An environmental metabolomics investigation of the effects of chiral pharmaceuticals and environmental pollutants on microorganisms

A thesis submitted to the University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Engineering and Physical Sciences

2010

Emma Sarah Wharfe

School of Chemistry

Table of Contents

Final word count: 59,529

List of Figures

List of Tables

List of Abbreviations

ANN	artificial neural network
API	active pharmaceutical ingredient
ATR	attenuated total reflectance
CaF_2	calcium fluoride
cm	centimetre
cm^{-1}	wavenumber
CV	canonical variate
CVA	canonical variates analysis
DF	discriminant function
DFA	discriminant function analysis
DGGE	denaturing gradient gel electrophoresis
DIMS	direct injection mass spectrometry
DNA	deoxyribonucleic acid
DTGS	deuterated triglycine sulphate
EM	electromagnetic
EMSC	extended multiplicative scatter correction
F	Fisher ratio
FIR	far infrared
FPA	focal plane array
FT-IR	Fourier transform infrared
g	grams
GA	genetic algorithm
GC	gas chromatography
GC-MS	gas chromatography mass spectrometry
GC-TOF-MS	gas chromatography time-of-flight mass spectrometry
h	hour
HCA	hierarchical cluster analysis
HeNe	helium neon
HPLC	high performance liquid chromatography
IR	infrared
JM	Jaworski's medium
L^{-1}	per litre
LC	liquid chromatography
LC-MS	liquid chromatography mass spectrometry

LHS	left hand side
MALDI	matrix assisted laser desorption ionisation
MCT	mercury-cadmium-telluride
mg	milligram
min	minute
MIR	mid infrared
mL	millilitre
mM	milli-molar
mm	millimetre
MS	mass spectrometry
MVA	multivariate analysis
ng	nanogram
NIPALS	non-iterative partial least squares
NIR	near infrared
OD	optical density
PC	principal component
PCA	principal components analysis
PCR	principal components regression
PLS	partial least squares
PLSR	partial least squares regression
pM	picomolar
PMF	peptide mass fingerprint
PR	pattern recognition
QUAD	quadrupole detector
RHS	right hand side
RI	retention index
RMS	root mean squared
RMSEP	root mean square error of prediction
RNA	ribonucleic acid
RNA-SIP	RNA stable isotope probing
rpm	revolutions per minute
RubisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
s	second
SNR	signal to noise ratio
STP	sewage treatment plant
SVD	singular value decomposition

TEAA	triethylamine acetate
TOF	time-of-flight
v/v	volume by volume
ZnSe	zinc selenide
℃	degrees Celsius
µg mL^{-1}	microgram per millilitre
µL	microlitre
µm	micrometre

Abstract

The University of Manchester
Emma Sarah Wharfe
Degree of Doctor of Philosophy in the Faculty of Engineering and Physical Sciences
An environmental metabolomics investigation of the effects of chiral pharmaceuticals and environmental pollutants on microorganisms
March 2010.

Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment and their occurrence in the aquatic environment is of growing concern. However, despite the fact that these may cause harmful effects in organisms found within this niche, little is currently known about the effects of APIs in the aquatic environment. Chiral pharmaceuticals are of particular concern as the enantiomers may be metabolised differently, with the potential for the production of an array of harmful compounds. There are many racemic APIs for treating human and animal conditions, and even in these target organisms the pharmacodynamic effects of the enantiomers are not always known. Within recent years the importance of the interactions of these compounds within the aquatic environment has been realised and information regarding the fate and biodegradation of such environmental pollutants is of great importance. The advent of post-genomic technologies has proved advantageous in the study of the effects of these environmental pollutants.

In this thesis, the effects of a range of chiral APIs, and other environmental pollutants, on environmentally relevant microorganisms were investigated at the metabolome level. The effects of chiral APIs were investigated in a number of prokaryotic and eukaryotic systems in order to provide a comprehensive study of the effects of the APIs in the aquatic environment. FT-IR spectroscopy was employed for metabolic fingerprinting of some environmentally relevant bacteria and GC-MS was subsequently employed for metabolite profiling of two pseudomonads that had shown differential chiral effects with Propranolol. In addition, FT-IR microspectroscopy was employed for the investigation of the phenotypic and localised effects of chiral APIs in a eukaryotic system. Furthermore, the effects of a range of environmental pollutants on a complex bacterial community were investigated with the use of FT-IR spectroscopy and multivariate analysis. Initial results indicated a large phenotypic response in relation to phenol, and this was further explored with a range of ageing experiments and metabolic fingerprinting. An FT-IR peak was found to be characteristic of the phenotypic changes in the actively degrading communities and this was likely to be a degradation product of phenol, and armed with this knowledge the activated sludge community was monitored during the active degradation of phenol with the use of GC-MS.

The work presented in this thesis has shown for the first time that metabolomics allows subtle phenotypes in microorganisms to be revealed when they are exposed to chiral forms of APIs which are commonly found in the aquatic environment. Despite these APIs not being designed for any interaction with bacteria and aquatic life in general these are significant findings and may have implications as more and more APIs become detectable and concentrated in the environment due to continued use in man and indeed animals or aquaculture.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Notes on Copyright:

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on presentation of Theses.

Acknowledgements

I thank my supervisor, Prof. Roy Goodacre, for his help, guidance and encouragement throughout the course of this work. I also thank all of my friends and colleagues at the University of Manchester for their help and support and BBSRC and AstraZeneca for funding for this work.

In particular I would like to thank Warwick Dunn for his help and advice and for running the GC-MS analyses carried out during the course of this work. I have had the pleasure of working with Cate Winder and Roger Jarvis who I would like to thank for sharing their wealth of scientific knowledge with me, great friendship and generally keeping me sane throughout this entire process!

Finally, I thank my parents, Margaret and Jeff Wharfe, who have offered me endless support (both financially and emotionally), without whom this experience would not have been possible.

1 Introduction

1.1 Pharmaceuticals in the Aquatic Environment

Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment (Escher *et al.*, 2005) and the occurrence of APIs in the aquatic environment is of growing concern (Fent *et al.*, 2006). There are a number of routes through which APIs and resulting metabolites / degradation products may enter these ecosystems (Figure 1.1), and a common avenue is through the excretion of the APIs and their metabolites in urine and faeces (Heberer, 2002). It is known that APIs have different rates of metabolism within man. For example, the β-blocker Propranolol is almost completely metabolised within the liver and only 1-4% of an oral dose is excreted as the unchanged API and its metabolites. By contrast, 40-50% of an oral dose of Atenolol (also a β-blocker), is excreted as the API or its constituent metabolites (Ashton *et al.*, 2004, Carlsson *et al.*, 2006a, Carlsson *et al.*, 2006b). The subsequent degradation of the APIs and their metabolites may also occur at sewage treatment plants (STPs); this is usually substrate specific and will vary greatly between APIs. The rate of adsorption to activated sewage sludge during treatment differs between APIs, and is dependant on the hydrophobic and electrostatic interactions of the API with the particulates and microorganisms, within the activated sewage sludge (Fent *et al.*, 2006). Any remaining API and relevant metabolites are diluted into the surface water when the effluent is released from the STP. Hence, many APIs are present in low concentrations (ng - μg L^{-1}) in aquatic environments such as rivers, streams and estuaries (Escher *et al.*, 2005, Ashton *et al.*, 2004, Carlsson *et al.*, 2006b). The majority of APIs are neither persistent nor highly bioaccumulative; however, the continuous release of APIs into the aquatic environment poses a potential risk to aquatic organisms even though the concentrations of APIs in receiving waters are quite low (Escher *et al.*, 2005). In addition, concentrations up to 31 mg L^{-1} have been reported in wastewater effluents of API production facilities, and 11 of the APIs detected in the effluent were found be present at concentrations >100 μg L^{-1} (Larsson *et al.*, 2007). Furthermore, the high concentrations of APIs in this effluent were found to elicit toxicity effects in the aquatic vertebrate *Xenopus tropicalis* and a fish species (*Danio rerio*) (Carlsson *et al.*, 2009).

Despite little being known about the effects of APIs in the environment, it must be taken into account that they are designed to elicit a specific mode of action in humans

(Escher *et al.*, 2005). Adverse side effects may be encountered in humans at higher doses of these APIs, and it can be expected that any beneficial or adverse effect may also be observed in aquatic organisms with similar biological functions or receptors. It must also be noted that similar targets may control different metabolic processes in different species (Seiler, 2002), and therefore APIs and their metabolites may act through additional modes of action in aquatic organisms. The effects of the APIs may be subtle due to the very low concentrations observed in the aquatic environment, and as a result these effects may go unnoticed (Escher *et al.*, 2005). It is also likely that the effect of the API will impact on the local population dynamics throughout the whole ecosystem, from bacteria up to higher organisms. To explore the effects of the APIs on biological systems a wide range of concentrations should be employed with the application of appropriate analytical platforms to profile the complement of biochemical components in the cell. Indeed it is known that APIs could become concentrated in the benthic environment of river beds and as bacteria inhabit this niche, this community may be exposed to higher levels than expected (Thacker, 2005, Halling-Sørensen *et al.*, 1998, Pouliquen *et al.*, 1992).

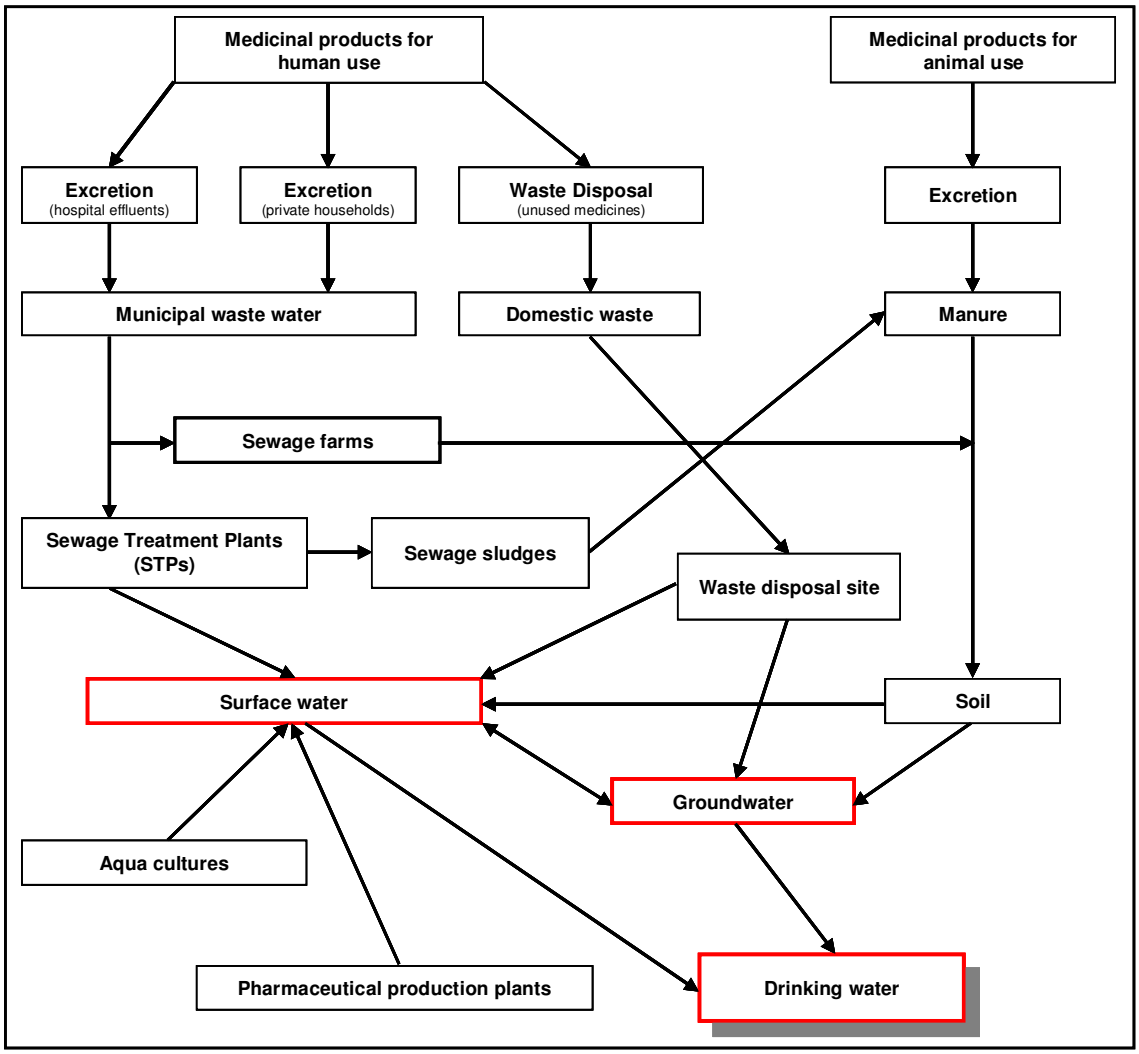


Figure 1.1: Scheme showing possible sources and pathways for the occurrence of API residues in the aquatic environment. Points outlined in red represent significant aquatic end-points in the pathways. Adapted from Heberer *et al.* (2002).

1.1.1 Chiral Pharmaceuticals

Whilst the effects of APIs in the environment is currently a growing area of research there is very little understanding of the environmental effects of chiral pharmaceuticals (Fono and Sedlak, 2005, Buser *et al.*, 1999). A chiral molecule is one that lacks an internal plane of symmetry. The non-superimposable mirror images are termed enantiomers and are labelled '(*R*)' or '(*S*)' according to a priority system (Cahn Ingold Prelog priority rules) based on the atomic number of the molecules substituents. Approximately 56% of the APIs currently in use are chiral compounds, 88% of which are administered therapeutically as the racemate (i.e., an equal mixture of the two enantiomers, represented by the symbol (±)). The chirality of environmental contaminants such as APIs must be taken into consideration in order to understand the environmental fate and effects of these compounds fully.

The enantiomers of a chiral API are able to interact differently with other chiral compounds such as enzymes, and therefore potentially exert different effects when released into the environment (Fono and Sedlak, 2005, Buser *et al.*, 1999, Nikolai *et al.*, 2006). It is widely known that the enantiomers of a chiral API may possess differing toxicological and biological effects both from each other, and the racemate (Lees *et al.*, 2003, Yang *et al.*, 2005). It has been shown that the (*S*)- enantiomers of the β-blocking agents Atenolol and Propranolol (Figure 1.2) are more potent in man than their respective antipodes (Davies, 1990, Pearson *et al.*, 1989, Kurt *et al.*, 1993, Barrett and Cullum, 1968) and that a number of the biotransformation pathways for β-blockers are stereoselective in humans (Mehvar and Brocks, 2001). The mode of action of the drugs and their enantiomers is not known in prokaryotic systems. It is therefore necessary to enhance our understanding of the fate and biological effects of chiral pharmaceuticals on typical microflora from the aquatic environment in order to fully appreciate the risks (Huggett *et al.*, 2002). Of particular interest is the group of APIs termed β-blockers as they all contain at least one chiral centre and are generally administered therapeutically as the racemate (Mehvar and Brocks, 2001). In addition, they are widely used with approximately 29 and 12 tonnes of Atenolol and Propranolol, respectively, consumed each year in the UK. Furthermore, these APIs have been detected in STP effluents and surface waters and are frequently found to be within the top 15 most concentrated APIs present within the aquatic environment (Jones *et al.*, 2002, Ashton *et al.*, 2004, Carlsson *et al.*, 2006a, Carlsson *et al.*, 2006b).

The chiral API Propranolol is a non-selective β-adrenergic blocking agent and whilst it is known that the (*S*)- enantiomer exerts the majority of the β-blocking effect, it is administered therapeutically as the racemate (Barrett and Cullum, 1968). Studies carried out in cats and on the atrial muscle taken from guinea pigs found the (*S*)-enantiomer to be approximately 100 times more potent and have a longer half-life than its respective antipode (Barrett and Cullum, 1968). In addition, (*R*)- Propranolol is reported to exert membrane stabilising activity (MSA) at therapeutic concentrations (Wood, 1984). The term MSA is employed to describe non-specific interactions that may occur between membrane lipid bilayers and lipophilic drugs and chemicals (Roth and Seeman, 1971). These interactions may inhibit membrane permeability to electrolytes, through the blockade of sodium fast channels. MSA may also be used to describe the non-specific physical protection of membranes, reports of which include red blood cell studies (Langslet, 1970, Seeman and Carl, 1966), and the release of lysosomal enzymes (Welman, 1979). However, studies in which properties of β-blockers may not be explained by competitive interaction with the adrenergic receptor frequently result in such effects being attributed to MSA (Arnim and Welman, 1981, Nayler *et al.*, 1980).

Toxicity studies carried out by Kim and co workers in the crustacean *Thamnocephalus platyurus* and a fish species (*Oryzias latipes*) reported that (±)- Propranolol caused acute toxicity in *T. platyurus* at a concentration of 10.61 μg mL^{-1} and in *O. latipes* at a concentration of 11.40 μg mL^{-1}. In addition, toxicity studies have been carried out on a range of APIs (including (±)- Propranolol) in the Japanese medaka fish (*Oryzias latipes*) an amphipod (*Hyalella azteca*) and two crustaceans (*Ceriodaphnia dubia* and *Daphnia magna*). It was found that Propranolol caused the greatest effect in the organisms studied. The crustacean *C. dubia* displayed toxicity responses at a concentration of 250 μg mL^{-1}. Propranolol was the only API investigated which was found to cause acute toxicity effects in the Japanese medaka fish. These effects were observed at a concentration of 0.5 μg mL^{-1} (Huggett *et al.*, 2002).

In contrast to the non-selective β-blocking agent Propranolol, Atenolol is a β_1-selective blocking agent. The (*S*)- enantiomer is reported to posses the majority of the pharmacological action in man (Kurt *et al.*, 1993, Pearson *et al.*, 1989). However, no additional action has been reported for either of the enantiomers. Atenolol is a hydrophilic β-blocker and as such does not posses MSA as it does not interact with the membrane lipid bilayer (Pearson *et al.*, 1989). Toxicity studies have shown reported

effects on polyp structure of the freshwater cnidarian *H. vulgaris* at 10 µg mL^{-1} following seven day exposure to the β-blocker (Pascoe *et al.*, 2003). In addition, reductions in heart beat rate and biomass were observed in *Daphnia magna* following chronic exposure to 0.11 µg mL^{-1} (±)- Atenolol (Fent *et al.*, 2006). To date, there have been no reports on the effects of the pure enantiomers of either Atenolol or Propranolol in aquatic organisms and this is an area of environmental concern.

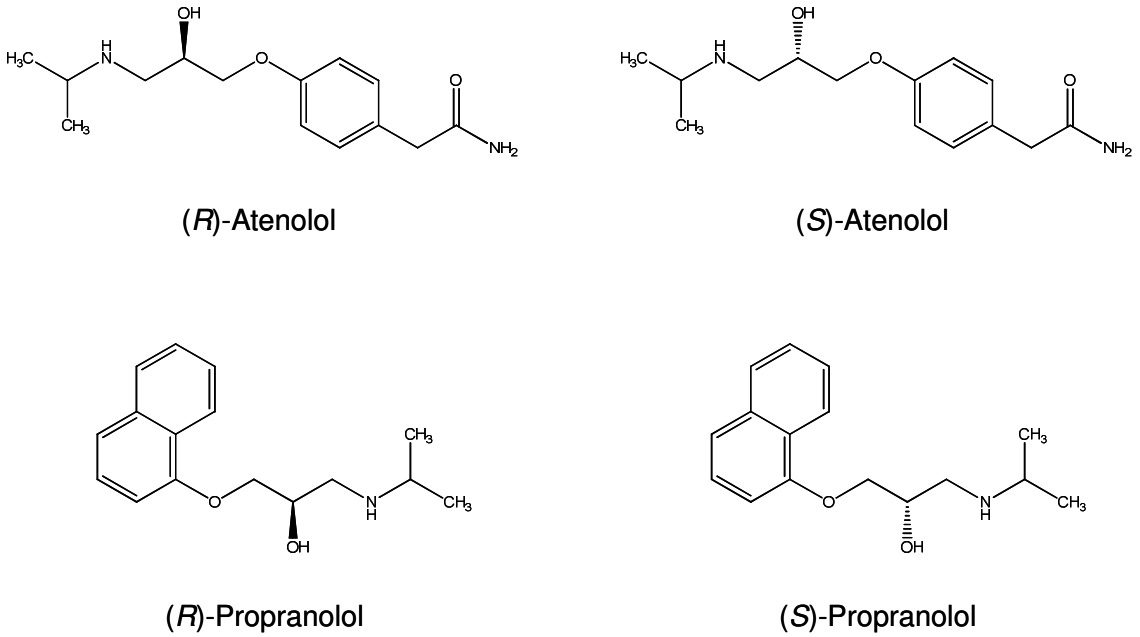


Figure 1.2: The chemical structures of the chiral β-blocking agents Atenolol and Propranolol.

1.1.2 Aromatic Hydrocarbons

The risks posed to the aquatic environment by aromatic hydrocarbons such as phenol and toluene are becoming increasingly apparent and many aromatic hydrocarbons are classed as environmental pollutants (Nahar *et al.*, 2000). Industrial processes such as the production of plastics, synthetic fibres and pesticides and the coking process result in the production of vast quantities of wastewaters containing aromatic hydrocarbons (Abuhamed *et al.*, 2004). Regulations regarding the discharge of such wastewaters are becoming increasingly more stringent worldwide (Lu *et al.*, 2008).

