey lines). Spectra were cut between 1800 and 900 cm\(^{-1}\), baseline corrected, Vector normalized and mean centred before the application of multivariate analysis (PCA-LDA). As determined using one-way ANOVA, the PCA-LDA values in each class were statistically significant (p < 0.0001). Test classes (Geg, Iko) were significant P < 0.05 when compared to reference class (Ify) using Dunnett's Multiple Comparison Test.
Table 1 Attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy distinguishing wavenumbers as shown in cluster vectors plots, and corresponding tentative chemical assignments: wavenumbers responsible for variance between tissue samples of the African Catfish (*Heterobranchus bidorsalis*) obtained in March 2013 from sampling points Ifiayong, Ikoro and Gelegele within the Niger Delta region. Distinguishing wavenumbers were derived following the application of multivariate analysis and using Ifiayong as reference site.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Site</th>
<th>Distinguishing wavenumbers (cm(^{-1}))</th>
<th>Tentative assignments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>Ikoro</td>
<td>1740</td>
<td>&gt;C=O ester stretching vibrations in triglycerides</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1620</td>
<td>Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1501</td>
<td>In-plane CH bending vibrations from phenyl rings</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1142</td>
<td>Phosphate and oligosaccharides; Oligosaccharide C – O bond in hydroxyl group that might interact with some other membrane components</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1030</td>
<td>Glycogen vibration; Collagen and phosphodiester groups of nucleic acids; stretching C – O ribose.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>961</td>
<td>C – O deoxyribose.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Gelegele</td>
<td>1747</td>
<td>C = O stretching vibration of Lipids, triglycerides, cholesterol esters</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1674</td>
<td>Anti-parallel β-sheet of Amide I, (\nu(C=C)) trans, lipids, fatty acids</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1616</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1474</td>
<td>Asymmetric CH(_3) bending of the methyl group of proteins</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1238</td>
<td>Asymmetric PO(_2) stretching</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1146</td>
<td>CO-O-C asymmetric stretching in glycogen and nucleic acids</td>
<td>3</td>
</tr>
<tr>
<td>Heart</td>
<td>Ikoro</td>
<td>1616</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1215</td>
<td>PO(_2) asymmetric (Phosphate I)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1153</td>
<td>Stretching vibrations of hydrogen-bonding C – OH groups</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1069</td>
<td>CO-O-C symmetric stretching of phospholipids and cholesterol esters</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1022</td>
<td>Glycogen</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>976</td>
<td>OCH(_3) (polysaccharides, pectin)</td>
<td>1</td>
</tr>
<tr>
<td>Gelegele</td>
<td>1701</td>
<td>Fatty acid esters</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td>---------------------------------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1620</td>
<td>Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1570</td>
<td>Amide II</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1524</td>
<td>Stretching C = N, C = C</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1466</td>
<td>CH2 scissoring mode of acyl chain of lipid, Cholesterol-methyl band</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>980</td>
<td>OCH3 (polysaccharides-cellulose)</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Ikoro</td>
<td>1717</td>
<td>Amide I (arises from C = O stretching vibration), C = O stretching vibration DNA and RNA.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1678</td>
<td>Stretching C = O vibrations that are H-bonded (changes in the C = O stretching vibrations could be connected with destruction of old H-bonds and creation of the new ones).</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1616</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1543</td>
<td>Amide II</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1470</td>
<td>CH2 bending of the methylene chains in lipids</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1204</td>
<td>Vibrational modes of collagen proteins-amide III C − O − C, C − O dominated by the ring vibrations of polysaccharides C − O − P, P − O − P collagen.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Gelegele</td>
<td>1717</td>
<td>Amide I (arises from C = O stretching vibration), C = O stretching vibration DNA and RNA.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1663</td>
<td>Amide I band, (\nu(C=C)) \textit{cis}, lipids, fatty acids</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1597</td>
<td>C = N, NH2 adenine</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1558</td>
<td>Ring base</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1524</td>
<td>Stretching C = N, C = C</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1466</td>
<td>CH2 scissoring mode of the acyl chain of lipid</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Wave Number (cm⁻¹)</td>
<td>Description</td>
<td>Frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1751</td>
<td>ν(C=C) lipids, fatty acids</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1443</td>
<td>δ(CH₂), lipids, fatty acids</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1366</td>
<td>Stretching C – O, deformation C – H, deformation N – H.</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1219</td>
<td>PO₂⁻ asymmetric vibrations of nucleic acids when it is highly hydrogen bonded, asymmetric hydrogen-bonded phosphate stretching mode</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1150</td>
<td>C – O stretching vibration, C – O stretching mode of the carbohydrates CH₈</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1061</td>
<td>CO – O – C symmetric stretching of phospholipids and cholesterol esters</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1732</td>
<td>C = O stretching in lipids</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1639</td>
<td>Amide I</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1443</td>
<td>δ(CH₂), lipids, fatty acids, Asymmetric CH₃ bending of the methyl groups of proteins</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1551</td>
<td>Amide II of proteins, N – H bending and C – N stretching</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1470</td>
<td>CH₂ bending of the methylene chains in lipids</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1173</td>
<td>C – O (stretching in malignant tissues), Non-hydrogen-bonded stretching mode of C – OH groups</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1740</td>
<td>&gt;C=O ester stretching vibrations in triglycerides</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1659</td>
<td>Amide I</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1612</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1562</td>
<td>CO₂⁻ asymmetric stretching possibly from glutamic acid</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1504</td>
<td>In-plane CH bending vibrations from the phenyl rings</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1458</td>
<td>CH₃ asymmetric bending</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1728</td>
<td>C = O band/ stretching</td>
<td>1, 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1651</td>
<td>80% C = O stretching; 10% C – N stretching; 10% N – H bending, Amide I absorption (predominantly the C = O stretching vibration of the amide C = O)</td>
<td>1, 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1589</td>
<td>Ring C – C stretch of phenyl, Ring stretching vibrations with little interaction with</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CH in-plane bending</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>-------------------------------------</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1462</td>
<td>CH$_2$ scissoring</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1408</td>
<td>CH$_3$ asymmetric deformation, (CH$_3$)$_3$N$^+$ symmetric bending</td>
<td>1, 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1234</td>
<td>Amide III/phosphate vibration of nucleic acids</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- v: stretching, δ: deformation
- References: (1) Movasaghi et al., 2008; (2) Stuart, 2005; (3) Obinaju et al., 2014
Table 2 Attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy distinguishing wavenumbers as shown in cluster vectors plots, and corresponding tentative chemical assignments: wavenumbers responsible for variance between tissue samples of the African Catfish (*Heterobranchus bidorsalis*) obtained in March 2013 from sampling points Ifiayong, Ikoro and Gelegele within the Niger Delta region. Distinguishing wavenumbers were derived following the application of multivariate analysis and using Ikoro as reference site.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Site</th>
<th>Distinguishing wavenumbers (cm⁻¹)</th>
<th>Tentative assignments</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>Ifiayong</td>
<td>1740</td>
<td>&gt;C=O ester stretching vibrations in triglycerides</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1620</td>
<td>Peak of nucleic acids due to the base carbonyl stretching and ring breathing mode</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1501</td>
<td>In-plane CH bending vibrations from phenyl rings</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1142</td>
<td>Phosphate and oligosaccharides; Oligosaccharide C – O bond in hydroxyl group that might interact with some other membrane components</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1030</td>
<td>Glycogen vibration; Collagen and phosphodiester groups of nucleic acids; stretching C – O ribose.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>961</td>
<td>C – O deoxyribose.</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Gelegele</td>
<td>1728</td>
<td>C = O band/ stretching</td>
<td>1, 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1674</td>
<td>Anti-parallel β-sheet of Amide I, ν(C=O) trans, lipids, fatty acids</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1234</td>
<td>Amide III/phosphate vibration of nucleic acids</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1146</td>
<td>CO-O-C asymmetric stretching in glycogen and nucleic acids</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1034</td>
<td>Collagen</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>964</td>
<td>C – C, C – O deoxyribose</td>
<td>1, 3</td>
</tr>
<tr>
<td>Heart</td>
<td>Ifiayong</td>
<td>1616</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1215</td>
<td>PO₂⁻ asymmetric (Phosphate I)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1153</td>
<td>Stretching vibrations of hydrogen-bonding C – OH groups</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1069</td>
<td>CO-O-C symmetric stretching of phospholipids and cholesterol esters</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1022</td>
<td>Glycogen</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>976</td>
<td>OCH₃ (polysaccharides, pectin )</td>
<td>1</td>
</tr>
<tr>
<td>Kidney</td>
<td>1717</td>
<td>Amide I (arises from C = O stretching vibration), C = O stretching vibration DNA and RNA.</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1678</td>
<td>Stretching C = O vibrations that are H-bonded (changes in the C = O stretching vibrations could be connected with destruction of old H-bonds and creation of the new ones).</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1616</td>
<td>Amide I (Carbonyl stretching vibrations in side chains of amino acids)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1543</td>
<td>Amide II</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1470</td>
<td>CH₂ bending of the methylene chains in lipids</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

| Gelegele | 1693 | A high frequency vibration of an antiparallel β-sheet of amide I (the amide I band is due to in-plane stretching of the C = O band weakly coupled to stretching of the C – N and in-plane bending of the N – H bond | 1 |
|         | 1647 | Amide I (α-helix) | 2 |
|         | 1612 | Amide I (Carbonyl stretching vibrations in side chains of amino acids) | 3 |
|         | 1543 | Amide II | 1 |
|         | 1169 | \(v_{at}\) CO – O – C | 1 |
|         | 1061 | CO – O – C symmetric stretching of phospholipids and cholesterol esters | 3 |
| Liver   | 1751 | $\nu$(C=C) lipids, fatty acids | 3 |
|         | 1443 | $\delta$(CH<sub>2</sub>), lipids, fatty acids | 1 |
|         | 1219 | PO<sub>2</sub> asymmetric vibrations of nucleic acids when it is highly hydrogen bonded, asymmetric hydrogen-bonded phosphate stretching mode | 1 |
|         | 1150 | C – O stretching vibration, C – O stretching mode of the carbohydrates CH<sub>3</sub> | 1 |
|         | 1061 | CO – O – C symmetric stretching of phospholipids and cholesterol esters | 3 |
| Gelegele| 1732 | C = O stretching in lipids | 1 |
|         | 1639 | Amide I | 1 |
|         | 1508 | In-plane CH bending vibration from the phenyl rings | 1 |
|         | 1470 | CH<sub>2</sub> bending of the methylene chains in lipids | 1 |
|         | 1153 | Stretching vibrations of hydrogen-bonding C – OH groups | 1 |
|         | 1057 | Stretching C – O deoxyribose | 1 |
| Gills   | 1740 | >C=O ester stretching vibrations in triglycerides | 3 |
|         | 1659 | Amide I | 1 |
|         | 1612 | Amide I (Carbonyl stretching vibrations in side chains of amino acids) | 3 |
|         | 1562 | CO<sub>2</sub> asymmetric stretching possibly from glutamic acid | 2 |
|         | 1504 | In-plane CH bending vibrations from the phenyl rings | 1 |
|         | 1458 | CH<sub>3</sub> asymmetric bending | 2 |
| Gelegele| 1663 | Amide I band, $\nu$(C=C) <i>cis</i>, lipids, fatty acids | 3 |
|         | 1589 | Ring C – C stretch of phenyl | 1 |
|         | 1543 | Amide II (protein N-H bend, C-N stretch) in α-helices | 3 |
|         | 1462 | CH<sub>2</sub> scissoring | 2 |
|         | 1178 | C-O asymmetric stretching of glycogen | 3 |
|         | 968  | C – C , C – O deoxyribose, DNA | 1 |