Phenol has been shown to cause toxicity responses in aquatic organisms including *Daphnia magna*, and a range of fish species. Kühn *et al* (1989) reported *Daphnia magna* exposed to phenol for 24 h displayed toxicity responses at 10 µg mL^{-1} and

these effects were noted in organisms exposed to only 4 µg mL^{-1} for 48 h. The reported increase in toxicity of phenol for extended exposure periods is highly relevant to aquatic organisms that are likely to be exposed to such pollutants during their entire lifespan. The range of observed toxicity responses in fish include genotoxic, immunotoxic, and physiological effects (Hori *et al.*, 2006, Saha *et al.*, 1999, Taysse *et al.*, 1995). In addition, phenol is highly bioaccumulative due to its lipophilic nature and environmental concentrations may exceed those previously expected (Hori *et al.*, 2006). Furthermore, the toxicity of mixtures of aromatic hydrocarbons such as phenol, toluene and benzene has been shown to increase in comparison with the isolated chemicals (Abuhamed *et al.*, 2004). This is of great relevance to the aquatic environment, which is likely to contain a variety of chemicals.

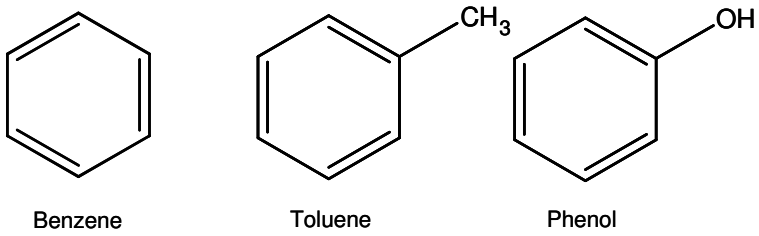


Figure 1.3: The chemical structures of the aromatic hydrocarbons benzene, toluene and phenol.

1.2 Ecotoxicogenomics

The term genomics encompasses a wide range of scientific disciplines including, genome sequencing, assigning gene function to known genes, determination of the genome architecture, investigating gene expression at the transcriptome level, protein expression at the proteome level and metabolite flux at the metabolome level (Snape *et al.*, 2004). Toxicogenomics combines the disciplines of genomics and toxicology, and may be described as the study of genes and their products which play a central role in the adaptive responses to abiotic perturbations within biological systems (Nuwaysir *et al.*, 1999, Iannaccone, 2001, Rockett and Dix, 1999, Lovett, 2000, Pennie *et al.*, 2000, Watanabe, 2007). This approach enables the further understanding of the underlying molecular responses to chemical contaminants (Bradley and Theodorakis, 2002, Moore, 2002). Similarly, the term ecogenomics refers to the application of genomics to the field of ecology (Chapman, 2001). This approach aims to develop further understanding of the interactions between organisms in the environment and the

interactions between organisms and the environment itself (Snape *et al.*, 2004, Zelikoff, 1993, Hutchinson *et al.*, 1999, Hutchinson *et al.*, 2003, Harvell *et al.*, 1999, Eggen and Suter, 2007).

The term *ecotoxicogenomics* was first proposed by Snape *et al.* (2004) to describe the integration of the 'omics technologies, transcriptomics, proteomics and metabolomics into the field of ecotoxicology. This is defined as the study of gene, protein and metabolite expression in biological systems that is important in response to environmental toxicant exposure (Snape *et al.*, 2004) and aims to understand the key mechanisms of toxicity (Poynton and Vulpe, 2009, Poynton *et al.*, 2008, Watanabe and Iguchi, 2006). In addition, ecotoxicogenomics aims to link molecular and cellular biomarkers with the higher level population and ecosystem responses; in order to predict potential ecological risks for new compounds providing an enhanced mechanistic understanding of aquatic ecotoxicology (Snape *et al.*, 2004, Watanabe and Iguchi, 2006, Crane *et al.*, 2006). Environmental contaminants are likely to produce genomic responses within an organism (Snape *et al.*, 2004) and within the field of biomedicine, it has been shown that exposure to toxicants (such as polycyclic aromatic hydrocarbons) almost always results in altered gene expression (Nuwaysir *et al.*, 1999). At present, the relationship between genotype and phenotype is poorly understood, and the application of functional genomics may provide insight into both the genotypic and phenotypic effects of contaminants in the environment (Snape *et al.*, 2004, Thompson, 1991, Depledge, 1994, Shugart and Theodorakis, 1996, Hood *et al.*, 2000).

A broad range of gene expression technologies have been developed in order to investigate organisms at a genome wide level (Patanjali *et al.*, 1991, Arnheim and Erlich, 1992, Schena *et al.*, 1995, Velculescu *et al.*, 1995, Rockett and Dix, 2000, Clark *et al.*, 2002). In addition, methods for the investigation of biological systems at the proteome (Gevaert and Vandekerckhove, 2000, Mann *et al.*, 2001) and metabolome level (Nicholson *et al.*, 1999, Thomas, 2001, Watkins and German, 2002, Buchholz *et al.*, 2002) have also been developed. However, few of these methods have so far been applied to the response of organisms to contaminants in complex environments (e.g. sediments or soil). A key objective for ecotoxicogenomics is the improvement of existing, and the development of new, technologies capable of measuring the genome, proteome and metabolome profiles for environmental organisms that are exposed to numerous perturbations in complex natural environments (Snape *et al.*, 2004).

Modern high-density DNA techniques are capable of simultaneously analysing thousands of genes and are becoming widely available for both model and non-model organisms (Snape *et al.*, 2004, Schena *et al.*, 1995, Rockett and Dix, 2000). Application of these techniques allows for the identification of transcripts which are up- or down- regulated. The major advantage of these techniques is the ability to rapidly analyse vast numbers of transcripts within a single analysis. However, such techniques are only semi-quantitative as they are dependent on the efficiency of the labelling and hybridisation of the individual clones (Snape *et al.*, 2004). In addition, the detection of low copy number transcripts is currently rather difficult (Bartlett, 2001). Furthermore, array-based techniques rely on *a priori* knowledge of the transcripts that are likely to be present within a sample in order that gene probes may be designed prior to analysis (Clark *et al.*, 2002).

Proteomics is the measurement of the complement of proteins within a biological system that are expressed under specific conditions (Snape *et al.*, 2004, Monsinjon and Knigge, 2007). This approach generally employs 2D gel electrophoresis for the separation of proteins (firstly by charge and then by size). However, for peptides or protein fragments, high-resolution chromatography techniques such as high performance liquid chromatography (HPLC) are more commonly employed. The separation of proteins is then followed by the chemical characterisation of the proteome, usually by mass spectrometry which has the sensitivity, selectivity and throughput required in order to identify each protein within the proteome (Snape *et al.*, 2004). The resultant peptide mass fingerprint (PMF) may then be compared with fragment databases in order to provide identification of the proteins (Gevaert and Vandekerckhove, 2000, Mann *et al.*, 2001). Proteomics approaches to ecotoxicology investigations, in which measurements of the proteomic response to disease or abiotic perturbation during a time course may be required, are both costly and time-consuming. A further disadvantage to the application of proteomics to the study of the effects of contaminants in environmental populations is the lack of relevant DNA sequence information and PMF databases meaning that identification is not possible for many of the proteins within a sample (Snape *et al.*, 2004). However, the continuing advances in the DNA sequencing of environmental metagenomic samples and the *de novo* sequencing of protein MS/MS spectra, may provide further insight into the effects of contaminants on environmental populations (Maroti *et al.*, 2009, Schulze *et al.*, 2005, Venter *et al.*, 2004).

Metabolomics is the non-biased identification and quantification of all metabolites within an organism or biological system. (Fiehn, 2002, Harrigan and Goodacre, 2003, Oliver *et al.*, 1998). Metabolomics and proteomics approaches possess two clear advantages with respect to the analysis of gene function in ecotoxicology studies; the total complements of proteins and metabolites change according to the physiological, developmental or pathological state of the biological system studied, and are functional entities within an organism. Analysis at the proteomic and metabolomic levels therefore enables the characterisation of a biological system at the functional level (Raamsdonk, 2001, Snape *et al.*, 2004). In addition, for many organisms the number of metabolites is far fewer than the number of genes or gene products (Raamsdonk, 2001).

Disease or abiotic effects on gene expression may be detected at the genomic, proteomic and metabolomics levels. However, some environmental contaminants may only act at a pharmacological level and will not affect gene regulation and expression or protein production, and may therefore only be detected at the metabolome level (Lindon *et al.*, 2001b). Drug induced and other pathophysiological changes result in alterations of the ratios, concentrations, binding and fluxes of endogenous biochemicals, through direct chemical reaction or binding to key metabolic macromolecules. Sufficiently large disturbances will however, affect the efficient functioning of the whole organism and will therefore be reflected in the genome, proteome and metabolome (Lindon *et al.*, 2001a).

Functional studies have emphasised analyses at the transcriptomics, proteomic and metabolomic levels with a view to a 'systems biology' approach of defining the phenotype and to identify the linkages between the genotype and phenotype (Goodacre *et al.*, 2004, Fiehn, 2002). The observed response of an organism to toxic or other stresses at all three levels gives a comprehensive description of the RNAs, proteins and metabolites within the system in response to such stresses. This approach enables analysis of the relevant associations between macromolecules, identification of functional linkages, and the production of models that quantitatively describe the dynamics of the biological system (Goodacre *et al.*, 2004, Snape *et al.*, 2004).

1.3 Introduction to metabolomics

In the post-genomic era, the high degree of complexity inherent in the information produced by genomic studies has exposed the limitations of solely genomic-based investigations. Recent research has extended to the transcriptome, the proteome; and most recently, the metabolome.

The metabolome is the final downstream product of the genome, and has been defined as the qualitative and quantitative collection of all low molecular weight molecules present in a cell that participate in general metabolic reactions (Dunn *et al.*, 2005, Oliver *et al.*, 1998, Harrigan and Goodacre, 2003, Goodacre *et al.*, 2004). Analysis of the metabolome can provide a comprehensive evaluation of the physiological state of an organism coupled with unique insights into specific biochemical processes (Oliver *et al.*, 1998, Fiehn, 2001, Fiehn, 2002). The size of the metabolome varies greatly depending on the organism studied, for example *Saccharomyces cerevisiae* contains approximately 600 metabolites (Forster *et al.*, 2003) whereas the plant kingdom is estimated to have a total of approximately 200,000 metabolites although individual species have significantly lower numbers (Fiehn, 2001). The metabolome is more diverse in chemical and physical properties than the genome or the proteome due to the greater variations in atomic arrangements (Dunn *et al.*, 2005). Consequently, metabolomic investigations comprise the analyses of an extensive range of chemical species, from low molecular weight polar volatiles such as ethanol, to high molecular weight polar glucosides, non-polar lipids and inorganic species (Dunn *et al.*, 2005, Lahner *et al.*, 2003). Furthermore, the range of metabolite concentrations present in a cell can vary over nine orders of magnitude (pM-mM), making analysis of the metabolome significantly more complex than that of the genome or proteome (Dunn *et al.*, 2005).

1.3.1 Applications of Metabolomics

Metabolomics is a complementary technique to the other ‘omics disciplines and is perceived as having a number of advantages (Goodacre *et al.*, 2004). As the final downstream product of the genome, the metabolome more accurately reflects the activities of the cell at a functional level and changes in the metabolome are amplified relative to the proteome or the transcriptome (Urbanczyk-Wochniak, 2003). In general the number of metabolites is lower than the number of genes or gene products (Raamsdonk, 2001). Changes in the quantities of enzymes have significant effects on

metabolite concentrations despite having little effect on metabolic fluxes (Raamsdonk, 2001). However, some research has found metabolic flux analysis to be very useful for understanding biochemical processes in cancer (Boros *et al.*, 2003, Boros *et al.*, 2002, Boros *et al.*, 2004, Lane *et al.*, 2009, Fan *et al.*, 2009, Forbes *et al.*, 2006). In addition, there is evidence that metabolic fluxes are regulated by environmental stresses as well as gene expression (Johnson *et al.*, 2003), and therefore the measurement of the metabolome may be a more sensitive technique to yield information about the biological system (Brown *et al.*, 2005, Harrigan and Goodacre, 2003). Global metabolomics approaches have been applied to many different areas of biological and clinical study including; clinical diagnosis (Lindon *et al.*, 2003, Nicholls *et al.*, 2001, Matsumoto and Kuhara, 1996, Rashed, 2001, Brindle *et al.*, 2002), mode of action investigations (Allen *et al.*, 2004, Aranìbar *et al.*, 2001), microbial systems (MacKenzie *et al.*, 2008, Smedsgaard and Nielsen, 2005), metabolic engineering studies (Fiehn *et al.*, 2000, Weckwerth *et al.*, 2004, Bro and Nielsen, 2004) and environmental stress (Viant, 2007, Viant *et al.*, 2003).

1.3.2 Instrumentation Employed in Metabolomics Investigations

The biochemical status of an organism can be revealed through the qualitative and quantitative measurements of its cellular metabolites and these data can be employed to both monitor, and determine gene function (Fiehn, 2001, Fiehn, 2002). The analytes studied vary greatly in both concentration and nature and this complexity makes the simultaneous profiling of the global metabolome particularly challenging to the analytical technologies employed in metabolomics investigations. At present no single analytical technology or extraction method is able to detect every metabolite (Hall, 2006). Therefore, a range of targeted extraction procedures and analytical technologies are employed (Sumner *et al.*, 2003, Weckwerth, 2003).

A range of terminologies have been applied to the varying metabolomics approaches and these are detailed in Table 1.1. Metabolomics studies may employ a wide range of metabolic fingerprinting and profiling methods. Metabolic fingerprinting methods involve the high-throughput analysis of crude samples or simple cellular extracts and provide a global metabolite fingerprint. These techniques are frequently employed in the classification or screening of samples (Johnson *et al.*, 2003, Winder *et al.*, 2004, Winder *et al.*, 2006). However, quantification and identification of metabolites is not generally carried out in this approach (Dunn *et al.*, 2005). Fingerprinting methods

present an inexpensive means of screening a biological system prior to metabolic profiling with more costly techniques (Kaderbhai *et al.*, 2003). Metabolic profiling methods commonly employ hyphenated mass spectrometry (MS) approaches with the aim of detecting, identifying and quantifying the metabolites within a biological sample (Bino *et al.*, 2004, Dunn *et al.*, 2005).

Table 1.1: Terms and Definitions used in Metabolomics.

Term	Definition
Metabolome	The complete biochemical complement present in an organism (Fiehn, 2002, Goodacre *et al.*, 2004).
Metabolomics	Non-biased identification and quantification of all metabolites within an organism or biological system. Requires the use of highly selective and sensitive analytical techniques. At present, no single analytical technique, or combination of techniques can determine all metabolites in microbial, plant or mammalian metabolomes (Dunn *et al.*, 2005, Goodacre *et al.*, 2004, Kell, 2006).
Metabonomics	The quantitative analysis of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification (Nicholson *et al.*, 1999). Often used synonymously with metabolomics
Metabolic profiling	The identification and quantification of metabolites related through similar metabolic pathways or chemistries. Chromatographic separation is normally employed prior to detection.
Targeted metabolite analysis	The quantitative determination of one or a few metabolites related to a specific metabolic pathway following extensive sample preparation and separation from the sample matrix. This approach employs chromatographic separation and sensitive detection
Metabolic flux analysis	The quantitative monitoring of whole network operation. MFA provides a global perspective on the integrated regulation at all levels (including gene expression, enzyme kinetics, and allosteric regulation) (Sauer, 2005).
Metabolic fingerprinting	Rapid and high-throughput methods where global metabolite profiles are obtained for sample classification or screening. Analysis is of crude samples or simple cellular extracts. Identification and quantification of metabolites is not performed.
Metabolic footprinting	The global measurement of metabolites secreted from the intracellular complement into the extracellular growth medium. High-throughput method which requires minimal metabolite quenching and extraction. Has been frequently employed in microbial metabolomics.

1.4 Metabolic Fingerprinting Technologies

Metabolic fingerprinting is the high-throughput screening of intact samples and cellular extracts. These biochemical fingerprints are usually analysed with chemometrics techniques and used to model relationships between samples and to classify them according to their origin or biological relevance (Fiehn, 2001). Metabolic fingerprinting provides "biochemical snapshots" of biological samples at a specific time (Dunn *et al.*, 2005).

One of the popular approaches for metabolic fingerprinting is based on vibrational spectroscopy. These techniques measure the vibrations of bonds within functional groups when they interact with electromagnetic (EM) radiation (light). The two vibrational spectroscopy techniques that are used routinely for metabolic fingerprinting (or what has been termed disease pattern recognition (Himmelreich *et al.*, 2003, Petrich *et al.*, 2000, Staib *et al.*, 2001)) include: Fourier transform infrared (FT-IR) spectroscopy, which is based on the absorption of EM radiation; and Raman spectroscopy, which is based upon inelastic light scattering (Choo-Smith *et al.*, 2001). These vibrational spectroscopy techniques are not as sensitive as mass spectrometry or nuclear magnetic resonance (NMR) spectroscopy methods and do not provide metabolite identification from complex biological samples, however, they do provide a relevant metabolomics tool. Fourier transform-infrared (FT-IR) spectroscopy is currently the most frequently employed technique in metabolic fingerprinting and enables the rapid, non-destructive and comparatively inexpensive analysis of complex biological systems (Dunn *et al.*, 2005, Kell, 2004, Oliver *et al.*, 1998). Therefore, FT-IR spectroscopy provides a first round screening, or starting point for metabolomic analyses, which may direct the focus of more extensive metabolomics studies in complementary and more sensitive analytical platforms (Dunn *et al.*, 2005).

1.4.1 FT-IR spectroscopy

Infrared spectrometers have been commercially available since the 1940s; however, the technology employed a conventional prism and grating monochromators and each wavelength within the IR region had to be measured in succession. This technology therefore, had several drawbacks with respect to analysis time, heat damage, sensitivity and reproducibility. The most significant advances in infrared spectroscopy

occurred with development of the interferometer and the introduction of the mathematical Fourier transform. The utilisation of these methods enables the rapid and simultaneous analysis of the entire IR spectrum (Banwell and McCash, 1994). Consequently, the field of FT-IR spectroscopy has made great strides within recent years (Stuart, 1996).

Infrared spectroscopy is based on the various vibrations of the bonds between the atoms of a molecule. Following interrogation with IR light the functional groups within a sample absorb the light at specific wavelengths and vibrate in bending, twisting and stretching motions thus causing changes in the dipole moment of that bond. FT-IR spectroscopy measures the frequency and intensity of the radiation which is absorbed by a sample and the resultant spectrum is therefore representative of these various vibrations (Banwell and McCash, 1994). The determination of such factors may provide insight into the chemical composition of a biological sample or allow for the identification of an unknown compound (Harrigan and Goodacre, 2003).

The infrared spectrum can be divided into three regions corresponding to wavenumbers between 14000 – 4000 cm^{-1} (near-infrared), 4000 – 400 cm^{-1} (mid-infrared) and 400 – 10 cm^{-1} (far-infrared). Measurement of the far-infrared (FIR) region is more difficult that that of the near- and mid-infrared regions due to the weakness of the source employed. A mercury lamp is most often employed as the source for measurement of the FIR; however the signal-to-noise ratio is often higher than that of the mid- or near-infrared regions. Consequently, the majority of infrared applications utilise the mid-infrared (MIR) and near-infrared (NIR) regions. Measurement of the NIR region provides information regarding the absorption characteristics of CH, OH and NH groups (Belton *et al.*, 1987, Dunn *et al.*, 2005). The spectra produced from the NIR region contain broad features which strongly overlap and therefore require high levels of statistical manipulation (Belton *et al.*, 1987). In contrast, measurement of the MIR region gives information regarding the chemical and structural composition of a sample which is open to direct interpretation (Griffiths and de Haseth, 2007). Thus, the use of the MIR in metabolomics FT-IR spectroscopy applications is currently favoured (Johnson *et al.*, 2003, Johnson *et al.*, 2004, Dunn *et al.*, 2005). The IR absorption of water is particularly strong within the MIR region resulting in the production of broad bands in the spectra which may mask the biologically important chemical information. These effects may be removed from the FT-IR spectra by post-analysis data processing methods, however, the most common approach employed is the drying of

samples prior to analysis (Johnson *et al.*, 2003, McGovern *et al.*, 1999, Timmins *et al.*, 1998, Schuster *et al.*, 2001, Curk *et al.*, 1994). Alternatively an optics accessory such as an attenuated total reflectance (ATR) cell may be employed for the sampling of liquids such as plasma or growth media. It is composed of a crystal such as zinc selenide and reduces the absorption of water enabling the direct and non-destructive analysis of liquids (Bouhedja *et al.*, 1997, Winder and Goodacre, 2004, Dunn *et al.*, 2005).

The MIR region may be further broken down into 5 major regions of biological significance (Forster *et al.*, 2003). These regions are termed the fatty acid region (3000 – 2800 cm^{-1}), the amide region (1700 – 1500 cm^{-1}) which can be divided into bands dominated by amide I and amide II, a region at wavenumbers 1450 - 1200; 1200 – 1250 cm^{-1} displaying vibrations from the carboxylic groups of proteins and PO_2^- of phosphodiesters, the polysaccharide region (1200 – 900 cm^{-1}) and the fingerprint region (<900 cm^{-1}) which consists of a variety of weak spectral features (Dunn *et al.*, 2005, Naumann *et al.*, 1996) (See Figure 1.4 for further details). The characteristics of the spectra; the frequency, intensity, number of bands and half widths are distinctive, therefore providing a unique fingerprint for the sample (Naumann *et al.*, 1991b).

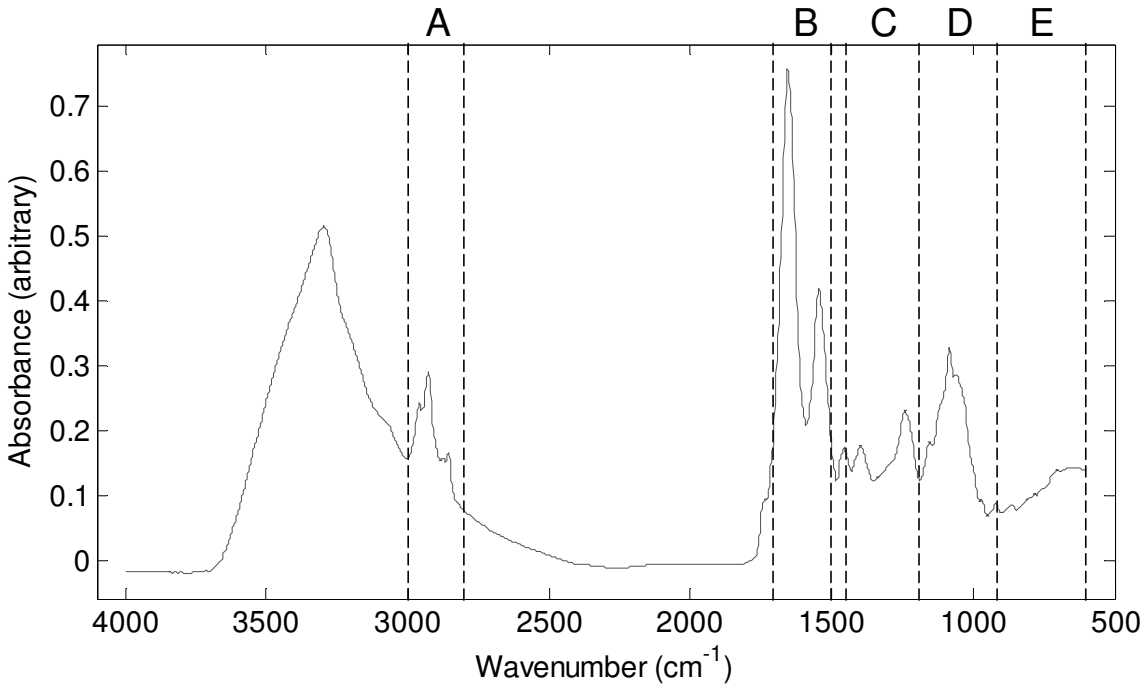


Figure 1.4: A typical FT-IR absorbance spectrum of *Pseudomonas aeruginosa* PA14. The major regions of biological interest are highlighted A-E. Refer to Table 1.2 for further details.

Table 1.2: Major spectral regions of biological interest.

	Wavenumber (cm^{-1}) range	**Dominant Compounds**
A	3000-2800	CH_x stretches from fatty acids
B	1700-1500	Proteins
	(1700-1600)	C=O from amide I
	(1600-1500)	C-N and C-N-H from amide II
C	1450-1200	Carboxylic groups of proteins, free amino acids, polysaccharides
	(1250-1200)	P-O from RNA/DNA, phospholipids
D	1200-900	C-O or O-H from polysaccharides
E	<900	Unassigned

Figure 1.5 highlights the basic components of an FT-IR spectrometer. Infrared radiation emerges from the source, passes through the interferometer to the sample. The IR radiation is absorbed by the sample before reaching the detector. The signal is then amplified prior to conversion to a digital form by an analogue to digital converter and then transferred to a computer when the Fourier transformation is carried out and the data are converted to an IR spectrum (Stuart, 1996).

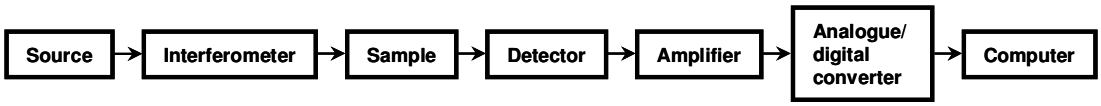


Figure 1.5: Flowchart of the components of an FT-IR spectrometer.

1.4.1.1 The Infrared Source

The source of infrared emission in IR spectrometers typically consists of an inert solid filament which is electrically heated to incandescence. The most commonly used mid-infrared source is the *Globar*, which is a resistively heated rod constructed of silicon carbide (Griffiths and de Haseth, 2007). In addition, the *Nernst glower*, which is composed of a mixture of refractory oxides, and a tungsten filament is commonly employed (Griffiths and de Haseth, 2007). Following heating of the source to incandescence, IR radiation is emitted and directed to a power mirror, which directs the IR beam through an aperture resulting in collimation of the beam. A series of flat mirrors then direct the IR beam to the interferometer (Stuart, 1996, Griffiths and de Haseth, 2007, Banwell and McCash, 1994).

1.4.1.2 The Michelson Interferometer

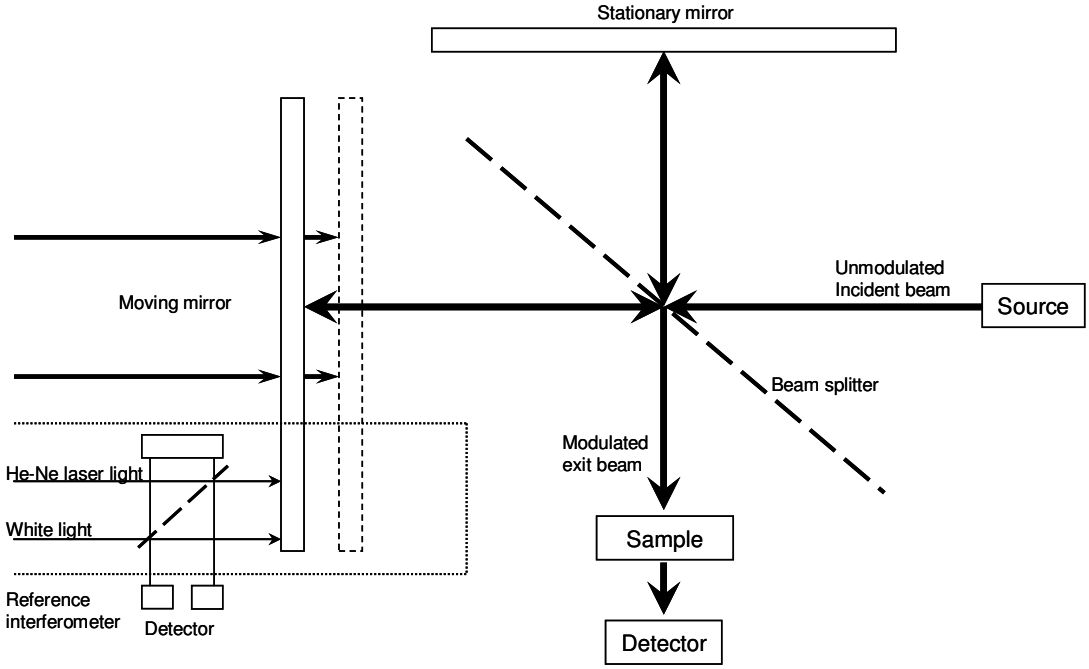


Figure 1.6: A Michelson interferometer showing planar mirrors and the beamsplitter. Adapted from Stuart (1997).

The Michelson interferometer is the basic optical component of the FT-IR spectrometer (Figure 1.6) (Stuart, 1996). It consists of two mirrors, one of which is stationary and the other of which can travel in a direction perpendicular to the plane. The planes of these two mirrors are bisected by a semi-reflecting film termed the beamsplitter. The beamsplitter is composed of material which is chosen according to the infrared region of investigation. For the near- or mid-infrared region, an infrared transparent substrate such as caesium iodide or potassium bromide is employed. The substrate is coated with germanium or iron oxide, and 50% of the radiation falling onto the beamsplitter is reflected by this coating (Stuart, 1996, Banwell and McCash, 1994).

Once the IR beam reaches the interferometer it is directed onto the beamsplitter, the coating of which reflects 50% of the beam to the stationary mirror. The mirror reflects the beam back to the beamsplitter, where it has travelled a fixed distance. The remaining 50% of the beam transmits through the beamsplitter to the moving mirror where it is reflected back to the beamsplitter and the two IR beams recombine and interfere. The transmitted beam interferes destructively and the reflected beam

interferes constructively (Stuart, 1996, Banwell and McCash, 1994). The spectral information is contained within this interference, which is then directed to the detector. The moving mirror produces a difference in the optical path between the two mirrors. It therefore has to be aligned in order to enable it to accurately scan distances. This alignment is carried out by a visible helium neon (HeNe) laser which is focused onto the corner of the mirror (Stuart, 1996, Griffiths and de Haseth, 2007, Banwell and McCash, 1994).

The interferometer enables the simultaneous and rapid analysis of all of the frequencies within the infrared spectrum. The function of the interferometer is to produce a novel signal of lower frequency than the original IR signal of a type that the detector can recognise. The resulting variations in beam intensity are then measured as a function of path difference. In FT-IR spectroscopy the spectral resolution is limited by the maximum path difference between the two infrared beams and the limiting resolution in wavelength (cm^{-1}) is calculated as the reciprocal of the pathlength difference (cm). A key advantage of rapid scanning instruments is the potential increase in the signal-to-noise ratio (SNR). This may be achieved by the collection of a number of interferograms which are then signal averaged (referred to as co-adds) by the computer.

1.4.1.3 The Infrared Detector

There are two types of detectors which are most commonly used in the mid-infrared region. The most common detector is a pyroelectric device which incorporates deuterated triglycine sulphate (DTGS) in a temperature resistant alkali halide window (Stuart, 1996). These thermal detectors are commonly employed due to their sensitivity across the entire infrared range (Banwell and McCash, 1994). Photoconductive detectors such as mercury-cadmium-telluride (MCT) are often employed in more sensitive applications. This type of detector contains conductive bands in which electrons are readily excited by the absorption of energy from an external source such as infrared radiation (Banwell and McCash, 1994). MCT detectors are highly sensitive and have a rapid response compared with DTGS detectors, requiring minute levels of excitation energy (Banwell and McCash, 1994). Consequently, excitation of electrons within the detector may effected by thermal sources other than the infrared radiation and are therefore cooled with liquid nitrogen (Griffiths and de Haseth, 2007).

1.4.2 FT-IR Microspectroscopy

FT-IR microspectroscopy, referred to as fast or hyperspectral imaging, is a relatively recent development, combining the data obtained from mid-infrared spectroscopy with the ability to acquire this data in a spatially resolved manner (Lewis *et al.*, 1995, Snively *et al.*, 1999, Treado *et al.*, 1994). A FT-IR imaging dataset contains both spatial and spectral information regarding the sample under investigation. The typical dimensions of a FT-IR imaging dataset range from 64 × 64 to 320 × 256 pixels, within which, each pixel contains an infrared spectrum (Snively *et al.*, 1999). This technique enables analysis of the spatial distribution of functional groups which may be connected to specific components within a cell or tissue sample (Patel *et al.*, 2008, Lewis *et al.*, 1995, Beardall *et al.*, 2001, Lasch *et al.*, 2002).

Traditionally FT-IR microscopes were employed to generate chemical maps by a single pixel at a time. In this approach the stage is moved over a raster in order to collect the spectrum for each individual pixel to generate a FT-IR map (Lewis *et al.*, 1995). This lengthy technique is both time consuming and limiting in respect to the fidelity of the images acquired. The use of a focal-plane array (FPA) detector significantly reduces the acquisition time, thus permitting the measurement of high density chemical maps in seconds to minutes (Lewis *et al.*, 1995). FT-IR imaging currently employs a step-scanning approach in which the moving mirror waits for the detector read-out prior to moving to the next position. In this approach all of the interferograms are collected simultaneously and subsequently transformed into IR spectra. The instrumentation is generally composed of a step-scan FT-IR spectrometer as a light source, a microscope with Cassegrainian optics, and an FPA detector (Snively *et al.*, 1999). The moving mirror of the interferometer modulates the light coming from the spectrometer and this is translated in order to collect a dataset. At specific intervals (i.e., retardations), the intensity values for all of the pixels are measured, resulting in the collection of a complete interferogram for each pixel in the array (Snively *et al.*, 1999). Subsequent Fourier transformation of all of the interferograms produces an array of infrared spectra which are spatially resolved (Snively *et al.*, 1999). The pixel size is diffraction limited and therefore the resolution is typically in the 4 - 10 µm range in the mid infrared region.

1.4.3 Applications of FT-IR spectroscopy

The principal applications of FT-IR spectroscopy in the study of complex biological systems have been in the field of microbiology. In general FT-IR spectroscopy has been applied to biomedical and industrial studies and has previously been used to generate informative metabolic fingerprints from biological material (Johnson *et al.*, 2004, Winder *et al.*, 2007). Previous studies have also proved its applicability to biological systems by successfully demonstrating the discrimination of bacteria to sub-species level (Winder *et al.*, 2006, Timmins *et al.*, 1998, Naumann *et al.*, 1991a). The high-throughput nature of the technique combined with the minimal sample preparation has proven the technique ideal for the rapid identification of clinically significant bacterial isolates such as, *Eubacterium* associated with oral infections (Alsberg *et al.*, 1998), and the differentiation between *Candida* isolates (Timmins *et al.*, 1998, Udelhoven *et al.*, 2000). FT-IR spectroscopy has also been applied to the field of disease pattern recognition (Petrich *et al.*, 2000, Staib *et al.*, 2001) and to the grading of lymphoid tumours (Andrus and Strickland, 1998). In addition, FT-IR spectroscopy has been successfully employed to monitor industrial fermentation processes for the quantification of metabolite production (McGovern *et al.*, 2002, Schuster *et al.*, 2001). The technique has also been employed in the field of plant science, with application to the discrimination and classification of plants, including the study of *Arabidopsis* cell wall mutants (Wang *et al.*, 1998, Mouille *et al.*, 2003) and plant responses to abiotic stresses (Johnson *et al.*, 2003, Johnson *et al.*, 2004). In more recent years FT-IR spectroscopy has been successfully applied to environmentally relevant investigations such as the study of microbial survival during extreme starvation (Barton and Northup, 2007, Bullen *et al.*, 2008), the discrimination of mycobacteria (Winder *et al.*, 2006) and the metabolic effect of salinity in both plant (Johnson *et al.*, 2003) and animal (Bussell *et al.*, 2008) species.

The application of FT-IR imaging has been successfully demonstrated in the study of biological (Heraud *et al.*, 2005, Kidder *et al.*, 1997, Lewis *et al.*, 1996) and chemical (Bhargava *et al.*, 1998, Oh and Koenig, 1998) systems and is popular for the study of spatially heterogeneous samples (Snively *et al.*, 1999). The high spatial resolution now available with FT-IR microspectroscopy has allowed investigations of the minute details in leaf tissues (Heraud *et al.*, 2007), the classification of malignant gliomas (Krafft *et al.*, 2007) the study of abiotic perturbations in environmentally relevant microorganisms (Patel *et al.*, 2008) and the investigation of single cells to the sub-cellular level (Lasch *et al.*, 2002, Giordano *et al.*, 2001). Furthermore, FT-IR spectroscopy has been

successfully applied to the investigation of sub-cellular changes in living algal cells induced by nutrient starvation (Heraud *et al.*, 2005).

1.5 Mass Spectrometry

Mass spectrometry (MS) is a powerful detection method and is widely employed in metabolomics investigations. MS is commonly combined with chromatographic techniques such as gas chromatography (GC) and liquid chromatography (LC) and the use of hyphenated-MS instrumentation improves quantification and compound identification (Dunn *et al.*, 2005). Identification of metabolites is possible by spectral library matching with standard compounds, interpretation of mass spectra or determination of the molecular formula from accurate mass measurements (Dunn *et al.*, 2005).

In GC-MS, compounds are separated by GC and then transferred online to a mass spectrometer for further separation and detection (Kopka *et al.*, 2004). The use of GC-MS combines two highly complementary techniques: GC separates metabolites on the basis of polarity and boiling point and is able to separate metabolites that have nearly identical mass spectra such as isomers, whilst the fragmentation patterns produced by MS can differentiate between co-eluting, but chemically diverse metabolites, which can then be identified by database searching (Kopka *et al.*, 2004). GC-MS is the most widely used technology in metabolomics and has been described as the 'gold' standard (Harrigan and Goodacre, 2003).

1.5.1.1 Metabolite Extraction

The extraction process should be both non-selective and comprehensive; however, each extraction method introduces a level of chemical bias towards a particular group of metabolites. For example metabolites such as carbohydrates have a high solubility and are therefore typically extracted in methanol:water. However, non-polar metabolites such as lipids have low solubilities and are therefore typically extracted in chloroform. Consequently, a range of extraction methods may be considered in order to obtain a comprehensive metabolite profile from a sample (Soga *et al.*, 2003, Winder *et al.*, 2008).

1.5.1.2 Derivatisation

Volatile low-MW metabolites may be sampled and analysed directly. However, the majority of metabolites analysed require chemical derivatisation prior to analysis to provide volatility and thermal stability (Roessner *et al.*, 2000, Dunn *et al.*, 2005). Derivatisation is a two step process: methoximation is carried out in order to convert carbonyl groups to oximes therefore eliminating slow and reversible silylation reactions. Trimethylsilylation is then carried out, during which exchangeable protons are replaced with trimethylsilyl esters. The stability of samples is of concern as residual water within the samples can result in the breakdown of trimethylsilyl esters. However, this effect can be reduced by extensive sample drying (Dunn and Ellis, 2005).

1.5.1.3 Ionisation and Detection

Electron impact (EI) ionisation is the most widely used ionisation method for GC-MS as it is highly reproducible, and provides molecular ion fragmentation. This fragmentation results in the production of a mass spectrum which is indicative of the metabolites structure, therefore aiding the identification of the metabolite (Dunn *et al.*, 2005, Kopka *et al.*, 2004). Typically, three types of detector may be employed in GC-MS couplings, and the sensitivity of detection may be influenced by the mass spectrometer employed (Dunn *et al.*, 2005). Mass detection may be conducted by either single quadrupole detectors (QUAD), ion-trap technology (TRAP), or time-of-flight detectors (TOF) (Kopka *et al.*, 2004). QUAD detectors may be employed in single ion monitoring mode in order to enhance sensitivity. This method also generally has a wider dynamic range than TRAP or TOF instruments. However, information regarding the metabolites present within a sample is required and so this method may introduce bias into metabolic profiling studies (Dunn *et al.*, 2005). TOF detectors do not require such information regarding the sample, are high-throughput and can provide good sensitivity (Dunn *et al.*, 2005, Kopka *et al.*, 2004).

1.5.1.4 Deconvolution

The chromatograms produced by GC-MS are complex, containing hundreds of metabolite peaks and are further complicated by multiple derivatisation products (Dunn and Ellis, 2005). The application of retention indices (RI) aids metabolite identification by the alignment of chromatograms. Deconvolution software acts by using pure mass spectra to define chromatographic peaks, including overlapping peaks. It also

potentially allows reductions in run time, from approximately 60 minutes to less than 15 minutes (Dunn *et al.*, 2005, Roessner *et al.*, 2000, Weckwerth *et al.*, 2004), as full chromatographic separation is not required, and, in many metabolomics analyses, is difficult or impossible to achieve (Dunn *et al.*, 2005). Metabolites are then identified by matching their chromatographic retention times and mass-spectral fragmentation patterns to known information available in databases (Wagner *et al.*, 2003).

Targeted GC-MS analyses can provide absolute quantification of the level of a given metabolite in a concentration range of up to four orders of magnitude, provided that appropriate external and internal standardisation have been carried out (Dunn *et al.*, 2005). However, each step during extraction, preparation and analysis can introduce general and substance-specific losses and these can vary with the biological material. Current methods are capable of detecting trisaccharides, steroids, diglycerides and some monophosphorylated metabolites (such as glycerol 3-phosphate and glucose 6-phosphate), but most polyphosphorylated and activated metabolic intermediates are presently not accessible to GC-MS analyses. Another major limitation of GC-MS is that many of the peaks are still unidentified. Some of these unidentified peaks may be analytes generated during extraction or by fragmentation in the MS step, but others may be significant and even novel metabolites (Kopka *et al.*, 2004). However the development of mass spectral libraries and databases aims to improve the chemical identification of these peaks (Allwood *et al.*, 2006, Brown *et al.*, 2009, Kopka *et al.*, 2005, Wishart *et al.*, 2009).