- $\nu$: stretching, $\delta$: deformation, References: (1) Movasaghi et al., 2008; (2) Stuart, 2005; (3) Obinaju et al., 201
Table 3 Mid-infrared absorbance peak/band areas calculated for each detected centroid in mean bio-fingerprint spectra of the African catfish (*Heterobranchus bidorsalis*) tissues samples collected from Ifiayong (Ify), Ikoro (Iko) and Gelegele (Geg) in March 2013, along Ovia River (Nigeria). Peak/band areas are representative of concentrations of biomolecule within the tissue. Peak areas were statistically tested using Ifiayong as the control.

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Band Assignment</th>
<th>Peak Centroid</th>
<th>Peak Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IFY</td>
<td>IKO</td>
</tr>
<tr>
<td>Brain</td>
<td>Lipids</td>
<td>1740</td>
<td>Not observed</td>
</tr>
<tr>
<td></td>
<td>Amide I</td>
<td>1643</td>
<td>1632</td>
</tr>
<tr>
<td></td>
<td>Amide II</td>
<td>1535</td>
<td>1528</td>
</tr>
<tr>
<td></td>
<td>COO symmetric stretch in fatty and amino acids, CH&lt;sub&gt;2&lt;/sub&gt; and CH&lt;sub&gt;3&lt;/sub&gt; deformation in lipids and protein</td>
<td>Not observed</td>
<td>1393</td>
</tr>
<tr>
<td></td>
<td>Amide III/Phosphate Vibrations of Nucleic Acids</td>
<td>1231</td>
<td>1231</td>
</tr>
<tr>
<td></td>
<td>Stretching / Vibrations of C – O ribose and Deoxyribose</td>
<td>1061</td>
<td>1065</td>
</tr>
<tr>
<td>Liver</td>
<td>Lipids</td>
<td>Not observed</td>
<td>1740</td>
</tr>
<tr>
<td></td>
<td>Amide I</td>
<td>1628</td>
<td>1628</td>
</tr>
<tr>
<td></td>
<td>Amide II</td>
<td>1520</td>
<td>1528</td>
</tr>
<tr>
<td></td>
<td>COO symmetric stretch in fatty and amino acids, CH&lt;sub&gt;2&lt;/sub&gt; and CH&lt;sub&gt;3&lt;/sub&gt; deformation in lipids and protein</td>
<td>Not observed</td>
<td>1389</td>
</tr>
<tr>
<td></td>
<td>Amide III/Phosphate Vibrations of Nucleic Acids</td>
<td>1234</td>
<td>1234</td>
</tr>
<tr>
<td></td>
<td>C – O Stretching</td>
<td>Not observed</td>
<td>1173</td>
</tr>
<tr>
<td></td>
<td>Stretching / Vibrations of C – O ribose and Deoxyribose</td>
<td>1065</td>
<td>Not observed</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------------------------------------</td>
<td>------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>Heart</strong></td>
<td>Amide I</td>
<td>1628</td>
<td>1632</td>
</tr>
<tr>
<td></td>
<td>Amide II</td>
<td>1524</td>
<td>1524</td>
</tr>
<tr>
<td></td>
<td>COO⁻ symmetric stretch in fatty and amino acids, CH₂ and CH₃ deformation in lipids and protein</td>
<td>1393</td>
<td>1393</td>
</tr>
<tr>
<td></td>
<td>Amide III/Phosphate Vibrations of Nucleic Acids</td>
<td>1234</td>
<td>1238</td>
</tr>
<tr>
<td></td>
<td>Glycogen</td>
<td>Not observed</td>
<td>1026</td>
</tr>
<tr>
<td><strong>Kidney</strong></td>
<td>Amide I</td>
<td>1628</td>
<td>1628</td>
</tr>
<tr>
<td></td>
<td>Amide II</td>
<td>1528</td>
<td>1520</td>
</tr>
<tr>
<td></td>
<td>COO⁻ symmetric stretch in fatty and amino acids, CH₂ and CH₃ deformation in lipids and protein</td>
<td>Not observed</td>
<td>1389</td>
</tr>
<tr>
<td></td>
<td>Amide III/Phosphate Vibrations of Nucleic Acids</td>
<td>1234</td>
<td>1234</td>
</tr>
<tr>
<td></td>
<td>C – O Stretching</td>
<td>1169</td>
<td>Not observed</td>
</tr>
<tr>
<td><strong>Gills</strong></td>
<td>Amide I</td>
<td>1628</td>
<td>1628</td>
</tr>
<tr>
<td></td>
<td>Amide II</td>
<td>1520</td>
<td>1528</td>
</tr>
<tr>
<td></td>
<td>COO⁻ symmetric stretch in fatty and amino acids, CH₂ and CH₃ deformation in lipids and protein</td>
<td>1389</td>
<td>1393</td>
</tr>
<tr>
<td></td>
<td>Amide III/Phosphate Vibrations of Nucleic Acids</td>
<td>1234</td>
<td>1234</td>
</tr>
</tbody>
</table>

***p < 0.0001, **p > 0.001, * p > 0.05; ¶ not observed: peak centroid not detected.
4. Discussion

The mean absorbance spectra present an overview of the changes occurring in an interrogated sample. These changes are recognised as intensity variations and shifts in peak centroids. The intensity of absorption bands in the IR spectra of biological samples is regarded as directly proportional to the concentration of the particular biomolecule (Cakmak et al., 2006; Severcan et al., 2005). The position of peak centroids such as Amide I (~1650 cm\(^{-1}\)) is considered to be sensitive to protein conformation (Obinaju et al., 2015; Palaniappan and Pramod, 2011) and thus, peak shifts are regarded as alterations to either total structure of the molecule or specific peptides (Holman et al., 2000; Obinaju et al., 2015). Within the scores plots, nearness of the individual sites to each other suggests a similarity of chemical structures and distance suggests dissimilarity. A positive index in LD1 suggests an increase in the total biomolecules present in the sample compared (Llabjani et al., 2014).

Based on the physical characteristics of the various sites, initial cluster vectors plots (Fig. 3) compared tissues sampled from Gelegele and Ikoro to tissues from Ifiayong. Spectral differences were observed in cluster vectors plots associated with lipid/protein (~1750 cm\(^{-1}\) - 1400 cm\(^{-1}\)) and DNA/RNA (~1300 cm\(^{-1}\) - 900 cm\(^{-1}\)) regions of the biofingerprint of all tissue types. These included alterations to C=O stretching vibrations in triglycerides and cholesterol esters, carbonyl stretching vibrations in the side chains of amino acids, asymmetric stretching of phosphate and carbon-to-oxygen vibrations in deoxyribose. However, from observations in mean absorbance spectra, and the differences between the sites in 1-D scores plots, a second cluster vectors plot (Fig 4) compared tissues sampled from Gelegele and Ifiayong to tissues from Ikoro. These plots revealed a similar pattern of spectral alterations across the IR fingerprint of the tissues from Gelegele and Ifiayong, with slight variations to
intensity in tissues obtained from Gelegele. Spectral alterations in cluster vectors plots closely matched the observations in the mean absorbance spectra for tissues sampled from Gelegele and Ifiayong, compared to Ikoro. This observation suggests the possibility of environmental contamination occurring at Ifiayong. More importantly, it suggests that the alterations in tissues from Ifiayong were induced by contaminants possibly similar to those present at Gelegele (e.g., PAHs). Tissues from Ifiayong and Ikoro were consistently distinguished by the same wavenumbers in cluster vectors plots. In contrast, distinguishing wavenumbers varied for tissue samples from Gelegele, when compared to tissues from either Ifiayong or Ikoro in the cluster vectors plots.

From previous observations (Obinaju et al., 2014; Obinaju et al., 2015) and the site characteristics of Gelegele, alterations observed within the DNA/RNA region of the biofingerprint are possibly PAH-mediated toxicity as a result of metabolite binding to macromolecules such as DNA. With no documentation of industrial activity in close proximity to Ifiayong, we hypothesize that the possibility of contamination occurring at Ifiayong may be due to 1) the sloppy topography of the community which predisposes it to erosions (Umoh, 2013) and the possible deposition of contaminants from urban runoff; or, 2) the location of the community along a river path, e.g., downstream, making it a recipient of environmental contaminants based on the direction of river flow. It is also possible that fish samples obtained at Ifiayong were pre-exposed and migrated from contaminated regions as a measure of adaptation to environmental change (Alemanni et al., 2003).
5. Conclusion

FTIR spectroscopy monitors the vibrational modes of functional groups within biomolecules and enables a correlation between chemical information and histological structures, where shifts in peak positions, changes in bandwidths, intensities and band area values of the IR bands are used to obtain valuable structural and functional information about the system of interest (Cakmak et al., 2006). We have previously shown the potential of ATR-FTIR spectroscopy to detect sub-lethal real-time exposure to environmental contaminants in sentinel organisms (Obinaju et al., 2014). More recently, we have shown that the patterns of spectral alterations in the IR spectra signature can be related to the presence of specific environmental contaminants (Obinaju et al., 2015).

This study aimed to understand the changes occurring in fish tissues as a result of PAH exposure at the sites Gelegele and Ikoro, by comparing the IR spectra of the tissues to those obtained from a relatively pristine site (Ifiayong). This aim was modified based on observations in the mean absorbance spectra of the tissues interrogated, and Ifiayong was classed as a ‘blind’ site (i.e., site with no prior information of contaminant levels or recorded history of contamination). Multivariate analysis revealed that PAH contamination could be occurring at Ifiayong. Although contaminant bioaccumulation in the tissues of most sentinels may not pose a direct health risk to the human population, monitoring bioaccumulation in these tissues are important to assessing environmental health of most ecosystems. Our results present the possibility of identifying contaminants and contaminant-induced changes in organisms of unknown origins, based on existing knowledge of IR spectra acquired from organisms with exposure to known compounds. Our results provide evidence supporting the use of ATR-FTIR spectroscopy in biomonitoring in sentinel organisms.
ACKNOWLEDGEMENTS: BEO is a Faculty for the Future Fellow of the Schlumberger Foundation, an independent non-profit entity that supports science and technology education. Such Fellowships support female academics from developing and emerging countries for advanced graduate study.
References

Ahmad, M., Mirza, B., Hussain, M., Hanif, M., Ali, S., Walsh, M., Martin, F., 2008. ATR-FTIR spectroscopy detects alterations induced by organotin(IV) carboxylates in MCF-7 cells at sub-cytotoxic-genotoxic concentrations. PMC Biophysics 1, 3.