1.6 Multivariate Data Analysis

Metabolomics experiments generate extensive amounts of multivariate data, and the differences between samples are not generally discernable by visual inspection. In addition, only a few of these data points may be required in order to describe the problem sufficiently (Goodacre *et al.*, 2004). It is therefore essential to extract the most meaningful elements of these data for the generation of new knowledge in a robust and interpretable manner (Goodacre *et al.*, 2004, Eriksson *et al.*, 2004). Multivariate statistical analyses are an important tool for the analysis of complex spectral information and are used to extract the maximum information from these datasets. Pattern recognition (PR) strategies may be employed for the reduction of the complexity of the datasets produced in metabolomics investigations, and in the generation and testing of novel hypotheses (Goodacre *et al.*, 2004). The use of PR algorithms may allow the identification and interpretation of non-random behaviour in a complex system, which may be obscured by noise or random variations within a biological system (Lindon *et al.*, 2001a). There are a number of different pattern recognition strategies employed in the analysis of metabolomic data, many of which are based on *unsupervised* methods such as principal components analysis (PCA) (Lindon *et al.*, 2001a, Manly, 1994). This multivariate approach uses clustering methods to assess the similarity in a dataset on the basis of their metabolite profiles (Goodacre *et al.*, 2004). In addition, a number of *supervised* methods may be employed in pattern recognition. Methods such as discriminant function analysis (DFA), principal components regression (PCR), partial least squares regression (PLSR), genetic algorithms (GAs) and artificial neural networks (ANNs) seek to transform multivariate data extracted from metabolite profiles into something of biological interest under the guidance of a "teacher" (Goodacre *et al.*, 2004, Manly, 1994, Martens and Naes, 1989). Within these techniques there are two types of data; inputs which are explanatory variables, and targets (or outputs). Supervised learning aims to find a "model" in which these data are correctly associated (Goodacre *et al.*, 2004).

1.6.1 Data Pre-processing

In metabolomics investigations the multivariate data are generally arranged in a matrix in which each row corresponds to a sample and each column corresponds to one

variable (e.g. metabolite, or wavenumber). The matrix, termed **X**, has N rows (samples) and K columns (variables), (Figure 1.7).

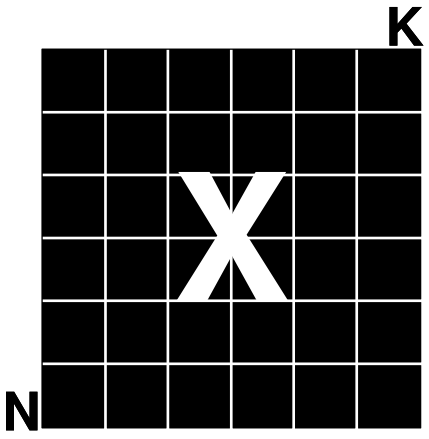


Figure 1.7: A visual representation of a dataset or matrix **X** (input data), employed in multivariate analyses. The dataset **X** consists of **N** rows and **K** columns. Each row in the matrix corresponds to one sample and each column to one variable.

Pre-processing of multivariate data prior to statistical analysis is often required in order to improve the reproducibility of the data. The variables within the dataset are often mean centred, in which the average of each variable is subtracted. In addition, the variables may be scaled to unit variance such that each variable has an equal chance of influencing subsequent statistical analyses. The pre-processing methods employed may be data dependant, and therefore a vast range of methods may be investigated for any one experiment. There are numerous pre-processing methods which may be employed in metabolomics data analyses and many of these methods are order-independent (Goodacre *et al.*, 2007). The methods most commonly employed in the pre-processing of multivariate datasets are detailed in Table 1.3 (Brown *et al.*, 2005, Lindon *et al.*, 2001a).

Table 1.3: Methods commonly employed in the pre-processing of multivariate datasets. Adapted from Goodacre *et al.* (2007).

Term	Description
Normalisation	An operation performed within or across rows in order to make the row profiles comparable in size.
Centering	An operation performed across the rows in order to translate the centre of gravity of the dataset (Keun *et al.*, 2004).
Mean-centering	A method which is commonly employed for centering. Each column is expressed in deviations from its mean, across the rows.
Scaling	An operation which is performed within a column in order to make the column profiles more comparable.
Autoscaling	A type of scaling which mean-centres each value of the column and then divides the row entries of a column by the standard deviation within the column.
Range scaling	A scaling method that mean-centres prior to dividing the row entries of a column through the range within the column (van den Berg *et al.*, 2006).
Pareto scaling	A scaling method that mean-centres prior to dividing the row entries of a column through the square root of the standard deviation within the column.
Transformations	Transformations that linearise or change the scale of the data. These could be Log, Square Root, Box-Cox.
Missing Values	Data in the table which are not available for analysis.
Outliers	Data points which deviate from the distribution of the majority of the data. These data points may be samples, variables or a specific combination of both.

1.6.2 Principal Component Analysis

Principal component analysis (PCA; (Jolliffe, 1986, Wold *et al.*, 1987)) is an unsupervised multivariate projection method employed as a data reduction strategy in which principal components (PCs) are derived that explain the majority of the variance within a complex dataset. The aim of the data compression is to obtain a set of novel PCs which describe as much of the variance in the original dataset as possible. Reducing the dimensionality of the data allows easier interpretation of the variation within a dataset, so that groups, trends and outliers may be identified within the samples. In many cases correlations between variables (e.g. metabolites) occur because they change according to some systematic underlying common factor e.g. a genetic modification. PCA has the ability to detect these underlying factors and compress the information based upon them.

PCs are new variables which are created from linear combinations of the starting variables with the appropriate weighting coefficients (Nicholson *et al.*, 1999). Each PC is orthogonal (uncorrelated) with all the other PCs and consists of one score vector **T** and one loading vector **P** (Figure 1.8). The score vector is considered as the new variable and the loading vector may be described as the link between the score vector and the original variable. The scores are linear combinations of the original variables within the dataset, and the loadings represent the influence (i.e. weighting) of the original variables on the scores.

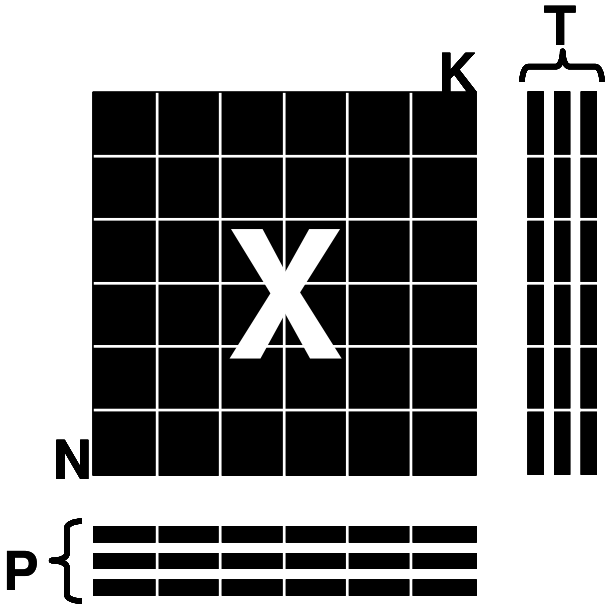


Figure 1.8: PCA compresses the information in the data matrix **X** into scores **T** and loadings **P**, describing the underlying systematic structures of **X**.

PCs are commonly calculated using one of two algorithms; the Non-Iterative Partial Least Squares (NIPALS) algorithm (Wold, 1966), in which the PCs are calculated sequentially; or by Singular Value Decomposition (SVD) (Jolliffe, 1986) in which all of the possible PCs are computed in one step. The length of a PC is termed the eigenvalue and is proportional to the amount of variation the PC describes. The first PC contains the greatest variance, with subsequent PCs containing correspondingly smaller amounts of variance and at some point the PCs will consist only of data noise (Nicholson *et al.*, 1999, Lindon *et al.*, 2001a).

PCA models are both linear and additive; as more PCs are used more of the variance within the dataset is explained. A plot of the first two or three PCs will generally provide maximum information content of the data in two dimensions (Jolliffe, 1986, Wold *et al.*, 1987). Thus, PCA offers a rapid method for visualising and comparing the variance within a dataset. Furthermore, PCA is an unsupervised method which does not require *a priori* knowledge of the dataset structure and therefore models the total variance within a dataset.

Scores plots may provide insight into the relationships between samples within a dataset and can be obtained by plotting the score values for different PCs against each other. In addition outliers within a dataset (i.e. data points which deviate from the distribution of the majority of the data. See Table 1.3) may be identified from PC scores plots and removed from the dataset prior to additional analyses. Plotting the corresponding loadings may illustrate the variables responsible for the trends observed in these scores plots.

1.6.3 Discriminant Function Analysis

Discriminant function analysis (DFA; (Manly, 1994)) is a supervised projection method which is also known as canonical variates analysis (CVA). DFA generates a number of linear discriminant functions by finding the eigenvalues and eigenvectors of the expression:

$$W^{-1}B \qquad \text{(Equation 1.1)}$$

Where W is the within group matrix of sums of squares and cross products, and B is the between group matrix of sums of squares and cross products.

The DFA algorithm cannot analyse datasets which contain colinear variables or too many variables, and it is therefore necessary to perform a data compression such as PCA, prior to the analysis of a multivariate dataset, and the combination is referred to as PC-DFA (Goodacre *et al.*, 1998). *A priori* knowledge of class structure is employed to achieve the supervision of the algorithm. Each group of samples is mean centred to each of the group means, from which the within group matrix of sums of squares and cross products (W) is calculated. In order to determine the corresponding between group matrix (B) the algorithm calculates the difference between W and the total sum of squares and cross products. These matrices are employed by the algorithm in Equation 1.1 in order to determine the Fisher (*F*) ratio (the ratio of between-group to within-group variance). The linear combination of discriminant functions that maximise *F* for the dataset is then sought by the algorithm.

The *a priori* class structure employed by the DFA algorithm may be based on technical experimental replicates, which do not bias analyses designed to observe or model the class structure of biological replicates in an experiment (e.g. for the classification of unknown microorganisms) (Jarvis and Goodacre, 2004b). Alternatively, for experiments designed to observe or model the class structure according to other parameters (e.g. exposure to abiotic perturbation) the *a priori* class structure may be based on the biological replicates (Lopez-Diez *et al.*, 2005).

As DFA is a supervised technique, appropriate validation (Section 1.6.6) methods need to be employed in order to ensure that the quality of the data is of a high standard, and that the subsequent conclusions drawn from the data are valid.

1.6.4 Principal Component Regression

Principal component regression (PCR; (Gemperline *et al.*, 1991)) is a supervised linear regression method the purpose of which, is to provide an estimation of the values of a response variable (**Y**; e.g. API concentrations) on the basis of selected PCs of the explanatory variables (**X**; e.g. FT-IR spectra). In this approach PCA is first carried out on the dataset and the PC scores are then regressed against the response variables (e.g. concentration of abiotic perturbation).

The explanatory variables in multivariate datasets are often highly correlated (multicollinearity) and this may lead to the inaccurate estimation of the least squares

regression coefficients. Such difficulties are avoided by decomposition of the data matrix (**X**) into latent variables by PCA, as PCs are *uncorrelated*. Furthermore, the use of PCs reduces the dimensionality of the regressors as a subset of PCs is typically employed in the regression. This subset usually consists of those PCs which account for the highest explained variance within the dataset; however, the PCs which account for smaller amounts of variance may also be significant for some analyses (Jolliffe, 1982). Moreover as PCA is used it is possible that the largest variation in the data is not directly correlated to the **Y** variable, so alternative regression methods may need to be employed.

1.6.5 Partial Least Squares Regression

The partial least squares regression (PLSR; (Martens and Naes, 1989)) method provides another biased (supervised) regression approach with improvements over PCR (Geladi and Kowalski, 1986, Höskuldsson, 1995, Wold *et al.*, 2001a, Wold *et al.*, 2001b, Geladi, 1988, Wold, 1984, Jolliffe, 1986). It has been shown that PLSR is comparable to finding successively linear functions of the predictor variables that have maximum covariance with the dependent variables (Jolliffe, 1986, Stone and Brooks, 1990). This is reliant upon each of the linear functions being uncorrelated with the previous linear functions (Stone and Brooks, 1990). In contrast to the PCR technique which models the variances derived from **X**, and least squares regression techniques which maximise the correlations between **Y** and **X**, PLSR combines both correlation and variance to model *covariance* between the predictor matrix (**X**) and the response matrix (**Y**) (Jolliffe, 1986). PLSR is particularly useful for the prediction of a set of dependent variables from a large set of independent variables, and can deal efficiently with multivariate datasets that are highly correlated and contain noise. PLSR allows the relationships between different blocks of data to be derived using modelling with "latent variables" which are derived by an iterative procedure. These latent variables are mutually orthogonal (uncorrelated) within each data block and are linear combinations of the original variables (Martens and Naes, 1989, Lindon *et al.*, 2001a).

PLSR employs the latent variables (i.e. scores) of **X** to explain the variance in both **X** and **Y** (whereas PCA only models the variance in **X**). Consequently, the aim is also different in the calculation of the latent variables in PLSR and the variance in **X** which is needed to predict the variance in the response variable (**Y**) is extracted (See Figure 1.9). The first score (**T**1) is calculated by the algorithm in order to maximise the

covariance between the **T**1 score vector and **Y**. The weight vector **W**1 contains the influence of the original variable in **X** on the **T**1 score. Information regarding the variance described by the **T**1 score vector is contained in the **P**1 and **C**1 loadings of **X** and **Y** respectively. The variance explained by the first score vector is removed from the **X** and **Y** matrices by the subtraction of **T**1**P**1 and **T**1**C**1 from **X** and **Y** resulting in the formation of the residuals **E** and **F** respectively. These residuals replace the original **X** and **Y**, and the method is repeated until the optimal number of components has been calculated.

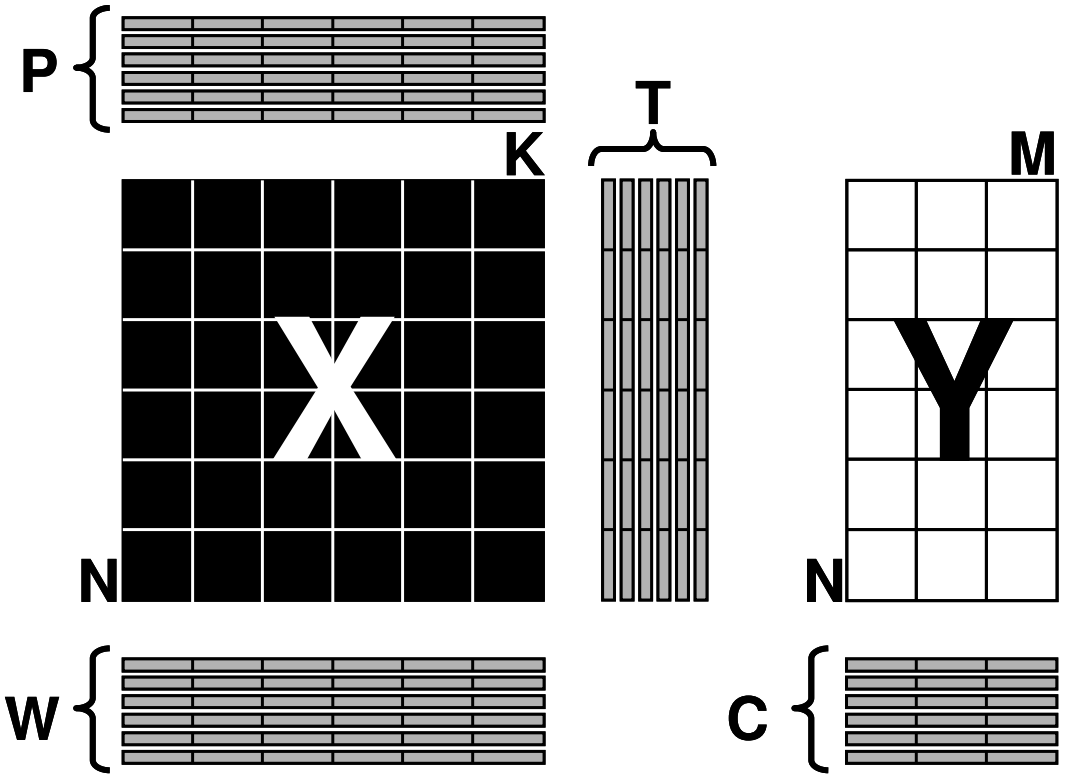


Figure 1.9: PLSR combines both correlation and variance to model covariance between the predictor matrix (**X**) and the response matrix (**Y**).

The number of significant components in a PLSR model can be estimated through the use of validation techniques (Wold, 1978). PLSR models may be validated through either internal or external techniques. Cross validation is commonly used for internal validation (Wold, 1978) and may be employed to estimate the number of model components and to estimate the predictive capability of the model. For external validation techniques, the PLSR model is used to predict a test set of samples with known responses (**Y**) and the accuracy of the predictions is assessed based upon the size of the root mean square error of prediction (RMSEP). For each response variable

in **Y** a RMSEP value is obtained and is presented in the same units as the response (e.g. µg mL^{-1} for concentration based studies). In external validation techniques, samples that have not been employed to build the model are predicted, therefore providing an independent measure of the predictive ability of the model. Consequently external validation is considered the best method for PLSR model validation (Broadhurst and Kell, 2006).

1.6.6 Model Validation

As highlighted for PLSR, model validation is an essential step in multivariate data analysis. It is carried out in order to ensure that the quality of the model is of a high standard, and that the subsequent conclusions drawn from the data are valid. Without the application of appropriate model validation techniques, models may over-fit the data, finding relationships between the data and the dependent variables that do not hold for subsequent analyses (Picard and Cook, 1984); that is to say, there is a lack of generalisation.

A common method of validation is to split the data into three sets; a training set which is used for training the model, a validation set which is used to avoid over training of the data, and a test set which is used to test the ability of the model to generalise. Careful consideration should be taken when splitting the data as the groups should represent the spread of the variance within the dataset (Brown *et al.*, 2005, Picard and Berk, 1990). For techniques such as PCR, PC-DFA and PLSR the models are generated using the training dataset and the test set is then projected into the same ordinate space. The optimum number of latent variables or PCs may then be selected for the model where the validation data falls within the bounds of the training and test datasets (Handl *et al.*, 2005, Kaderbhai *et al.*, 2003).

1.7 Research Objectives

There are a number of classes of environmental pollutants of which APIs - with specific focus on those which contain a chiral centre - are just one. Within recent years the importance of the interactions of these compounds within the aquatic environment has been realised (Escher *et al.*, 2005, Fent *et al.*, 2006, Heberer, 2002). Information regarding the fate and biodegradation of such environmental pollutants is now required.

The intention of this study is to investigate the effects of chiral APIs and other environmental pollutants on the metabolome(s) of environmentally relevant microorganisms (both axenic cultures and communities) and the metabolic pathways in which these pollutants undergo degradation. The effects of these pollutants will be tested in both prokaryotic and eukaryotic systems, and a variety of technologies will be employed, including FT-IR spectroscopy and GC-MS. The chemical analyses of these systems will be combined with multivariate analysis methods in order to model the metabolic effect of chiral APIs and other environmental pollutants on microbial species.

A combination of multivariate methods will be employed for the effective analysis of the analytical data obtained, in order to provide classification of the samples. In addition it is intended that the spectral and chromatographic regions important in the discrimination and quantification of the metabolic effects may be identified in order to fully appreciate the effects of the pollutants on the microbial systems. Furthermore, any possible degradation products will be identified.

In particular the environmental metabolomics investigation presented here will investigate the following areas;

1. The growth dynamics of a range of environmentally relevant bacterial species will be monitored following exposure to a range of concentrations of the chiral APIs Atenolol and Propranolol. FT-IR spectroscopy - combined with the relevant multivariate analysis methods - will be employed as a metabolic fingerprinting tool in order to monitor any phenotypic effects exerted on the bacterial species at a range of concentrations and enantiomeric ratios. In addition, GC-MS will be employed for metabolic profiling of the observed of chiral specific effects.

2. The potential chiral specific effects of the APIs Atenolol and Propranolol will be investigated in a eukaryotic system. FT-IR microspectroscopy will be employed in order to monitor the general phenotypic effects and more specific localised effects of a range of enantiomeric ratios of the APIs at a given concentration in the green alga *Micrasterias hardyi*. In addition a wide range of multivariate analysis methods will be employed to interrogate the hyperspectral data.

3. The phenotypic effects of the chiral APIs Atenolol and Propranolol and the aromatic hydrocarbons toluene and phenol on a complex activated sludge community will be investigated. FT-IR spectroscopy will be employed as a phenotypic fingerprinting technique in order to model the phenotypic differences in the community following exposure to the various environmental contaminants. In addition, any possible degradation of these compounds will be monitored with the use of HPLC or a colorimetric assay.

4. Biochemical changes in the activated sludge community will be monitored during the degradation of phenol over a 48 h period through the use of FT-IR spectroscopy. In addition the phenotypic effect of storing the activated sludge for extended periods of time (without added phenol), will be determined by FT-IR analysis in order to understand the ability of this microbial community to degrade phenol. In addition, GC-MS will be employed for metabolic profiling of the activated sludge community in order to monitor the changes in the meta-metabolome during the active degradation of phenol.

2 Monitoring the effect of chiral pharmaceuticals on aquatic microorganisms by metabolic fingerprinting and metabolite profiling

The work presented in this chapter has been published:

E. S. Wharfe, C. L. Winder, R. M. Jarvis and R. Goodacre (2010) Monitoring the Effects of Chiral Pharmaceuticals on Aquatic Microorganisms by Metabolic Fingerprinting. *Applied and Environmental Microbiology* **76,** 2075-2085.

2.1 Abstract

The effects of the chiral pharmaceuticals Atenolol and Propranolol on *Pseudomonas putida*, *Pseudomonas aeruginosa*, *Micrococcus luteus*, and *Blastomonas natatoria* were investigated. The growth dynamics of exposed cultures were monitored using a Bioscreen instrument. In addition, Fourier-transform infrared (FT-IR) spectroscopy with appropriate chemometrics and high-performance liquid chromatography (HPLC) were employed in order to investigate the phenotypic changes and possible degradation of the drugs in exposed cultures. For the majority of the bacteria studied there was not a statistically significant difference in the organism's phenotype when it was exposed to the different enantiomers or mixtures of enantiomers. In contrast, the pseudomonads appeared to respond differently to Propranolol, and the two enantiomers had different effects on the cellular phenotype. Furthermore, GC-MS metabolic profiling results corresponded with these findings, implying that there were different metabolic responses in the organisms when they were exposed to the different enantiomers. We suggest that our findings may indicate that there are widespread effects on aquatic communities in which active pharmaceutical ingredients are present.