Chapter 7

Discussion

Ecotoxicology primarily aims to identify patterns describing population and community responses to contaminants as it is now generally acceptable that chemical characterization of a compound by itself does not provide specific biological information about potential hazards to organisms (Chapman, 2007). The ability of researchers to predict these responses is generally greatest for communities that change consistently in response to a specific contaminant or class of contaminants, thereby providing a direct path to extrapolation, hypothesis testing, and scientific inference (Clements et al., 2012). Validated biomarkers are important tools for ecotoxicologists where they are early warning signals to pre-empt dire environmental consequences (Eason and O'Halloran, 2002). As sensitive and ecologically relevant measures of environmental conditions, bioindicators can be used to assess the health of aquatic ecosystems which may be compromised by a variety of environmental stressors such as contaminants, sediments, nutrients, and varying temperature, salinity, and hydrologic regimes (Adams and Greeley, 2000).

The bioindicator approach uses responses of key (sentinel) aquatic organisms both as integrators of stress effects and as sensitive response (early-warning) indicators of environmental health (Phillips and Rainbow, 1993). It involves measuring a suite of selected biological and ecological responses at several levels of biological organization from the biomolecular and biochemical to the community levels (Bodin et al., 2011; Brown et al., 1973; 1977; Malins, et al., 1984). In the context of environmental monitoring studies bioindicators describe organisms (or
parts of organisms or communities of organisms) that provide information on quality of the environment (or a part of the environment) (Markert et al., 2003).

Environmental contamination occurs or is more frequently encountered as complex mixtures and the behaviour of chemicals in a mixture rarely corresponds to that predicted from data on the pure compounds (Llabjani et al., 2010). Relating low environmental exposures to actual effects in organisms is often difficult and requires lab-based dose response type assays often involving sophisticated techniques and protocols requiring expensive kits. This understanding creates the need for simple, cost effective yet highly sensitive techniques and robust protocols (Baker et al., 2014) that are applicable to environmental biomonitoring and able to detect real-time contaminant exposure in these organisms even at very low doses.

Molecular bonds with an electric dipole moment that can change by atomic displacement owing to natural vibrations are IR active and the various vibrational modes can be quantitatively measured using vibrational spectroscopy (Griffiths and De Haseth, 2007). Vibrational spectroscopic techniques especially FTIR spectroscopy techniques, have become potential tools for non-invasive optical tissue diagnosis and have been applied to study a wide variety of pathologic states especially in clinical fields. A variety of the techniques have been applied to study chemical induced toxicity in sentinels (Cakmak et al., 2006; Llabjani et al., 2012; Severcan et al., 2005; Toyran et al., 2004).

The application of spectroscopy to biological samples generates information, based on the vibrational modes of functional groups within biomolecules and enables the correlation between chemical information and histological structures. The IR absorbance spectra for any biological sample contains important information
regarding the structure and conformation of molecules within the sample. The IR spectra with wavenumbers from 1800 – 900 cm\(^{-1}\) is regarded as the biochemical cell fingerprint of a biological sample, with each wavenumber corresponding to a specific biological molecule (Baker et. al., 2014).

Each acquired spectrum consists of hundreds of variables and therefore requires the use of computational analyses to extract the required information. There are several computational approaches that are potentially applicable to spectral datasets. However, the most commonly used include PCA, LDA, PLS and PCA-LDA. These techniques allow for the derivation of possible biomarkers (Fig. 8) which distinguish between sample types and treatment conditions.
7.1 Understanding chemical induced changes in cells and potentially subcellular components of cells

Obinaju et. al. (2015a) illustrates the potential application of ATR-FTIR spectroscopy to study, signature and understand the changes occurring in subcellular components of the cell as a result of exposure to potential mutagens. Chemically induced changes to biological molecules, particularly genetic molecules such as DNA are known to predispose organisms to pathologic disease conditions e.g. cancers (Malins et. al. 2006). The ability of infrared spectroscopy techniques to distinguish intact cell and subcell i.e. nucleus has been previously shown (Pang et al., 2012; Pijanka et al., 2009). The ability to detect slight changes in the IR spectra of samples at wavenumbers representative of biomolecules, e.g., symmetric (1088 cm\(^{-1}\)) and asymmetric (1234 cm\(^{-1}\)) PO\(_2\) bands, which typically can be associated with nucleic acids, are significant for understanding the differences between normal and malignant conditions (Lasch et al., 2002). Study presented in chapter three employed the use of ATR-FTIR spectroscopy techniques to detect and understand the response of cultured cell populations of the human mammary carcinoma (MCF-7), to very low dose exposures of B[a]P. The project observed the effect of exposure in both intact cells and isolated nuclei of cells in the G\(_0\)/G\(_1\) phase of the cell cycle.

*In vitro* experiments using very low doses (< 10\(^{-4}\) M) are more realistic models to concentrations obtainable in real world scenarios. Isolated nuclei of cells treated with the 10\(^{-6}\) M of B[a]P were observed to be slightly deformed and observed effects were attributed to the possible biotransformation/activation of B[a]P as well as its interaction with DNA molecules as previously documented by Malins et.al. (2006). B[a]P induced dose dependent alterations in cell populations in G\(_0\)/G\(_1\) phase and a bimodal response in S-phase. All observed responses were statistically significant (\(p <\)
0.01) when compared to control cell populations. The ability to extract potential biomarkers as wavenumbers using multivariate analysis was important to distinguishing each treatment condition from the control.