2.2 Introduction

Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment (Escher *et al.*, 2005), and the occurrence of APIs in the aquatic environment is of growing concern (Fent *et al.*, 2006). There are a number of routes through which APIs and resulting metabolites / degradation products may enter these ecosystems, and a common avenue is through the excretion of the APIs and their metabolites in urine and faeces. It is known that APIs have different rates of metabolism within man. For example, the β-blocker Propranolol is almost completely metabolised within the liver and only 1-4% of an oral dose is excreted as the unchanged API and its metabolites. By contrast, 40-50% of an oral dose of Atenolol (also a β-blocker), is excreted as the API or its constituent metabolites (Ashton *et al.*, 2004, Carlsson *et al.*, 2006a, Carlsson *et al.*, 2006b). The subsequent degradation of the APIs and their metabolites may also occur at sewage treatment plants (STPs); this is usually substrate specific and will vary greatly between APIs. The rate of adsorption to activated sewage sludge during treatment differs between APIs, and is dependant on the hydrophobic and electrostatic interactions of the API with the particulates and microorganisms within the activated sewage sludge (Fent *et al.*, 2006). Any remaining API and relevant metabolites are diluted into the surface water when the effluent is released from the STP. Hence, many APIs are present in low concentrations (ng - µg L^{-1}) in aquatic environments such as rivers, streams and estuaries (Escher *et al.*, 2005, Ashton *et al.*, 2004, Carlsson *et al.*, 2006b). The majority of APIs are neither persistent nor highly bioaccumulative; however, the continuous release of APIs into the aquatic environment poses a potential risk to aquatic organisms even though the concentrations of APIs in receiving waters are quite low (Escher *et al.*, 2005).

Despite little being known about the effects of APIs in the environment, it must be taken into account that they are designed to elicit a specific mode of action in humans (Escher *et al.*, 2005). Adverse side effects may be encountered in humans at higher doses of these APIs and it can be expected that any beneficial or adverse effect may also be observed in aquatic organisms with similar biological functions or receptors. It must also be noted that similar targets may control different metabolic processes in different species (Seiler, 2002) and therefore APIs and their metabolites may act through additional modes of action in aquatic organisms. The effects of the APIs may be subtle due to the very low concentrations observed in the aquatic environment and as a result these effects may go unnoticed (Escher *et al.*, 2005). It is also likely that the

effect of the API will impact on the local population dynamics throughout the whole ecosystem, from bacteria up to higher organisms. To explore the effects of the APIs on biological systems a wide range of concentrations should be employed with the application of appropriate analytical platforms to profile the complement of biochemical components in the cell. Indeed it is known that APIs could become concentrated in the benthic environment of river beds and as bacteria inhabit this niche this community may be exposed to higher levels than expected (Thacker, 2005, Halling-Sørensen *et al.*, 1998, Pouliquen *et al.*, 1992).

Whilst the effects of APIs in the environment is currently a growing area of research there is very little understanding of the environmental effects of chiral pharmaceuticals (Fono and Sedlak, 2005, Buser *et al.*, 1999). A chiral molecule is one that lacks an internal plane of symmetry. The non-superimposable mirror images are termed enantiomers and are labelled '(*R*)' or '(*S*)' according to a priority system (Cahn Ingold Prelog priority rules) based on the atomic number of the molecules substituents. Approximately 56% of the APIs currently in use are chiral compounds, 88% of which are administered therapeutically as the racemate (i.e. an equal mixture of the two enantiomers, represented by the symbol (±)). The chirality of environmental contaminants such as APIs must be taken into consideration in order to fully understand the environmental fate and effects of these compounds. The enantiomers of a chiral API are able to interact differently with other chiral compounds such as enzymes, and therefore potentially exert different effects when released into the environment (Fono and Sedlak, 2005, Buser *et al.*, 1999, Nikolai *et al.*, 2006). It is widely known that the enantiomers of a chiral API may possess differing toxicological and biological effects both from each other, and the racemate (an equal mixture of the two enantiomers) (Lees *et al.*, 2003, Yang *et al.*, 2005). It has been shown that the (*S*)-enantiomers of the β-blocking agents Atenolol and Propranolol are more potent in man than their respective antipodes (Davies, 1990, Pearson *et al.*, 1989, Kurt *et al.*, 1993, Barrett and Cullum, 1968) and that a number of the biotransformation pathways for β-blockers are stereoselective in humans (Mehvar and Brocks, 2001). The mode of action of the drugs and their enantiomers is not known in prokaryotic systems. It is therefore necessary to enhance our understanding of the fate and biological effects of chiral pharmaceuticals on typical microflora from the aquatic environment in order to fully appreciate the risks (Huggett *et al.*, 2002). Of particular interest is the group of APIs termed β-blockers as they all contain at least one chiral centre and are generally administered therapeutically as the racemate (Mehvar and Brocks, 2001). In addition,

they are widely used with approximately 29 and 12 tonnes of Atenolol and Propranolol respectively consumed each year in the UK (Ashton *et al.*, 2004, Carlsson *et al.*, 2006a, Carlsson *et al.*, 2006b).

In order to explore the effect of the APIs on biological systems we employed Fourier-transform infrared (FT-IR) spectroscopy, this is a phenotyping technique which has previously been used to generate metabolic fingerprints from bacteria (Johnson *et al.*, 2004, Winder *et al.*, 2007). Previous studies have proved successful in the discrimination of bacteria to sub-species level (Winder *et al.*, 2006, Timmins *et al.*, 1998, Naumann *et al.*, 1991a) through the detection of subtle changes in the biochemical phenotype of the bacteria and the use of FT-IR coupled with suitable chemometrics to allow the physiological assessment of bioprocesses has been demonstrated. In addition, the combination of FT-IR and trajectory analysis has proved successful in the identification of metabolic changes in fermentations (Johnson *et al.*, 2004). FT-IR is an automated high-throughput technique (10-60 s per sample is typical); requiring minimal sample preparation, and this makes it relatively inexpensive. It is therefore an ideal screening method to explore the effect of the APIs on a number of bacterial systems.

In this study the chiral specific metabolism of the β_1 – selective adrenergic blocking agent Atenolol and the non-selective β-adrenergic blocking agent Propranolol by a range of environmental microorganisms was investigated (Ternes, 1998, Fono and Sedlak, 2005, Roberts and Thomas, 2006). FT-IR spectroscopy was employed to monitor biochemical changes in the spectral fingerprint of the whole bacterial cells during the growth of these microorganisms in the presence of the selected APIs. The fate of the API was monitored with chiral HPLC, enabling quantification of the enantiomers. In addition, GC-MS was employed in order to monitor the changes in the metabolite profiles of two pseudomonads following exposure to a range of enantiomeric ratios of Propranolol.

2.3 Materials and methods

2.3.1 Cultivation of Bacteria

In order to monitor the effects of the chosen APIs within the aquatic environment a variety of microorganisms were selected for investigation. All of the microorganisms employed within this study have been reported to be commonly found within the aquatic environment, and are amenable to growth in the laboratory. Four bacteria were selected for this study, *Pseudomonas putida* KT2440 which has been known to inhabit freshwater streams and activated sewage sludge (John and White, 1998, Martínez-Bueno *et al.*, 2002), *P. aeruginosa* PA14 which is commonly isolated from freshwater streams (Tan *et al.*, 1999), and *Micrococcus luteus* 2.13 (Rickard *et al.*, 2000), and *Blastomonas natatoria* 2.1 (Sly and Cahill, 1997, Rickard *et al.*, 2000) which have been isolated from freshwater biofilms. The bacteria were cultured in R2A medium (Reasoner and Geldreich, 1985) at 15 °C for 24 h, 200 r.p.m. in a Multitron (INFORS HT, Switzerland) orbital shaker unless otherwise stated. The pure enantiomers (i.e. (*R*)- and (*S*)-) of both Atenolol and Propranolol (as hydrochloride) were purchased from Sigma-Aldrich Company Limited, (Poole, Dorset, U.K.).

2.3.2 Screening of microorganisms for growth in the presence of APIs

The growth of the bacteria was monitored (at 600 nm using a Bioscreen spectrophotometer (Labsystems, Basingstoke, U.K.)) for each microorganism at a range of concentrations (10-130 μg mL^{-1}) of each enantiomer of the API and the racemate. The data collected from these investigations were used to calculate the specific growth and death rates (this is when the rate of cell death or lysis exceeds the rate of growth so that a decrease is observed in the turbidity of the culture (Rice and Bayles, 2003)) from the exponential phases using the following equation:

$$\mu = 2.303 \frac{(\log_{10}OD2 - \log_{10}OD1)}{t_2 - t_1}$$

Where: $\log_{10}OD1$ = $\log_{10}$ value of OD reading at time point 1

$\log_{10}OD2$ = $\log_{10}$ value of OD reading at time point 2

t_1 = time point 1 and t_2 = time point 2

The growth rate data (*vide infra*) was used to select a reduced range of concentrations for further investigation; this was based on identifiable difference in growth rate, but which did not result in cell death.

2.3.3 Batch growth

Bacterial cultures were exposed in triplicate to a range of concentrations of the chosen API, the pure enantiomers and a range of enantiomeric mixtures were used (see Table 2.1). An aliquot (1 mL) of sterile water was added to an additional set of bacterial samples to act as a control. Samples were maintained at 15 °C, 200 r.p.m. in a Multitron orbital shaker for 24 h. Aliquots (2 mL) were taken in triplicate from each flask and centrifuged (5 min, 0 °C, 16089 *g*) to harvest the biomass. The supernatant and pelleted biomass was stored at -80 °C for further analysis.

Table 2.1: Microorganisms and conditions for batch growth.

Microorganism	API	Concentration of API (µg mL^{-1})	Growth Temperature (°C)	Incubation Period (h)
P. putida KT2440	Propranolol	(±): 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120	15	24
	Propranolol	50 µg mL^{-1} of (*R*) : (*S*) in the following ratios: (0, 25:75, 50:50 75:25 100)	15	24
	Atenolol	(±), (*R*), (*S*): 90, 100, 110, 120, 130	15	24
P. aeruginosa PA14	Propranolol	(±), (*R*), (*S*): 40, 50, 60, 70, 80	15	48
	Propranolol	50 µg mL^{-1} of (*R*) : (*S*) in the following ratios: (0, 25:75, 50:50 75:25 100)	15	48
	Atenolol	(±), (*R*), (*S*): 40, 50, 60, 70, 80	15	48
M. luteus 2.13	Propranolol	(±), (*R*), (*S*): 40, 50	15	336
B. natatoria 2.1	Propranolol	(±), (*R*), (*S*): 40, 50	15	336

2.3.4 Quantitative analysis of API concentration with HPLC

Concentrations of Atenolol and Propranolol were determined by HPLC (Agilent 1100 series). The supernatant samples were allowed to thaw at room temperature and were filtered (0.22 µm, Millipore™) in order to remove any microbial cells remaining in the medium. Aliquots (25 µL) were injected onto the HPLC column in a random order. Each sample was injected three times during the analysis, resulting in three analytical

replicates for each biological sample. The HPLC system was equipped with a Chirobiotic V2 column (250 mm × 4.6 mm i.d.), particle size 5 µm (ASTEC, Whippany, NY, USA) and a UV detector operating at a wavelength of 230 nm. The column was eluted with an isocratic mixture of methanol and water (90:10, v/v) and 1.0% triethylamine acetate (TEAA) buffer, pH 5.0. The pH of the buffer was adjusted with acetic acid prior to the addition of methanol. The measurements were carried out a 25 ± 1 °C at a flow rate of 1 mL min^{-1} (Bosakova *et al.*, 2005).

2.3.5 FT-IR spectroscopy

A 96-well zinc selenide (ZnSe) plate was cleaned by rinsing with 2-propanol and deionised water (three times) and allowed to dry at room temperature (Winder *et al.*, 2006, Harrigan *et al.*, 2004). The cell pellets stored at -80 °C were allowed to thaw at room temperature and washed in order to remove any traces of residual API. Ice cold sterile water (2 mL) was added to each sample and gently vortexed. The samples were centrifuged for 10 min (0 °C, 16089 *g*), and the supernatant was discarded; this cycle was repeated 3 times. A final aliquot of 100 µL sterile water was added to each sample and the solution was vortexed. Aliquots (20 µL) of each resuspended sample were applied to the ZnSe plate and oven dried at 50 °C for 10 min. Drying was used to minimise any signal arising from the absorption of water in the mid-IR region, which would mask the biologically important chemical information in the spectra. Three replicates of each of the samples were randomly applied to the ZnSe plates and triplicate spectra were obtained from different positions of each well, a total of nine spectra (so called technical replicates) were collected per sample. The plate was loaded onto a motorised microplate module HTS-XT under computer control by the OPUS software version 4 (Winder *et al.*, 2006). Spectra were collected using an Equinox 55 FT-IR spectrometer (Bruker Optics Ltd), in transmission mode using a deuterated triglycine sulphate (DTGS) detector over the wavelength range of 4000-600 cm^{-1} and with a resolution of 4 cm^{-1}. 64 spectra were co-added to improve the signal-to-noise ratio. The spectra are displayed in terms of absorbance (see Figure 2.1 for typical example spectra).

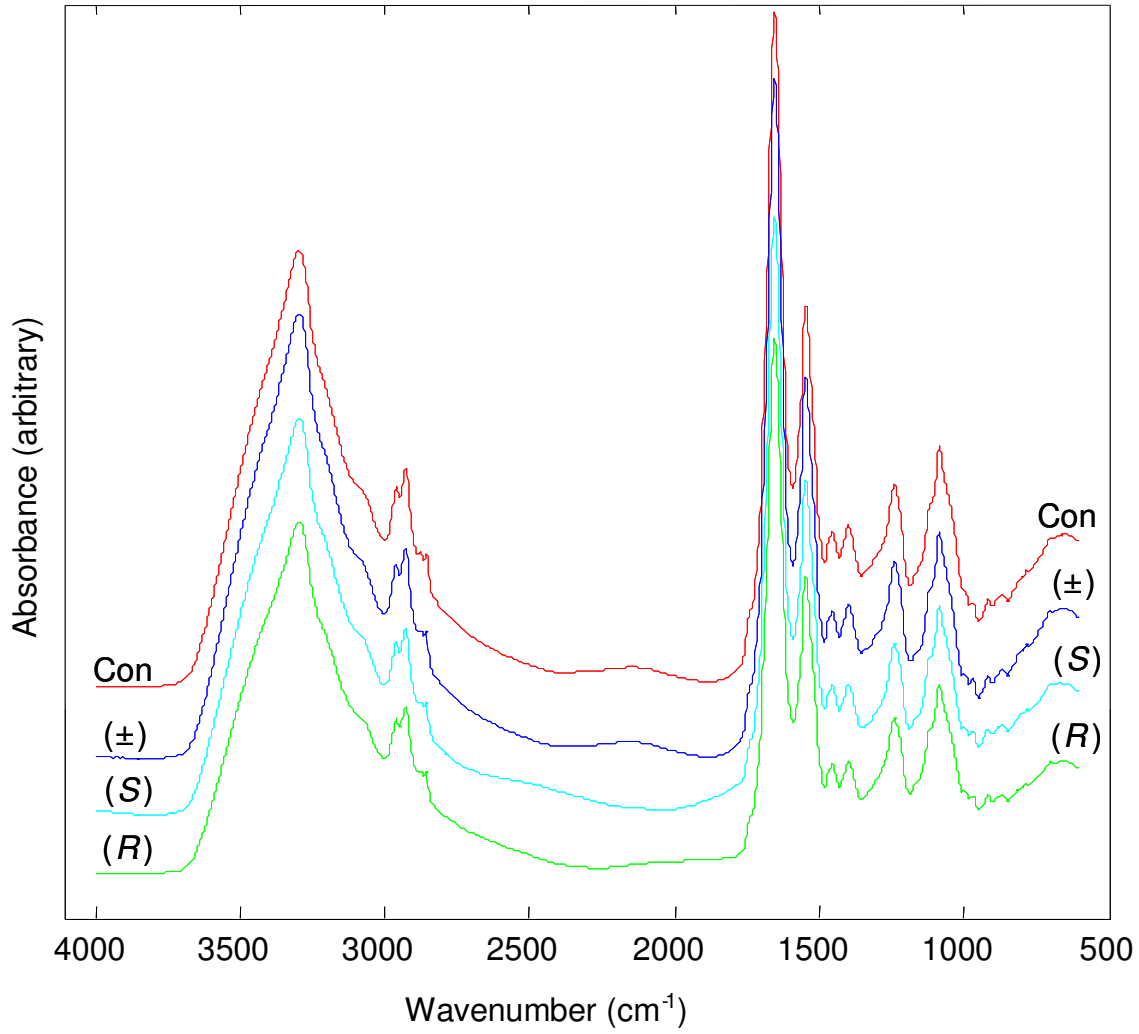


Figure 2.1: Typical processed FT-IR spectra for *Pseudomonas aeruginosa* PA14 exposed to 80 µg mL^{-1} (*R*), (*S*), and (±)- Propranolol. Additional control samples are included which were not exposed to Propranolol. The spectra are offset for clarity.

2.3.6 FT-IR Spectral pre-processing

The ASCII data were imported into Matlab version 7.1 (The MathWorks, Inc., Natick, MA, USA) and as an initial step spectral regions which are dominated by CO_2 vibrations arising from the atmosphere (2403 – 2272 cm^{-1} and 683 – 656 cm^{-1}) were removed and filled with a linear trend. The spectra were corrected using extended multiplicative scatter correction (EMSC) which normalises and smoothes the spectra by application of a polynomial smoothing function (Martens *et al.*, 2003). These pre-processed spectra were used for subsequent multivariate analyses.

2.3.7 Metabolite profiling with GC-MS

In order to quench metabolism, culture samples (20 mL) were plunged rapidly into an equal volume of 60% aqueous methanol solution (–48 °C). The quenched biomass was centrifuged for 10 min (3000 *g*, –9 °C) and the supernatant was removed. The biomass pellets were stored at –80 °C for further analysis (Winder *et al.*, 2008).

For metabolite extraction, the biomass pellets were suspended in 550 µL of 80% aqueous methanol solution (–48 °C), frozen in liquid nitrogen, and allowed to thaw on dry ice. The freeze thaw cycle was carried out three times in order to permeabilise the cells and allow leakage of the metabolites into the supernatant. The suspensions were centrifuged for 7 min (16,060*g*, –9 °C) and the supernatant was retained and stored on dry ice. An additional aliquot (550 µL) of 80% aqueous methanol solution was added to the biomass pellet and the procedure was repeated. The second aliquot of the supernatant was combined with the first and the sample was stored on dry ice for further analysis (Winder *et al.*, 2008). In order to prepare the metabolite extracts for GC-MS analysis, aliquots (1 mL) of each extract was spiked with 100 µL of internal standard solution (0.19 mg mL^{-1} succinic-d_4 acid, 0.27 mg mL^{-1} malonic-d_2 acid, 0.22 mg mL^{-1} glycine-d_5 in HPLC-grade water) and lyophilised (HETO VR MAXI vacuum centrifuge attached to a HETO CT/DW 60E cooling trap; Thermo Life Sciences, Basingstoke, U.K.) (Winder *et al.*, 2008).

Samples were subsequently derivatised in two stages. An aliquot (40 µL) 20 mg mL^{-1} *O*-methylhydroxylamine solution in pyridine was added and heated at 40 °C for 80 min followed by addition of 40 µL of MSTFA (*N*-acetyl-*N*-(trimethylsilyl)-trifluoroacetamide) and heating at 40 °C for 80 min. A retention index solution was added for chromatographic alignment (20 µL, 0.6 mg mL^{-1} $C_{10}/C_{12}/C_{15}/C_{19}/C_{22}$ *n*-alkanes).

The gas chromatography time-of-flight mass spectrometry (GC-TOF-MS) method used to analyse the metabolites has been shown to be suitable for detection and semi-quantification of a wide range of metabolite classes as detailed in available mass spectral / retention index libraries (Kopka *et al.*, 2005) or genome-scale reconstructions of metabolic networks (Feist *et al.*, 2007). Samples were analysed in a random order employing GC-TOF-MS (Agilent 6890 GC coupled to a LECO Pegasus III TOF mass spectrometer) and using the optimal settings determined previously for yeast analysis (O'Hagan *et al.*, 2004). The raw data were processed using LECO ChromaTof v2.12

and its associated chromatographic deconvolution algorithm, with the baseline set at 1.0, data point averaging of 3, and average peak width of 2.5. A reference database was prepared which incorporated the mass spectrum and retention index of all metabolite peaks detected in a random selection of samples in order to allow the detection of all metabolites present. Each metabolite peak in the reference database was searched for in each sample, and if matched, the peak area was reported and the response ratio relative to the internal standard was calculated (Winder *et al.*, 2008).