Wavenumbers such as 1740 cm\(^{-1}\) in isolated nuclei of the treated cell population suggested that B\([a]\)P induced changes to nuclear lipids of isolated nuclei of the cells (Balasubramanian et al., 2007). Nuclear lipids are believed to play a role in the proliferation, differentiation and apoptotic processes in the cell cycle and changes to nuclear lipids may be one of the mechanisms by which high doses of B\([a]\)P exerts cytotoxicity/cell death (Ledeen and Wu, 2006; Lin and Yang, 2008). C = O guanine deformation was also identified in isolated nuclei of G\(_0\)/G\(_1\) cells. This could be regarded a potential biomarker for B\([a]\)P exposure in cells as B\([a]\)P metabolites are known to possess a high affinity for guanine and covalently alter the structure of this molecule. Thus, understanding chemical induced changes in the various phases of the cell cycle is important for extrapolations to the changes occurring in quiescent and rapidly dividing cell populations e.g. neurons and epithelial cell populations.

7.2 IR spectroscopy to study real-time exposure in organisms

The ability to identify changes to cell and sub-cell using ATR-FTIR spectroscopy as documented in Obinaju et.al. (2015a) as well as previous applications particularly in clinical diagnostics to differentiate between normal and diseased tissues (Wong et al., 1991; Fung et al., 1996) and cell types (German et al., 2006) has informed the application of ATR-FTIR in environmental toxicology, particularly to track and understand effects of environmental compounds in real-time.

Impact and consequences of environmental contamination in Nigeria on the resident population is fairly known due to the scanty literature on risk assessment
studies, (eco) toxicological studies or epidemiology. Although there are studies documenting PAH contamination within the Niger Delta, these studies document mainly concentrations of PAHs in particle and sediment phase of the water column (Ana et al., 2009; Essien et al., 2011; Okafor and Opuene, 2007) and no study documenting concentrations in dissolved phase.

Compared to studies monitoring the bioaccumulation of heavy and trace metals in sentinels, very few studies (Anyakora and Coker, 2007; Benson et al., 2008; Eduok et al., 2010) have attempted to monitor PAHs in sentinels within the Niger Delta region. Furthermore, these studies only document the concentrations of PAHs within sentinel tissues with no specific documentation of possible exposure effects, e.g., DNA damage in the observed sentinels as a response to PAH exposure.

Taking these knowledge gap into account, real-time exposure effects in the tissues of the African catfish (*Heterobranchus bidorsalis*) and the water spinach (*Ipomea aquatica*) were observed and the first evidence of concentration levels for PAHs in the dissolved phase of a river in the Niger Delta presented (Obinaju et al. 2015b; Obinaju et al. 2014). The studies employed ATR-FTIR spectroscopy combined with multivariate analysis to signature effect and extract the possible biomarkers in the various tissues. Effect of contaminant exposure was observed in fish tissues sourced along the Ovia River which plays host to petroleum exploration activities and as such is thought to be contaminated with PAHs.
7.3 Differentiating between samples of varying exposure conditions using ATR-FTIR

Obinaju et. al. (2014) documents that in both sampling seasons (dry and rainy), ATR-FTIR mean spectra was able to detect subtle but clear variations between sampling sites. The application of multivariate analysis (PCA-LDA) to the bio-fingerprint region derived scores plots and corresponding cluster vector plots for the various tissues interrogated and multivariate analysis was able to make clear distinctions between the tissues based on seasonal variation.

Increased concentration of ester groups belonging to triglycerides within exposed tissues, particularly in the liver was observed by the increased intensity of lipid peaks in liver tissues samples. This observation was more remarked in liver tissues obtained in the dry season and was possibly due to an increased concentration of PAHs available for absorption due to reduced river current. The increase absorption of the available PAH compounds and subsequent metabolism in the liver may possibly have induced changes in lipid metabolism, resulting in the accumulation of lipids within the liver tissues and possibly the onset of fatty liver. Metabolic activities of enzymes in liver tissues resulted in changes to intensity and area of glycogen band (1177 cm\(^{-1}\)), a possible measure of oxidative stress in liver tissues. These findings were in agreement with chemical induced changes previously documented to have occurred in rainbow trout following exposure to 17β-estradiol (Cakmak et. al. 2006).

In plant leaves, senescence and environmental stresses are accompanied by changes in the cell surfaces and pigments which determine the optical characteristics of plant tissues (Ribeiro da Luz, 2006). Changes may arise from age of leaves and/or season variation. Obinaju et.al. (2014) was able to correlate variation in leaf
pigmentation with the characteristics of the sampling site after sample fixation. These variations were very well explained by the intensity variations to absorption band for cellulose (1030 cm\(^{-1}\)) in the IR spectra of the individual biofingerprint of the samples. Changes to absorption bands between absorption bands between 1650 cm\(^{-1}\) and 1500 cm\(^{-1}\) were considered physiological stress markers in plant leaves as previously suggested in Ivanova and Singh (2003). IR spectroscopy’s ability to detect the subtle changes in structure and dynamics of biological molecules in sentinel organisms exposed to varying degrees of environmental contamination presents the possibility for real-time evaluation of contaminant toxicity and could be important to effective environmental monitoring.

7.4 PAH concentration relation to IR biofingerprint

There is no literature documenting the concentrations of PAHs in the dissolved phase of the water column of any of the aquatic environments within the Niger Delta. Hence, Obinaju et. al. (2015b) measured PAH concentrations in the dissolved phase of the water column of the Ovia River and report concentrations ranging from 0.1 - 1055.6 ng.L\(^{-1}\) for both dry and rainy seasons. Detected concentrations decreased with increased distance from the known pollution source. Seasonal influence (e.g., heavy rainfall), which causes a change in river current and influences the dispersal of compounds was suspected as a possible factor responsible for the increased concentrations of compounds such as 2,6-dimethylnaphthalene, 2,3,6-trimethylnaphthalene, phenanthrene, fluorene, anthracene, 1-methylphenanthrene, fluoroanthen, pyrene at relatively pristine sites in the rainy season.