2.3.8 Multivariate Analysis

Multivariate analysis followed a protocol developed previously (Goodacre *et al.*, 1998, Alsberg *et al.*, 1998). Principal components analysis (PCA) is an unsupervised method for reducing the dimensionality of multivariate data whilst preserving the variance. This transformation was performed prior to canonical variates analysis (CVA). CVA is a supervised learning method that seeks to minimise within group variance whilst maximising between group variance, and can be used in conjunction with PCA to discriminate between groups on the basis of retained principal components (PCs), given *a priori* knowledge of group membership of the spectral replicates (Manly, 1994, Winder *et al.*, 2004). In this study, PC-CVA models were constructed with *a priori* knowledge of the biological replicates. In order to make sure these models were not over- or under-trained, validation was performed using the method of full cross-validation, where two of the biological replicates were used for model training with the third projected into the model for cluster validation purposes (Jarvis and Goodacre, 2004a). Finally, CVA also allows statistical significance to be displayed on the scores plots and circles were used to represent the 95% χ^2 confidence region constructed around each group mean based upon the χ^2 distribution with two degrees of freedom (Krzanowski, 1988).

Partial least squares (PLS (Martens and Naes, 1989)) regression is a multivariate linear regression method which allows the quantitative relationship between different variables (i.e., API concentration and FT-IR spectra) to be modelled, and can deal efficiently with datasets that are highly correlated. In this study, PLS regression was employed to predict the API concentration values from the FT-IR data. As for PC-CVA, the regression models were calibrated with two of the three biological replicates and the third replicate was used as an independent test set to validate the model and establish whether the models could generalise.

2.4 Results and Discussion

2.4.1 The effect of the chiral APIs on the bacterial growth rates

A number of aquatic microorganisms were exposed to the chiral APIs Atenolol and Propranolol, and growth rates, death rates and maximum optical density (OD) were determined to monitor the effect of the APIs on culture progression (see Figure 2.2 for examples of 0, 10, 50, 90 and 130 µg mL^{-1} for (*R*)-, (*S*)- and the racemic mix). Slight variations were observed in the specific growth rates within the *Pseudomonas* species exposed to varying concentrations (10-130 µg mL^{-1}) of (*R*)-, (*S*)- and (±)- Propranolol. There was a considerable difference in the growth rates observed between species, and a marked effect was observed in *P. aeruginosa* PA14 exposed to Propranolol. By contrast, minimal changes were detected in the growth rates, death rates or maximum amount of biomass of both organisms exposed to 10-130 µg mL^{-1} of (*R*)-, (*S*)- and (±)- Atenolol.

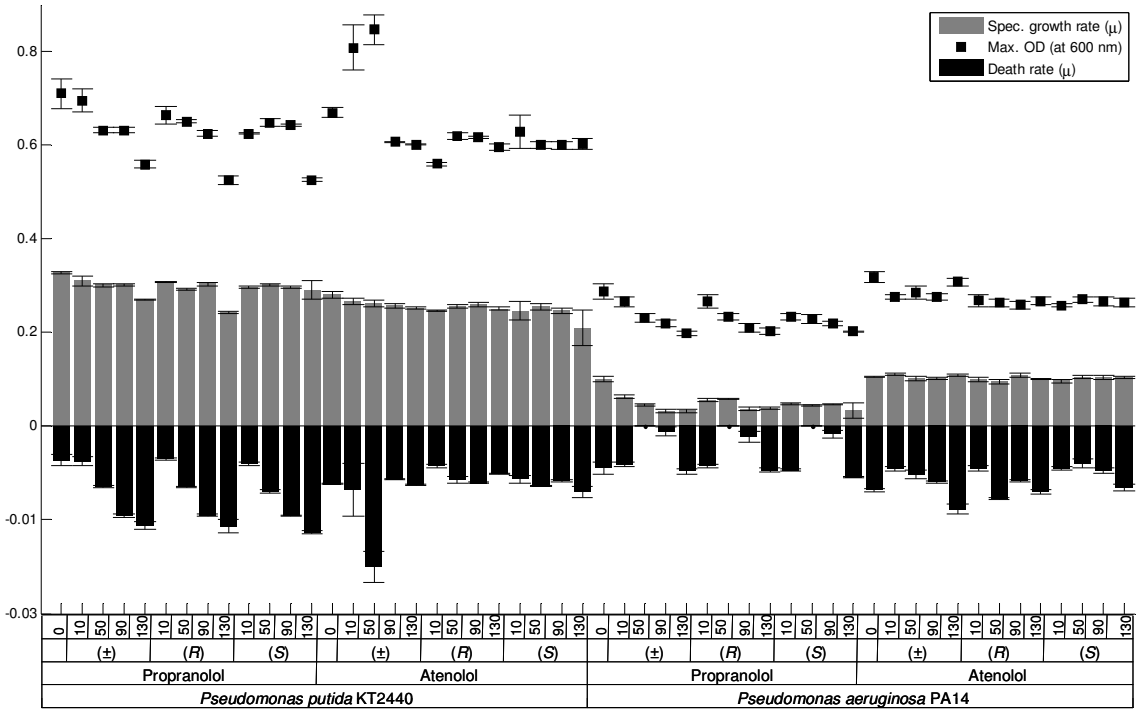


Figure 2.2: Specific growth rate data for *Pseudomonas putida* KT2440 and *P. aeruginosa* PA14 exposed to 0-130 µg mL^{-1} of Propranolol or Atenolol. Maximum OD (600 nm) and "specific death rates" are also shown. Data are averaged from 5 biological replicates and error bars show standard deviation.

An interesting effect was observed for *P. aeruginosa* PA14 exposed to both of the Propranolol enantiomers and the racemate. At concentrations of 50-70 μg mL^{-1} there appeared to be no death of the microbial cells. By contrast, for cells exposed to 10-40 and 80-130 μg mL^{-1} the death rate was equivalent to that from the control cells. This is likely to be due to the lower concentrations (<40 μg mL^{-1}) of Propranolol having very little effect on metabolism so cells quickly reach stationary and death phases, higher concentrations (>80 μg mL^{-1}) having a negative impact on metabolism and killing cells (as also seen in the final turbidity measurements being significantly lower than control cells), whilst the intermediate concentrations (50-70 μg mL^{-1}) might slow growth but the cells have not yet entered the death phase. Inspection of the growth curves indicated a second phase of growth several hours into the stationary phase. Whilst this may be indicative of the utilisation of a secondary carbon source this was not observed in the control cultures and as such is not the likely explanation. The biomass of the culture decreased as the concentration of the API increased and thus the original carbon source was potentially not depleted at the onset of stationary phase. Whilst at the higher concentrations of Propranolol a slight increase in the biomass was immediately followed by a noticeable decrease in the OD of the culture (death phase) the maximum biomass was severely inhibited by the presence of the API. Our observations would suggest that a differential effect is exerted by the API depending on the concentration applied. At lower concentrations the growth trend is unaffected by the API, and at high concentrations a death is observed during the growth period. However, at intermediate concentrations the production of biomass is extended throughout the growth period (the onset of death phase may be observed if the growth was monitored for extended periods). The APIs were not metabolised during the growth (Table 2.2) and we would hypothesise that the intermediate concentrations of Propranolol were either affecting the transport of nutrients into the cell or the rate of metabolism.

Table 2.2: Quantification of Propranolol from HPLC data for bacterial cells exposed to varying ratios of (*R*)- and (*S*)- Propranolol at 50μg mL^{-1}.

Microorganism or Medium	Ratio of Enantiomers (*R:S*)[b]							
	100% (*R*)	75:25		50:50		25:75		100% (*S*)
		(*R*)	(*S*)	(*R*)	(*S*)	(*R*)	(*S*)	
Medium[a]	49.6 (1.3)	36.9 (1.2)	11.3 (1.4)	23.8 (1.6)	24.3 (1.2)	11.6 (0.8)	37.2 (1.5)	48.4 (2.1)
P. putida KT2440	48.9 (1.7)	37.1 (0.9)	12.3 (1.1)	24.3 (0.8)	24.7 (1.1)	12.1 (0.7)	36.8 (1.4)	48.8 (1.8)
P. aeruginosa PA14	48.2 (1.9)	36.7 (1.4)	11.9 (1.8)	23.9 (1.9)	24.6 (1.0)	11.9 (1.5)	37.0 (0.9)	49.3 (1.5)

a Control experiments (labelled medium) were performed to determine the experimental effect on the drug concentration (i.e., loss of API during incubation in growth medium).
b The values are averages for five measurements. The values in parentheses are standard deviations.

It can be clearly noted in the growth rate data for the Propranolol exposed cells that the specific growth rate decreases as the concentration of the API increases. This trend is also observed in the maximum optical density data and death rates for both pseudomonads. Our findings indicate that Propranolol has considerably different affects on the two *Pseudomonas* species. These findings are rather surprising as these species are genetically closely related. Estimates show that greater similarity (60% of the predicted coding sequences) is shown between the two pseudomonads than any other completed microbial genomes to date (Nelson *et al.*, 2002). In addition, comparative genome analysis has shown that 85% of the *P. putida* KT2440 genome has homologues in the *P. aeruginosa* PAO1 genome (Stover *et al.*, 2000).

The toxicity effects of the APIs observed here in aquatic organisms have been previously reported. Toxicity studies carried out by Kim and co workers in the crustacean *Thamnocephalus platyurus* and a fish species (*Oryzias latipes*) reported that Propranolol caused acute toxicity in *T. platyurus* at a concentration of 10.61 μg mL^{-1} and in *O. latipes* at a concentration of 11.40 μg mL^{-1}. In contrast to the results presented here, they found that Atenolol did not cause toxicity effects in the aquatic organisms at the concentrations used (<100 μg mL^{-1}) (Carlsson *et al.*, 2009). In addition, toxicity studies have been carried out on a range of APIs (including Propranolol) in the Japanese medaka fish (*Oryzias latipes*) an amphipod (*Hyalella azteca*) and two crustaceans (*Ceriodaphnia dubia* and *Daphnia magna*). It was found that Propranolol caused the greatest effect in the organisms studied. The crustacean *C. dubia* displayed toxicity responses at a concentration of 0.25 μg mL^{-1}. Propranolol was the only API investigated which was found to cause acute toxicity effects in the Japanese medaka fish. These effects were observed at a concentration of 0.5 μg mL^{-1} (Huggett *et al.*, 2002).

Previous studies have suggested that β-blockers would not affect microbes due to the absence of the API receptors within microbial organisms (Jones *et al.*, 2002, CSTEE, 2001). However, in another study conducted by our group we reported that (±)-Propranolol significantly reduced the lipid storage components of the algae *Micrasterias hardyi* 649/15 and caused a marked reduction in the cellular protein content (Patel *et al.*, 2008). In addition, the findings from the metabolic fingerprinting suggested that the phenotype was altered during exposure to the API (Patel *et al.*, 2008). To our knowledge no further studies have been carried out on the metabolic effect of Propranolol within aquatic microorganisms.

The effects on the growth dynamics of the bacteria are likely to reflect changes in the metabolic potential of these cells and this quantitative drug effect was explored further using FT-IR spectroscopy and *P. putida* KT2440 exposed to (±)- Propranolol.

2.4.2 Quantitative effects of APIs on bacteria measured using FT-IR spectroscopy

In order to asses any possible quantitative effects of Propranolol on the phenotype of *P. putida* KT2440 we employed partial least squares (PLS) regression analysis to investigate whether the effect on the phenotype as measured using FT-IR spectroscopy was directly proportional to the concentration of API applied (Figure 2.3). A clear linear relationship was observed between the concentration of (±)- Propranolol to which the *P. putida* KT2440 cells had been exposed to and the metabolic fingerprint. In addition, we were able to predict the concentration of Propranolol to which the bacterial cells had been exposed with an accuracy of 95.45%. This is perhaps not surprising as the inhibitory effect of the Propranolol on the cells was proportional to the concentration of API. This was a clear phenotypic effect as we were unable to collect a spectrum of Propranolol at these concentrations using FT-IR spectroscopy. We also performed PLS regression on the profiles of *P. aeruginosa* PA14 exposed to the intermediate concentrations of Propranolol, to determine if the secondary growth effect was proportional to the concentration of API applied (data not shown). Under these conditions it was not possible to generate a correlation between drug exposure concentration and the FT-IR data. Therefore, the presence of the drug may have led to more complex biochemical perturbations in the organisms that we are unable to model using PLSR.

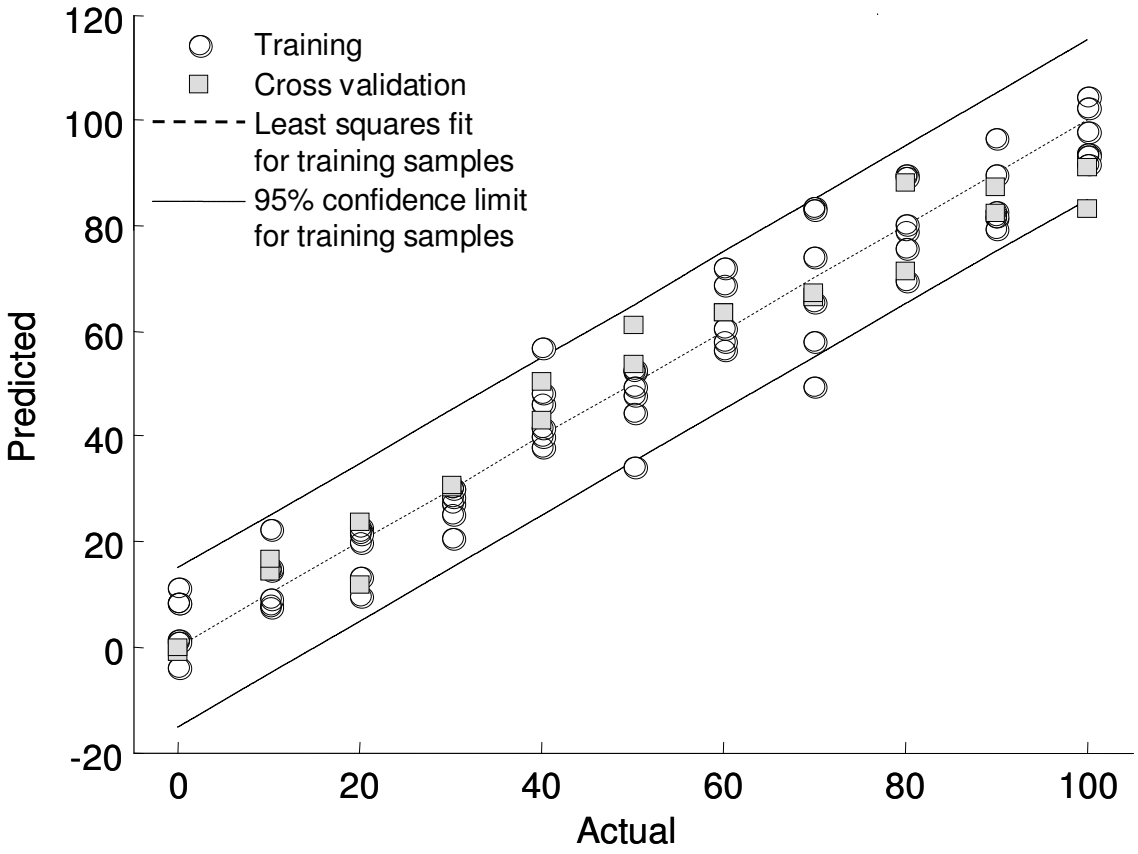


Figure 2.3: Partial least squares regression model for *Pseudomonas putida* KT2440 exposed to varying concentrations (0 – 100 µg mL^{-1} in steps of 10 µg mL^{-1}) of (±)- Propranolol. The model was trained with FT-IR data using two of the biological replicates and validated using the third biological replicate. The PLSR model was built using 10 factors.

2.4.3 Quantitative analysis of API concentration with HPLC

Chiral HPLC analysis was performed to quantify the amount of the enantiomers remaining at the end of the growth period. The analysis was targeted to explore the effects observed in the growth rate data. In addition, control experiments were performed to determine the experimental effect on the drug concentration; i.e., the loss of API during incubation in growth medium. To determine the effect of the enantiomers on the growth of the pseudomonads a range of ratios of Propranolol were employed at 50 µg mL^{-1}. The findings of the HPLC analysis (Table 2.2) demonstrated that neither of the enantiomers were degraded during the batch growth. In addition, the pseudomonads were exposed to a range of concentrations of the API Atenolol (Table 2.3), the concentrations selected were chosen from the findings of the growth rate data. There was no notable indication of API degradation during the growth of these bacteria.

Table 2.3: Quantification of Atenolol from HPLC data for bacterial cells exposed to varying ratios of (*R*)- and (*S*)- Atenolol at a range of concentrations.

Microorganism or Medium	Concentration (µg mL^{-1})	HPLC Quantification				
		Racemate			Pure Enantiomers	
		(*R*)	(*S*)	Total	(*R*)	(*S*)
Medium*	80	38.7 (1.9)	39.1 (1.6)	77.8 (3.5)	79.0 (1.4)	78.2 (2.6)
P. putida KT2440	90	42.1 (1.8)	43.8 (1.2)	85.9 (3.0)	88.9 (2.1)	88.7 (2.6)
	100	48.3 (1.9)	47.8 (2.4)	96.1 (4.3)	99.3 (1.9)	98.8 (2.2)
	110	53.9 (1.8)	54.2 (1.1)	108.1 (2.9)	108.7 (2.6)	109.3 (1.4)
	120	59.4 (1.3)	58.7 (1.8)	118.1 (3.1)	118.9 (2.4)	119.6 (1.2)
	130	62.9 (2.3)	63.8 (1.4)	126.7 (3.7)	129.7 (1.8)	128.9 (1.9)
P. aeruginosa PA14	40	18.6 (1.5)	19.1 (1.2)	37.7 (2.7)	39.7 (1.5)	38.5 (2.1)
	50	23.8 (1.6)	22.9 (2.2)	46.7 (3.8)	47.6 (4.2)	49.3 (1.6)
	60	29.2 (1.3)	28.6 (2.1)	57.8 (3.4)	58.4 (1.9)	59.1 (1.7)
	70	32.9 (2.6)	33.5 (1.4)	66.4 (4.0)	68.9 (1.8)	68.7 (2.2)
	80	39.2 (1.7)	38.8 (2.3)	78.0 (4.0)	79.9 (1.3)	78.6 (1.9)

2.4.4 Effects of chiral APIs on FT-IR metabolic fingerprints

In order to investigate whether there were any *chiral* specific phenotypic changes on the various aquatic bacteria, a single drug concentration in which there was no observable difference in growth rate between the enantiomers was chosen for investigation.

During the investigation into the chiral API specific effects on microorganisms the four bacteria were exposed to a number of drug concentrations (detailed in Table 2.1). A

summary of the statistically significant differences between the drug enantiomers, enantiomeric ratios, or the racemate is shown in Table 2.4.

Table 2.4: The concentration of Propranolol and Atenolol at which a significant effect on the bacterial phenotype was observed using FT-IR spectroscopy.

Microorganism	Concentration of observed effect *	
	Propranolol	Atenolol
P. putida KT2440	50µg mL^{-1}	> 130µg mL^{-1}
P. aeruginosa PA14	50µg mL^{-1}	> 80µg mL^{-1}
M. luteus 2.13	50µg mL^{-1}	ND
B. natatoria 2.1	40 and 50µg mL^{-1}	ND

* Effects are deemed significant in PC-CVA plots based on confidence regions being statistically different at the 95% χ^2 limit. This is to say at the concentration given, below this all bacteria had an equivalent phenotype (overlapping clusters) and above this there was clear differentiation.
ND; not determined as Propranolol appeared to produce the most notable effects during earlier experiments.

PC-CVA was carried out in order to investigate any chiral specific effects on the microorganisms as determined by FT-IR spectroscopy. The distance between samples plotted on a PC-CVA scores plot represents the degree of similarity or dissimilarity between those samples. A smaller distance indicates greater similarity and a larger distance would indicate that there are greater differences between samples. Loading plots provide an indication of which regions of the spectrum are used to define these patterns of separation, which allows for meaningful biochemical interpretation of the results. FT-IR analysis demonstrates that each of the Propranolol enantiomers exerts a metabolic effect on the microbial cells of *P. aeruginosa* PA14 at 80 µg mL^{-1} (Figure 2.4a) compared to the control. It can clearly be observed in the PC-CVA scores plot that the control samples separate across PC-CV 1, which accounts for the greatest variance within the data according to the putative class assignment, and as discussed above this finding was perhaps not surprising given the effect the of the API on the bacterial growth dynamics (Figure 2.2). In addition, the samples exposed to (±)-Propranolol are clearly separate from those exposed to each of the enantiomers ((*R*)- and (*S*)-), which in this analysis show no separation across the first two PC-CV scores. As described above, 2 of the 3 biological replicates were used for calibration (shown in black) and the 3rd projected into the model (grey font). The majority of these projected data are grouped with the appropriate calibration samples indicating that the separation displayed in the model was valid. Moreover, the 95% confidence intervals for the groups are also plotted showing that for 3 of the groups there is a distinct separation in CVA scores space. The analysis demonstrated that the microbial cells exposed to both (*R*)- and (*S*)- Propranolol cluster together in this analysis, indicating that there was no

metabolic difference in the microbial cells exposed to each of the pure enantiomers. It was very surprising to observe that the cultures exposed to the racemate formed a distinct cluster separate from the cultures exposed to the pure enantiomers or control samples in the CVA space. The loadings plot (Figure 2.4b) indicates that very specific changes in the metabolic fingerprint of the microbial cells account for the patterns of separation observed between the control and drug exposed samples in the scores plot. The major chemical changes arise in the protein (1681 – 1629 cm^{-1}) and carbohydrate (1155 – 999 cm^{-1}) regions of the FT-IR spectrum, with a less pronounced contribution from lipid species (2951 – 2845 cm^{-1}) changing in the same direction as the carbohydrates. To further investigate the effect on the biochemical components of the cell between the racemate and enantiomers, we calculated difference spectra (for example the metabolic fingerprint of the (*R*)- enantiomer was divided by that of the racemate) for each combination of interest. This data was then used to determine the relative changes in the lipid and amide components of the cells with respect to the API exposure. Inspection of the difference spectra (Figure 2.5) revealed that the bacterial cells exposed to the racemate contained lower levels of amides and higher levels of lipid compared to those exposed to either of the enantiomers. This suggests that the racemate exerts less of a metabolic effect on the bacterial cells, and this is supported by the PC-CVA scores plot in which the racemate lies between the control and the enantiomer exposed cells across PC-CV 1. As discussed above, HPLC analysis suggests that degradation or a significant uptake of the API does not occur within the microbial cells. Therefore, it is unlikely that the increase in proteins observed in the FT-IR spectra of Propranolol exposed cells is due to the expression of enzymes in order to metabolise the API. It is more probable that this effect is due to the expression of an efflux system to remove the API from the bacterial cell. In addition, Propranolol is a lipophilic API which is known to interact with cell membranes of mammalian cells and the observed reduction of lipids in exposed cells is likely to be due to interactions of the API with the bacterial cell membrane. Propranolol is routinely administered to man as the racemate. The (*S*)- enantiomer accounts for the majority of the β-blocking effect, while the (*R*)- has a predominantly membrane stabilising effect (Barrett and Cullum, 1968, Potter and Sweetland, 1967, Hanna and Evans, 2000, Walle *et al.*, 1984). One can hypothesise that the reason the racemate (±) was different compared to either of the enantiomers is because of the difference in the physical properties between the racemate and enantiomers (Chickos *et al.*, 1981, Secor, 1963).

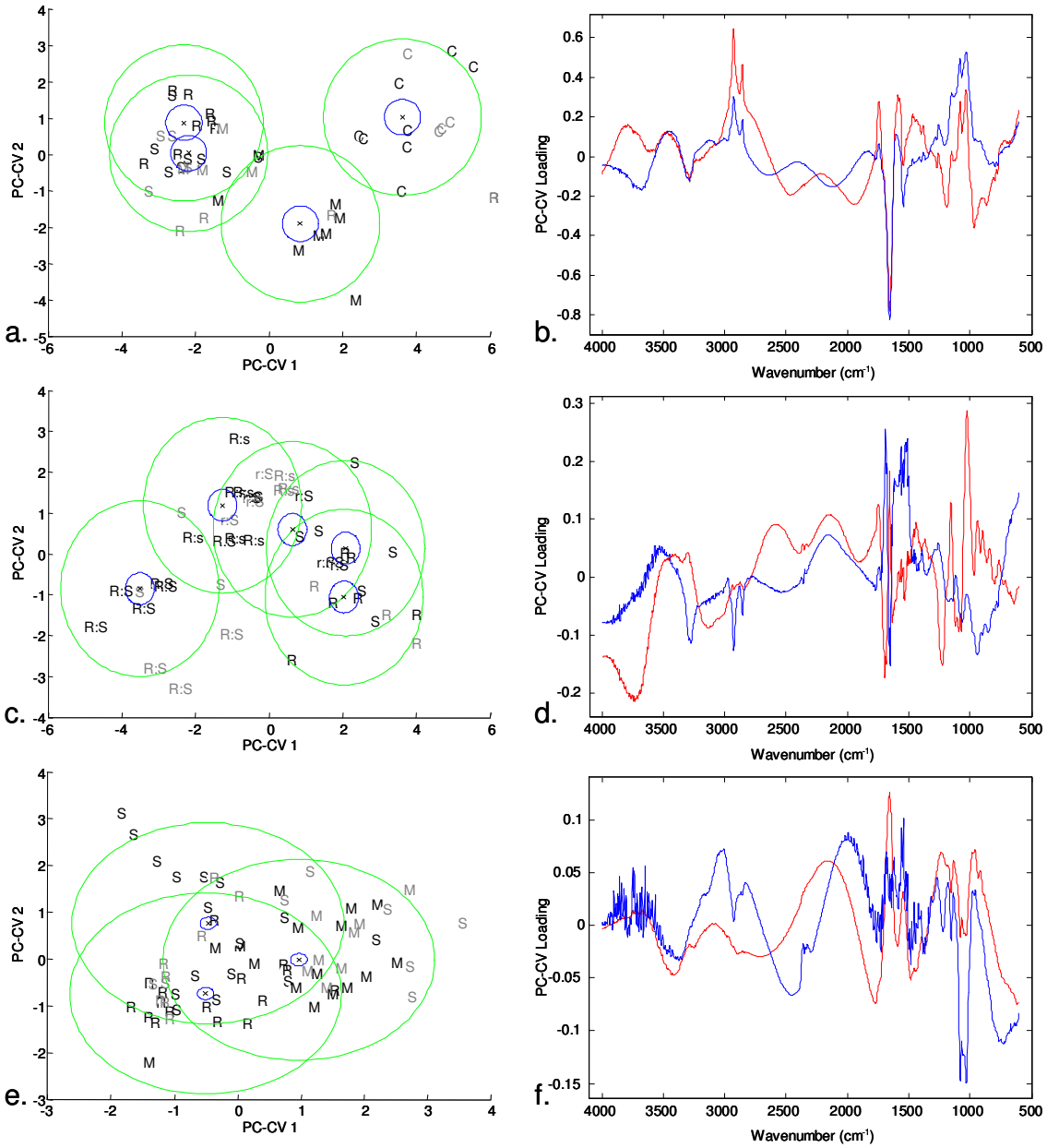


Figure 2.4: PC-CVA scores (LHS) and loadings (RHS) plots performed on FT-IR data for *Pseudomonas aeruginosa* PA14 exposed to 80µg mL^{-1} (*R*)-, (*S*)- and (±)- Propranolol (a and b), *P. putida* KT2440 exposed to varying ratios of 50 µg mL^{-1} (*R*), (*S*)- Propranolol (c and d), and *P. aeruginosa* PA14 exposed to 80µg mL^{-1} (*R*)-, (*S*)- and (±)- Atenolol (e and f). In the scores plots points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Blue circles represent the 95% confidence interval about the group centroid, and green circles the 95% confidence region about the group sample population. In the loadings plots the loadings for PC-CV 1 is represented in red and PC-CV 2 in blue. Points C, M, R and S represent control, racemic mixture, (*R*)- and (*S*)- respectively. The enantiomeric ratios 75:25 and 25:75, (*R*) : (*S*) are represented by R:s and r:S respectively.

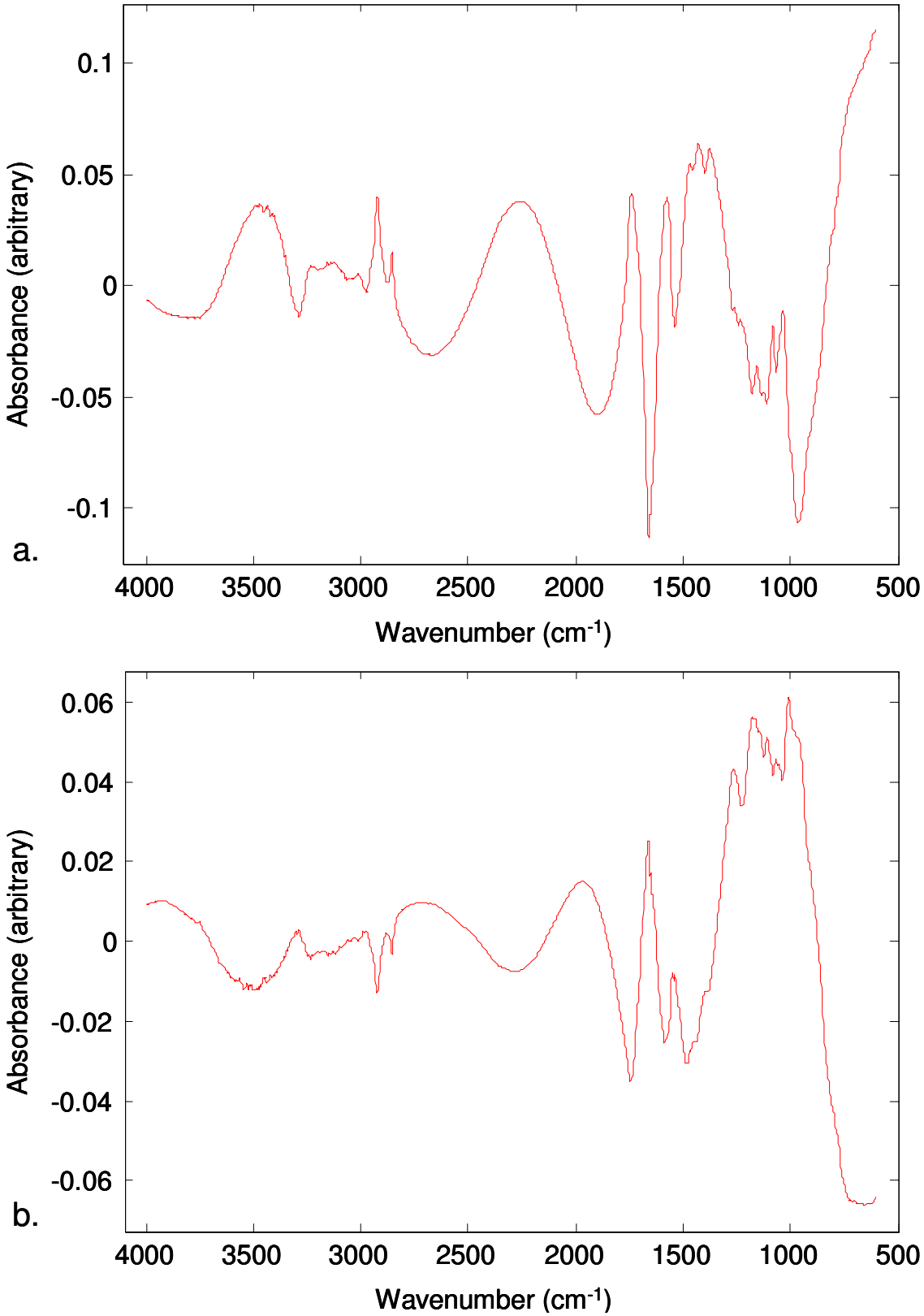


Figure 2.5: FT-IR difference spectra for *Pseudomonas aeruginosa* PA14 exposed to 80 µg mL^{-1} Propranolol. The average FT-IR spectrum for control samples divided by the average spectrum for the pure enantiomers ((*R*)- and (*S*)-) is shown in (a.). The average FT-IR spectrum for the pure enantiomers divided by the average spectrum for the bacterial cells exposed to the racemate is shown in (b.).

To explore the chiral specific effect noted in earlier experiments further the pseudomonads were exposed to varying ratios of 50 µg mL^{-1} (*R*) : (*S*)- Propranolol. The results of the chemometric analysis of the FT-IR spectra again illustrated a metabolic difference in the microbial cells exposed to varying ratios of (*R*)-, and (*S*)- Propranolol from the control samples (data not shown) and these controls were removed prior to PC-CVA so that only chiral specific changes were observed. The results of this PC-CVA are shown in Figure 2.4c where *P. putida* KT2440 exposed to 50 µg mL^{-1} (±)- Propranolol (racemate, labelled R:S) are recovered on the left hand side of this plot and the 25:75 mixture (r:S) as well as the pure enantiomers are located on the right hand side. The 75:25 (R:s) mixture falls between these two groups. The clear separation of *P. putida* KT2440 exposed to the racemate supports the earlier observations on the effect of Propranolol on the metabolic fingerprints of *P. aeruginosa* PA14.

By contrast, there was no phenotypic variation observed in the metabolic fingerprints of the samples exposed to the (*R*)-, (*S*)- and (±)- Atenolol (Figure 2.4e and f) as the 95% χ^2 confidence regions overlap. The analysis clearly illustrates that no discrimination was observed between the three conditions. This result differs from the analysis of *P. aeruginosa* PA14 exposed to Propranolol (Figure 2.4a) and indicates that Atenolol does not have a chiral specific metabolic effect on *P. aeruginosa* PA14.

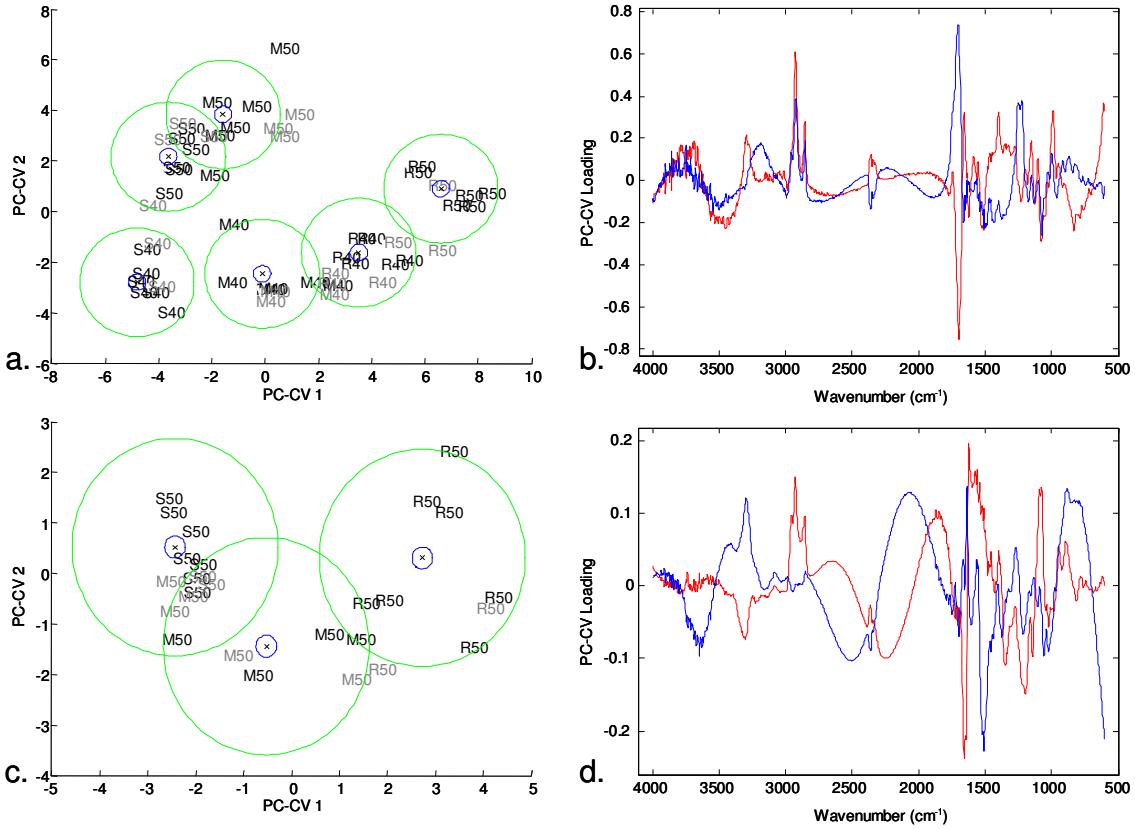


Figure 2.6: PC-CVA scores (LHS) and loadings (RHS) plots performed on FT-IR data for *Blastomonas natatoria* 2.1 exposed to 40 and 50µg mL^{-1} (*R*), (*S*) and (±)- Propranolol (a and b) and *Micrococcus luteus* 2.13 exposed to 50µg mL^{-1} (*R*), (*S*) and (±)- Propranolol. In the scores plots points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Blue circles represent the 95% confidence interval about the group centroid, and green circles the 95% confidence region about the group sample population. In the loadings plots the loadings for PC-CV 1 is represented in red and PC-CV 2 in blue. Points M, R and S represent the racemic mixture (±), (*R*)- and (*S*)- respectively.

Due to the chiral specific effects observed in the pseudomonads when exposed to Propranolol, two additional bacteria were selected to investigate the effects of Propranolol. Propranolol had a very noticeable metabolic effect on *B. natatoria* 2.1 at 40 and 50 µg mL^{-1} (Figure 2.6a and b) and on *M. luteus* 2.13 at 50 µg mL^{-1} (Figure 2.6c and d). The greatest difference observed in these analyses was between the control and API exposed samples. To investigate the more subtle differences between the cultures exposed to the different enantiomers and the racemate the control samples were removed from the analysis. *B. natatoria* 2.1 samples in the PC-CVA scores plot are separated across the 1st CV with respect to the enantiomers. The (*S*)- and (*R*)-enantiomers are clearly separated in the CVA space, with the racemate located between them. A concentration effect was also observed in the metabolic fingerprints across PC-CV 2. This is in contrast to the chiral specific effects for this API on the two pseudomonads, in which the greatest variation was noted between the cells exposed to the racemate and the enantiomers. The loadings data for *B. natatoria* 2.1 indicate the major chemical changes arise in the lipid (2936 – 2851 cm^{-1}) region of the FT-IR spectrum and at wavenumbers 1748 – 1654 cm^{-1}. Vibrations in this region may be attributed to the C=O stretching of esters and carboxylic acids, however this region is dominated by amide I. The FT-IR difference spectra (Figure 2.7) demonstrate that cells exposed to (*S*)- Propranolol contained lower levels of lipids but higher levels of amide and carbohydrate compared to those exposed to the (*R*)- enantiomer. This suggests that (*S*)- Propranolol exerts a greater biological effect on the bacterial cells. The effect of the API on the *M. luteus* 2.13 and *B. natatoria* 2.1 is perhaps more predictable as the differing metabolic effects of the enantiomers is linearly additive. The difference in phenotypic effect following exposure to Propranolol compared to the pseudomonads is likely to be a consequence of the metabolic differences between the bacteria.

To our knowledge, the β-blockers Atenolol and Propranolol have not previously been studied for chiral specific effects within microbial systems. Nevertheless, the effects of APIs within these systems are highly relevant, as microorganisms populate the lower trophic levels within foodwebs. Therefore, differences in the population dynamics could represent a significant effect on the whole of the freshwater community (Jones *et al.*, 2002).

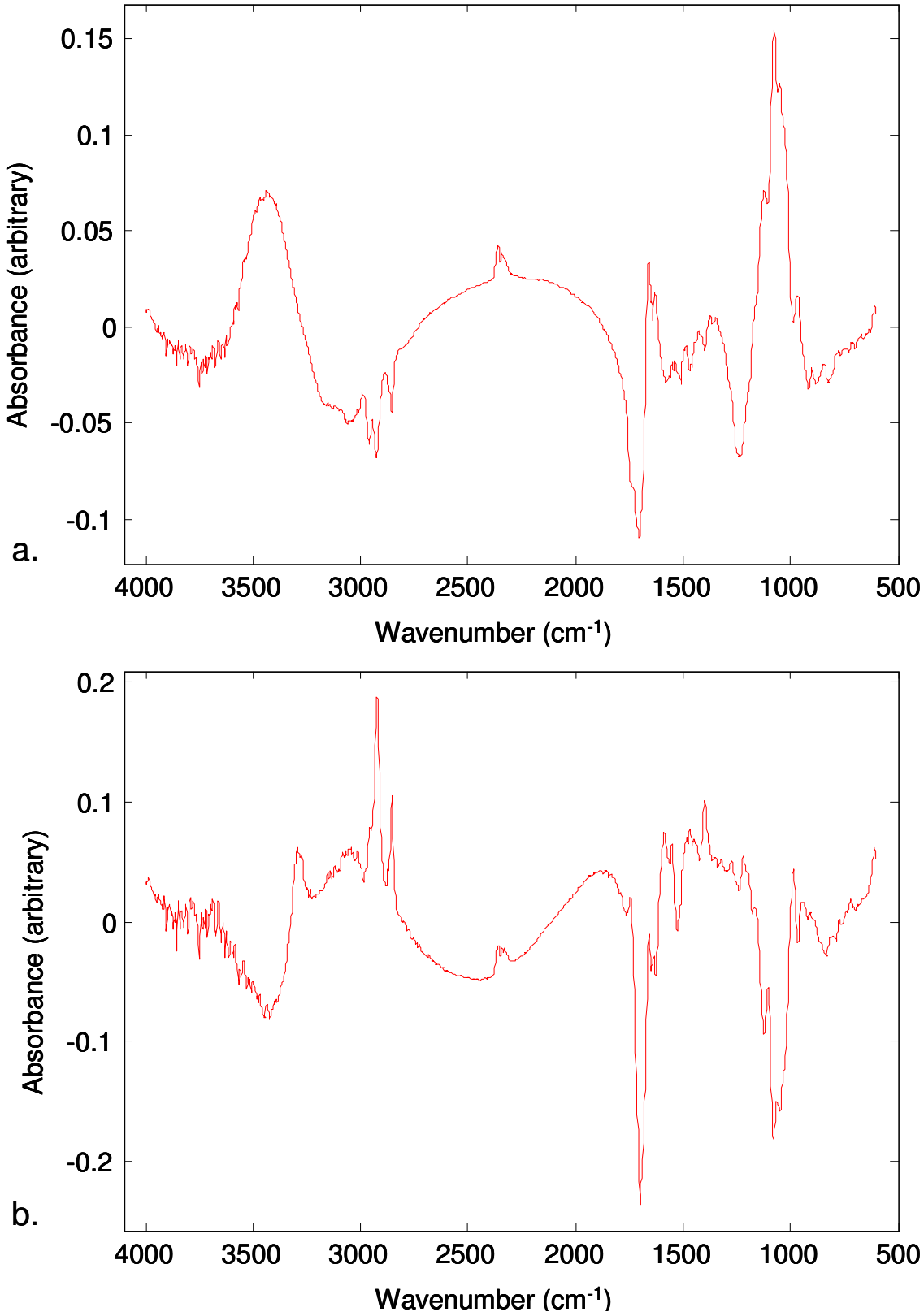


Figure 2.7: FT-IR difference spectra for *Blastomonas natatoria* exposed to 40 and 50 µg mL^{-1} Propranolol. The average FT-IR spectrum for samples exposed to 40 µg mL^{-1} divided by the average spectrum for samples exposed to 50 µg mL^{-1} Propranolol is shown in (a.). The average FT-IR spectrum for bacterial cells exposed to (*R*)- Propranolol divided by the average spectrum for the bacterial cells exposed to the (*S*)- enantiomer is shown in (b.).