The concentrations of PAH detected in Ovia river were either similar or much higher than concentrations detected in the dissolved phases of the water column in
comparable locations such as Pearl River and Macao harbour of the Pearl River Delta in South China, the Seine Estuary in France and the Southern Chesapeake Bay, USA (Cailleaud et al., 2007; Gustafson and Dickhut, 1997; Luo et al., 2004).

In order to exhibit carcinogenicity, chemical carcinogens require metabolic activation by cytochrome P450 (CYP) enzymes to more reactive metabolites. The liver of most organisms is the primary site of metabolic activation. Thus, a comparison of mean absorbance spectra and cluster vector plots for B[a]P treated MCF-7 cells to those of H. bidorsalis liver tissues was used to determine if alterations observed in fish tissues were PAH-induced. The mean absorbance spectra showed similar patterns of alterations in B[a]P treated MCF-7 cells and fish liver tissue. These alterations were observed in Amide I (1650 cm\(^{-1}\)), Amide II (1550 cm\(^{-1}\)) and Amide III/asymmetric phosphate stretching vibrations of nucleic acids (1234 cm\(^{-1}\)) regions of the biofingerprint. Peak areas calculated for each detected centroid in the mean spectra biofingerprint of interrogated H. bidorsalis tissues showed statistically significant decrease and increase in most band areas across all tissues interrogated and centroid positions of most peaks were observed to shift to lower or higher values in tissues interrogated. Based on the known source of contaminants, the results suggested that alterations in the biofingerprint spectra of tissues particularly to regions representative of proteins (Amide I and II) were PAH induced, a possible consequence of PAH-induced protein oxidation in the tissues and possibly mediated by ROS.
Fig. 8 Summary of the various mechanisms of genotoxic and non-genotoxic carcinogenic environmental compounds, including polycyclic aromatic hydrocarbons. Compounds directly or indirectly affect the regulation and expression of genes involved in cell cycle control, DNA repair, cell differentiation or apoptosis (cell death). DNA damage or altered signal transduction processes may lead to the loss of growth control and genome instability, the major hallmarks of cancer.
7.5 Identification of contaminant exposure in tissues based on IR spectra.

Both Obinaju et. al., (2014) and Obinaju et. al., (2015b) all showed that IR spectroscopy, particularly ATR-FTIR spectroscopy was able to differentiate between real-time exposure effects in tissues from sites with varying degrees of environmental contamination in both animal and plant tissues based on spectral variation. Thus, comparing tissues of the African Catfish (*Heterobrachus bidorsalis*) from sites with a documented history of polycyclic aromatic hydrocarbon (PAH) contamination from petroleum exploration activities, to tissues from a relatively pristine site with no documented history of contamination with PAHs and no industrial activity, Obinaju and Martin in chapter 6 aimed to show that it was potentially possible to determine the nature of compounds present at a site, based on spectra similarity or difference.

The mean absorbance spectra for tissues showed relatively marked alterations within the DNA/RNA region, as well as alterations to lipid and protein regions of the biofingerprint for all tissues excluding gills tissues. These alterations were reflected as increased or reduced intensities in the affected spectral region of the tissues. As have been previously suggested, the intensity and/or the area of the absorption bands is directly related to the concentration of the molecules (Cakmak et al., 2006; Severcan et al., 2005; Toyran et al., 2004).

However, the mean absorbance for most tissues (brain, liver, and gills) sampled from the relatively pristine site were more similar to the mean absorbance for tissues sampled from the site regarded as most contaminated based on detected PAH concentrations in previous projects (Obinaju et. al., 2015b) and the industrial activity at site. Spectral differences were observed in cluster vector plots associated with lipid/protein (~1750 cm\(^{-1}\) - 1400 cm\(^{-1}\)) and DNA/RNA (~1300 cm\(^{-1}\) - 900 cm\(^{-1}\)) regions of the biofingerprint of all tissue types. These included alterations to C = O
stretching vibrations in triglycerides and cholesterol esters, carbonyl stretching vibrations in the side chains of amino acids, asymmetric stretching of phosphate and carbon to oxygen vibrations in deoxyribose. Cluster vector plots which compared tissue samples from both pristine and most contaminated sites to tissues obtained from a site of known lesser contamination revealed a similar pattern of spectral alterations across the IR fingerprint of the tissues from pristine and most contaminated site, with slight variations to intensity in tissues obtained from the most contamination.

Based on the known characteristics of the pristine site, the possibility of contamination occurring may be due to the sloppy topography of the site which predisposes it to erosions, or the location of the site along a river path e.g. downstream, making it a recipient of environmental contaminants based on the direction of river flow. There is also the possibility that fish samples obtained were pre- exposed and had migrated from contaminated regions as a measure of adaptation to environmental change. Whichever the case, the ability of IR spectroscopy and multivariate analysis to discriminate these exposures in the tissues illustrates the immense potential of vibrational spectroscopy especially the techniques involved in IR spectroscopy, in environmental biomonitoring.
Conclusion

IR spectroscopy monitors the vibrational modes of functional groups within biomolecules and enables a correlation between chemical information and histological structures. The bands within an IR spectrum are used to obtain valuable structural and functional information about the system of interest. The various methods employed in IR spectroscopy are specific and sensitive to changes such as shifts in peak positions, changes in bandwidths, intensities and band area values within the biochemical constituents of cells and tissues at certain wavelengths. These changes can be correlated to the exposure to specific chemicals and potential biomarkers can be extracted based on the variation to specific IR bands.

Every study documented herein was designed to test the following hypotheses

1) the technique ATR-FTIR is sensitive to and able to detect minimal cellular changes occurring in tissues exposed to potential mutagens. 2) ATR-FTIR can extract potential biomarkers to signature chemical induced changes in tissues. Studies were designed to investigate the effects of low dose PAH exposure in intact cells and isolated nuclei as a baseline for real-time environmental exposure scenarios, to detect seasonal variations in exposure and real-time exposure effect in fish tissue and plant leaves, to investigate the correlation of PAH concentrations detected in dissolved phase of water column to the alterations in fish tissues and to explore the potential identification of biomarkers of PAH exposure in fish tissues using ATR-FTIR spectroscopy.

Results show that ATR-FTIR was able to detect changes in cell populations exposed to very low doses of B[a]P including the treatment-induced changes in the nuclei of the exposed population. The study investigating the possible application of ATR-FTIR to observe contaminant effect in sentinels showed that the technique was able to
detect real-time sub-lethal exposures in the tissues of organisms studied and the variations to spectral regions of the biofingerprint were reflective of the concentrations of known and detected contaminants at the various sampling sites. ATR-FTIR was able to extract of wavenumbers (potential biomarkers) which are representative of biomolecules that were possibly chemically altered. ATR-FTIR was able to detect exposure in tissues obtained without prior knowledge of contaminant exposure or the nature of the possible contaminating compound and presented the possibility of describing the nature of compound present.

Retrospectively, a single year of sampling may not be the best possible representation of long term contaminant effects in organisms. It would have been interesting to measure a variety of potential contaminants, including metals in the different phases of the water column. The sampling size could definitely be increased and expanded to account for changes occurring as a result of sex and lifecycle in the organisms studied. These are areas that could be improved in subsequent experimental designs given that projects discussed herein are the first of its kind to be conducted within the region.

That said and finally, vibrational spectroscopy is becoming a valuable tool to understand molecular pathways and a potential tool for clinical diagnosis. Its application in other areas such as agriculture and environmental monitoring hold equal promise. Technological advancement and the optimisation of protocols hold these promises for even more applications of vibrational spectroscopy in both biological and other scientific studies.
Appendix I

Using Fourier transform IR spectroscopy to analyze biological materials


Contribution:

- I wrote the materials section
- As a group, I contributed to other sections of the manuscript

.................................................. ..................................................
Blessing E. Obinaju Prof Francis L. Martin
Appendix II

In vitro protective effects of quercetin in MCF-7 cells despite an underlying toxicity profile

Blessing E. Obinaju and Francis L. Martin

Mutagenesis 27 (2012) 789-816

Contribution:

- I collected and presented the data at the 35th Annual Meeting of the United Kingdom Environmental Mutagen Society.

Blessing E. Obinaju

Prof Francis L. Martin
Appendix III

Attenuated total reflection Fourier-transform infrared spectroscopy detects real-time polyaromatic hydrocarbon toxicity in fish tissues.

Blessing E. Obinaju and Francis L. Martin

Mutagenesis 29 (2014) 79-96

Contribution:

- I collected and presented the data at the 36th Annual Meeting of the United Kingdom Environmental Mutagen Society.
Appendix IV

Alterations in infrared spectral signature of *Heterobrachus bidorsalis* reflects polyaromatic hydrocarbon concentrations in Ovia River, Nigeria.

**Blessing E. Obinaju** and Francis L. Martin

*Mutagenesis 29* (2014) 497-559

Contribution:

- I collected and presented the data at the 43rd Annual Meeting of the European Environmental Mutagen Society.

........................................

Blessing E. Obinaju                                      Prof Francis L. Martin
Bibliography


Baker, M.J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H.J., Dorling, K.M., Fielden, P.R., Fogarty, S.W., Fullwood, N.J., Heys, K.A., Hughes, C., Lasch,


Bi, X., Thomas, G.O., Jones, K.C., Qu, W., Sheng, G., Martin, F.L., Fu, J., 2007. Exposure of electronics dismantling workers to polybrominated diphenyl ethers,
polychlorinated biphenyls, and organochlorine pesticides in South China. Environmental Science and Technology 41, 5647-5653.


pollution and nonmalignant respiratory mortality in 16 cohorts within the ESCAPE project. American Journal of Respiratory and Critical Care Medicine 189, 684-696.


subjects exposed to urban air pollution and environmental tobacco smoke. Toxicology Letters 149, 269-280.


Glotfelty, D., Williams, G., Freeman, H., Leech, M., 1990. Regional atmospheric transport and deposition of pesticides in Maryland. Lewis Publishers: Chelsea, MI.


for determining freely dissolved water concentrations. Environmental Toxicology and Chemistry 16, 452-461.


Hallgren, S., and Darnerud, P. O., 2002. Polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs) and chlorinated paraffins (CPs) in rats—testing interactions and mechanisms for thyroid hormone effects. Toxicology, 177, 227-243.


formation in human skin fibroblasts: preventive effects of anti-oxidant agents.
Environmental Toxicology and Pharmacology 12, 37-42.


polycyclic aromatic hydrocarbon photoproducts. Environmental Toxicology and Chemistry 25, 1079-1087.


Mackay, D., Shiu, W.-Y., Ma, K.-C., Lee, S.C., 2006. Handbook of physical-chemical properties and environmental fate for organic chemicals. CRC/Taylor and Francis, Boca Raton, FL.


Martin, F.L., 2007. Complex mixtures that may contain mutagenic and/or genotoxic components: a need to assess in vivo target-site effect(s) associated with in vitro-positive(s). Chemosphere 69, 841-848.


Tintos, A., Gesto, M., Míguez, J.M., Soengas, J.L., 2007. Naphthalene treatment alters liver intermediary metabolism and levels of steroid hormones in plasma of rainbow trout (Oncorhynchus mykiss). Ecotoxicology and Environmental Safety 66, 139-147.


discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice. Metabolomics 1, 65-73.