2.4.5 Effects of the chiral API Propranolol on selected pseudomonas analysed by GC-MS

In order to investigate the potential chiral specific effects of Propranolol on the two pseudomonads, GC-MS was employed for metabolic profiling of the bacterial metabolomes following exposure to a range of enantiomeric ratios of (*R*) : (*S*)-Propranolol at 50 μg mL^{-1}. The multivariate analysis method PC-DFA was employed in order to visualise the trends within the bacterial metabolic profiles in response to the varying enantiomeric ratios. Figure 2.8 shows the PC-DFA scores plot (PC-DF 1 *vs*. PC-DF 2) constructed from the GC-MS profiles of both pseudomonads following exposure to the varying enantiomeric ratios of Propranolol and the control samples. PC-DF 1 accounts for the greatest variation within the dataset, and this appeared to be due to the biological differences between the two species (*P. aeruginosa* PA14 and *P. putida* KT2440). The samples appeared to display some separation across PC-DF 2 according to the enantiomeric ratio of Propranolol to which the bacteria had been exposed.

In order to explore the differential metabolic effect further, PC-DFA was carried out on the data for each of the pseudomonads separately. A notable chiral specific effect was observed in the metabolic profile for *P. aeruginosa* PA14 (Figure 2.9) and the samples formed distinct clusters according to the enantiomeric ratio of Propranolol to which the cells had been exposed. The samples exposed to the racemate appeared to display the greatest difference in comparison with the control samples according to the GC-MS data, and this finding corresponded with the earlier FT-IR spectroscopy analysis (Figure 2.4). However, it was not possible to differentiate between the pure enantiomers with FT-IR spectroscopy. By contrast, this GC-MS analysis demonstrated a notable difference in the metabolic response to the enantiomers, which showed a clear separation across PC-DF 2 (Figure 2.9). In addition, the remaining enantiomeric ratios (75:25 and 27:75 (*R*) : (*S*)- Propranolol) also appeared to form distinct clusters, suggesting that all of the various enantiomeric ratios exerted differing metabolic effects on the bacterial cultures.

Figure 2.10 shows the PC-DFA scores plot for the GC-MS metabolome data of *P. putida* KT2440 following exposure to the enantiomeric ratios of Propranolol and the control cultures. It can be clearly observed that the greatest variation within the dataset was between the control and Propranolol exposed cells. The control samples were

therefore removed from the analysis in order to investigate more subtle metabolic responses to the enantiomeric ratios within the bacterial cells (Figure 2.11). As observed for *P. aeruginosa* PA14, the bacterial cells appeared to display differing metabolic responses according to the enantiomeric ratios. Cultures exposed to the pure enantiomers displayed distinct clusters from the remaining samples. Previous analyses of *P. putida* KT2440 exposed to Propranolol were not able to differentiate fully between the experimental conditions based on the FT-IR spectra (Figure 2.4c). However, distinct clusters were observed for each of the enantiomeric ratios in the PC-DFA of the metabolic profile data (with the exception of one possible outlier labelled r:S which appeared to co-cluster with the racemate exposed cultures, labelled R:S). These findings indicated that differential effects were exerted on *P. putida* KT2440 when exposed to the enantiomeric ratios and corresponded with the findings for the closely related bacterial species, *P. aeruginosa* PA14.

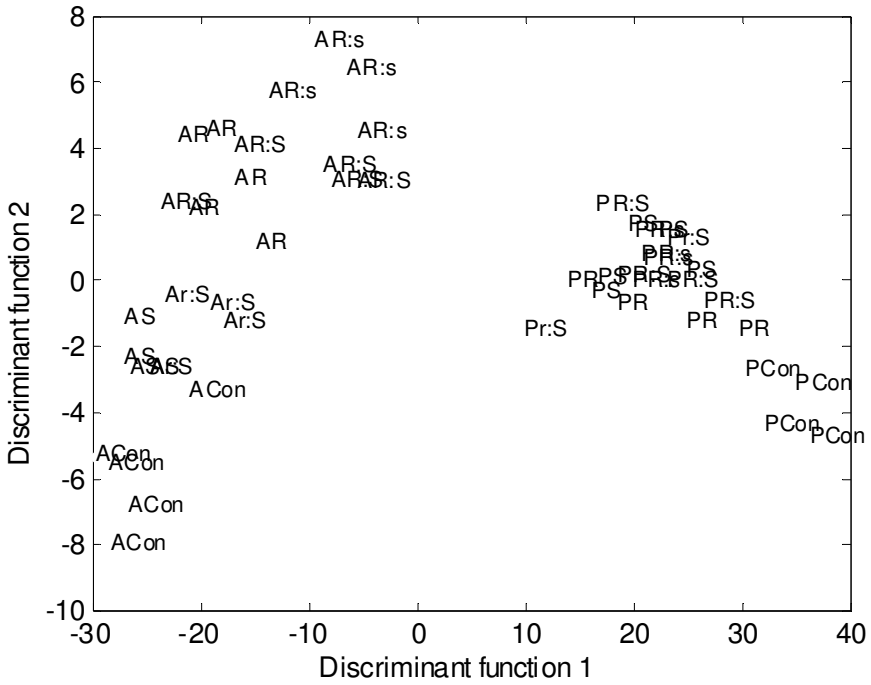


Figure 2.8: PC-DFA scores plots performed on GC-MS metabolome profiles from *Pseudomonas aeruginosa* PA14 and *P. putida* KT2440 exposed to varying enantiomeric ratios of 50 µg mL^{-1} (*R*) : (*S*)- Propranolol. The model was constructed using 20 PCs. Points A and P represent *P. aeruginosa* PA14 and *P. putida* KT2440 samples respectively. The enantiomeric ratios 100:0, 75:25, 50:50 25:75 and 0:100, (*R*) : (*S*)- Propranolol are represented by R, R:s, R:S, r:S and S respectively. Control samples are represented by Con.

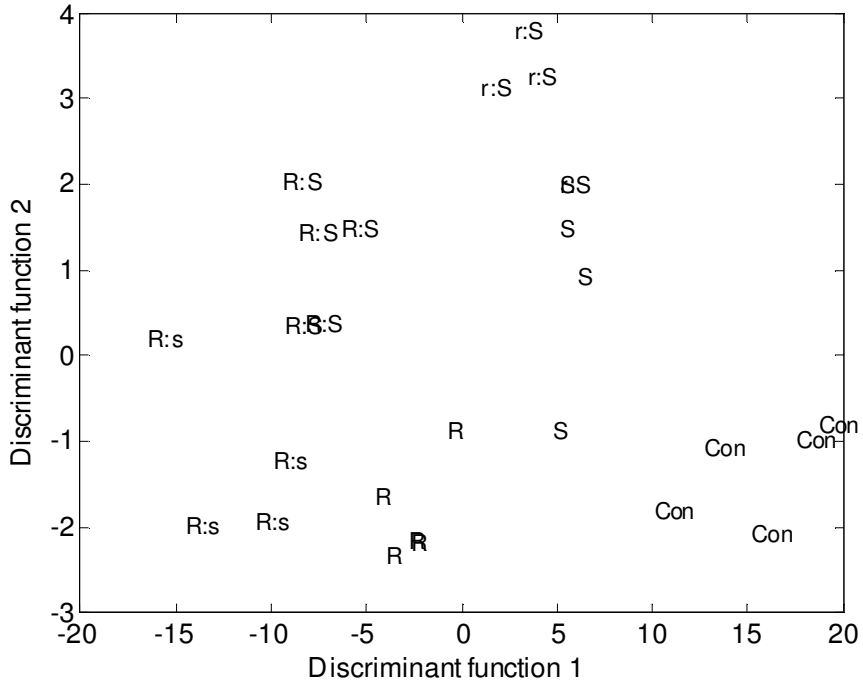


Figure 2.9: PC-DFA scores plots performed on the GC-MS metabolome profiles from *Pseudomonas aeruginosa* PA14 exposed to varying ratios of 50 µg mL^{-1} (*R*) : (*S*)- Propranolol. The model was constructed using 15 PCs. The enantiomeric ratios 100:0, 75:25, 50:50 25:75 and 0:100, (*R*) : (*S*)- Propranolol are represented by R, R:s, R:S, r:S and S respectively. Control samples are represented by Con.

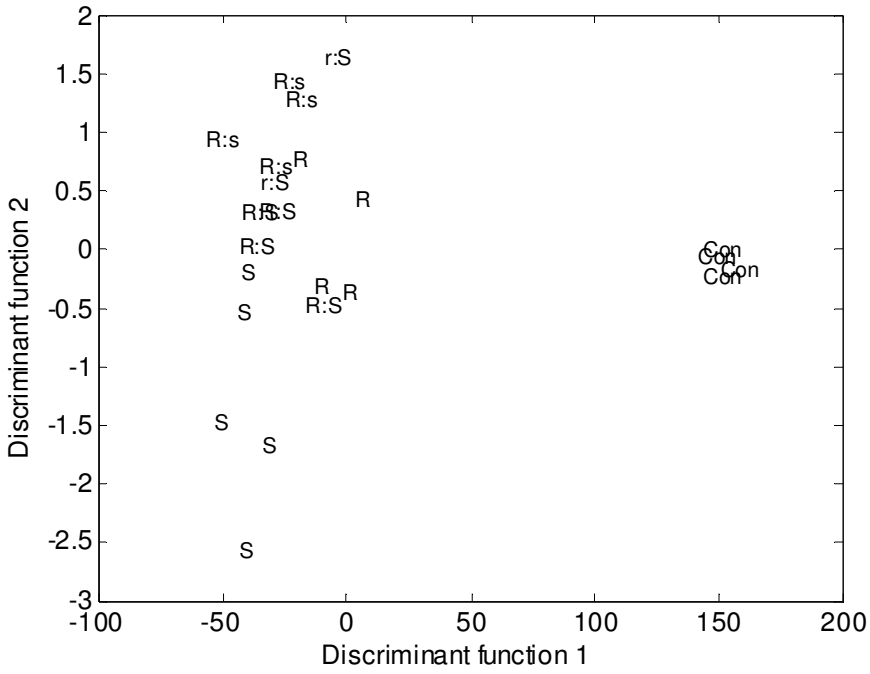


Figure 2.10: PC-DFA scores plots performed on the GC-MS metabolome profiles from *Pseudomonas putida* KT2440 exposed to varying ratios of 50 µg mL^{-1} (*R*) : (*S*)- Propranolol. The model was constructed using 10 PCs. The enantiomeric ratios 100:0, 75:25, 50:50 25:75 25:75 and 0:100, (*R*) : (*S*)- Propranolol are represented by R, R:s, R:S, r:S and S respectively. Control samples are represented by Con.

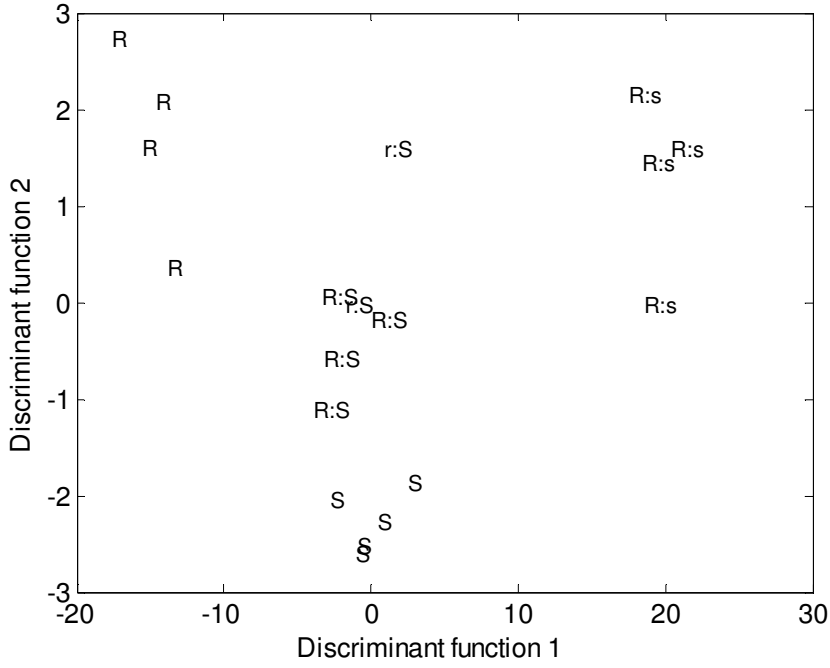


Figure 2.11: PC-DFA scores plots performed on the GC-MS metabolome profiles from *Pseudomonas putida* KT2440 exposed to varying ratios of 50 µg mL^{-1} (*R*) : (*S*)- Propranolol. The model was constructed using 13 PCs. The enantiomeric ratios 100:0, 75:25, 50:50 25:75 25:75 and 0:100, (*R*) : (*S*)- Propranolol are represented by R, R:s, R:S, r:S and S respectively. The control samples were removed for this analysis.

The Lilliefors test for normality was carried out on the GC-MS profiling data in order to assess whether the data followed a normal distribution (Lilliefors, 1967). The profiling data was not found to be normally distributed and therefore appropriate non-parametric tests were employed for subsequent analysis of the data. The Friedman test is a non-parametric two-way analysis of variance which does not make assumptions regarding the distribution of the data (Theodorsson-Norheim, 1987). A number of pair wise comparisons ($p < 0.02$) were carried out on the GC-MS profiling data in order to determine the significant changes in metabolites within the pseudomonads following exposure to the API in comparison with the control samples. Table 2.5 and Table 2.7 summarise the metabolites which displayed a significant difference compared to the control cultures in *P. aeruginosa* PA14 and *P. putida* KT2440 respectively. A total of 34 metabolites were found to change significantly in *P. aeruginosa* PA14 following exposure to the pure enantiomers or the racemate in comparison with the control cultures. Unfortunately, identification was only possible for 12 of these metabolites by means of matching to authentic standards by retention time and mass spectrum or by mass spectrum only (Winder *et al.*, 2008). Comparison of the *P. putida* KT2440 metabolic profiles following exposure with the control cultures revealed significant changes in 29 of the observed metabolites, of which 16 were identified.

The univariate data analysis of the *P. putida* KT2440 metabolic profile revealed significant reductions in the levels of fatty acids such as myristic acid, hexadecanoic acid and octadecanoic acid in cultures exposed to (*S*)- Propranolol. By contrast, the pure (*S*)- enantiomer appeared to exert a different metabolic effect in *P. aeruginosa* PA14, and levels of octadecanoic acid were significantly higher in cultures exposed to the enantiomer compared to the control cultures. These findings supported the earlier hypotheses that Propranolol exerts a metabolic effect on the lipid components within the pseudomonads and in addition, the API appeared to exert differing metabolic effects in the two species. Furthermore, levels of the phospholipid head group ethanolamine, commonly found in bacterial cell membranes, was significantly increased in *P. aeruginosa* PA14 cultures exposed to (±)- Propranolol. By contrast, *P. putida* cultures exposed to (*S*)- and (±)- Propranolol displayed lower levels of glycerol-3-phosphate; a component of glycerophospholipids and cultures exposed to the (*S*)-enantiomer displayed lower levels of N-acetylglucosamine, which is a component of the peptidoglycan bacterial cell wall. These results suggest that (*S*)- and (±)- Propranolol exerted an enhanced metabolic effect on the bacterial cell membrane and cell wall in *P. putida* KT2440 in comparison to the (*R*)- enantiomer. In addition, these findings

indicate that the API exerts a greater effect on the cell membranes in *P. putida* KT2440 than in the closely related *P. aeruginosa* PA14.

A number of amino acids including valine, leucine, isoleucine and tyrosine were significantly reduced in *P. putida* KT2440 following exposure to (*S*)- Propranolol. In addition, glutamic acid / pyroglutamic acid was reduced in (*R*)- exposed cultures. A reduction in valine was also observed in the *P. aeruginosa* PA14 cultures exposed to (*R*)- Propranolol and the racemate and an increase in tryptophan was observed following exposure to the racemate. However, no other significant changes in amino acids were observed in the metabolic profiles of *P. aeruginosa* PA14. A number of intermediates in the urea cycle (citrulline, ornithine and xanthine) were found to increase significantly in *P. aeruginosa* PA14 following exposure to (±)- Propranolol. However, no notable difference was observed for these metabolites in the metabolic profile of *P. putida* KT2440 providing further evidence of the different metabolic effects of Propranolol in the two pseudomonad species.

Kruskal-Wallis univariate data analysis (Kruskal and Wallis, 1952) was employed in order to determine any significant changes in the metabolite profiles of the pseudomonads following exposure to the pure enantiomers and the racemate. The Venn diagrams in Figure 2.12 and Figure 2.13 provide a summary of the numbers of metabolites displaying a significant change between the experimental conditions for *P. aeruginosa* PA14 and *P. putida* KT2440 respectively ($p < 0.02$), and Table 2.6 and Table 2.8 provide identifications for these metabolites. The univariate analyses indicated that only one metabolite displayed a significant difference in the metabolic profiles of *P. aeruginosa* PA14 exposed to (*R*)- and (*S*)- Propranolol. However, it was not possible to identify this metabolite. By contrast, 5 metabolites displayed significant differences in the *P. putida* KT2440 cultures exposed to the pure enantiomers, of which the amino acids tyrosine and tryptophan and the disaccharide trehalose were identified. Further inspection of the data revealed significantly lower levels of all of the metabolites in the (*S*)- exposed cultures compared to the (*R*)- enantiomer, suggesting that (*S*)- Propranolol exerts an enhanced metabolic effect on *P. putida* KT2440. These findings suggest that (*S*)- Propranolol exerts an enhanced metabolic effect on *P. putida* KT2440 and this is not perhaps surprising as the (*S*)- enantiomer exerts the majority of the intended β-blocking activity in man (Barrett and Cullum, 1968, Hanna and Evans, 2000, Potter and Sweetland, 1967, Walle *et al.*, 1984).

Comparison of the cultures exposed to (*R*)- or (±)- Propranolol revealed significant changes in a number of metabolites in the metabolic profile of *P. aeruginosa* PA14 (Figure 2.12). The racemate appeared to have a greater metabolic effect on the glycolysis intermediate, Fructose-6-phosphate, whilst the (*R*)- enantiomer appeared to exert the greatest effect on gluconolactone. In contrast, only one metabolite was identified as significantly different between the (*R*)- and (±)- Propranolol exposed cultures of *P. putida* KT2440 in the metabolic profile data (Figure 2.13). Unfortunately it was not possible to identify this metabolite. In addition, comparison of the metabolic profiles of the pseudomonads exposed to (*S*)- Propranolol or the racemate revealed reductions in the levels of glucose-6-phosphate in the *P. putida* KT2440 metabolic profile of the (*S*)- exposed cultures. By contrast, *P. aeruginosa* PA14 cultures displayed lower levels of lactose following exposure to the racemate. An additional two metabolites were identified as significantly different for this comparison and inspection of the data revealed significant reductions in the levels of these metabolites following exposure to the racemate in comparison to the (*S*)- enantiomer, however it was not possible to identify them from the retention time and mass spectra. These results suggest that (±)- Propranolol exerts a greater effect on the *P. aeruginosa* PA14 metabolic profile than (*S*)- Propranolol, whereas the pure (*S*)- enantiomer appeared to exhibit a greater metabolic effect in *P. putida* KT2440.

Table 2.5: Metabolites found to be significantly altered in *Pseudomonas aeruginosa* PA14 following exposure to (*R*)-, (*S*)- or (±)- Propranolol. Metabolites are listed in order of significance (highest first) according to the Friedman test. *

Metabolite	R	±	S
Valine[d]			
Amine[d]			
Tryptophan[d]			
2-hydroxyglutarate[d]			
Octadecanoic acid[d]			
Hexadecanoic acid[d]			
Ethanolamine[d]			
Citrulline[d]			
Nicotinamide[d]			
Ornithine[d]			
Xanthine[p]			
Myristic acid[d]			
Unknown peak (24)			
Unknown peak (163)			
Unknown peak (69)			
Unknown peak (93)			
Unknown peak (134)			
Unknown peak (190)			
Unknown peak (181)			
Unknown peak (8)			
Unknown peak (91)			
Unknown peak (195)			
Unknown peak (182)			
Unknown peak (108)			
Unknown peak (172)			
Unknown peak (5)			
Unknown peak (159)			
Unknown peak (206)			
Unknown peak (29)			
Unknown peak (135)			
Unknown peak (137)			
Unknown peak (62)			
Unknown peak (129)			
Unknown peak (143)			

* Green points represent a significant increase and red points represent a significant decrease in the metabolites identified in the Friedman test ($p<0.02$).
[d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only. For the unknown peaks the database identifier is provided in parentheses.

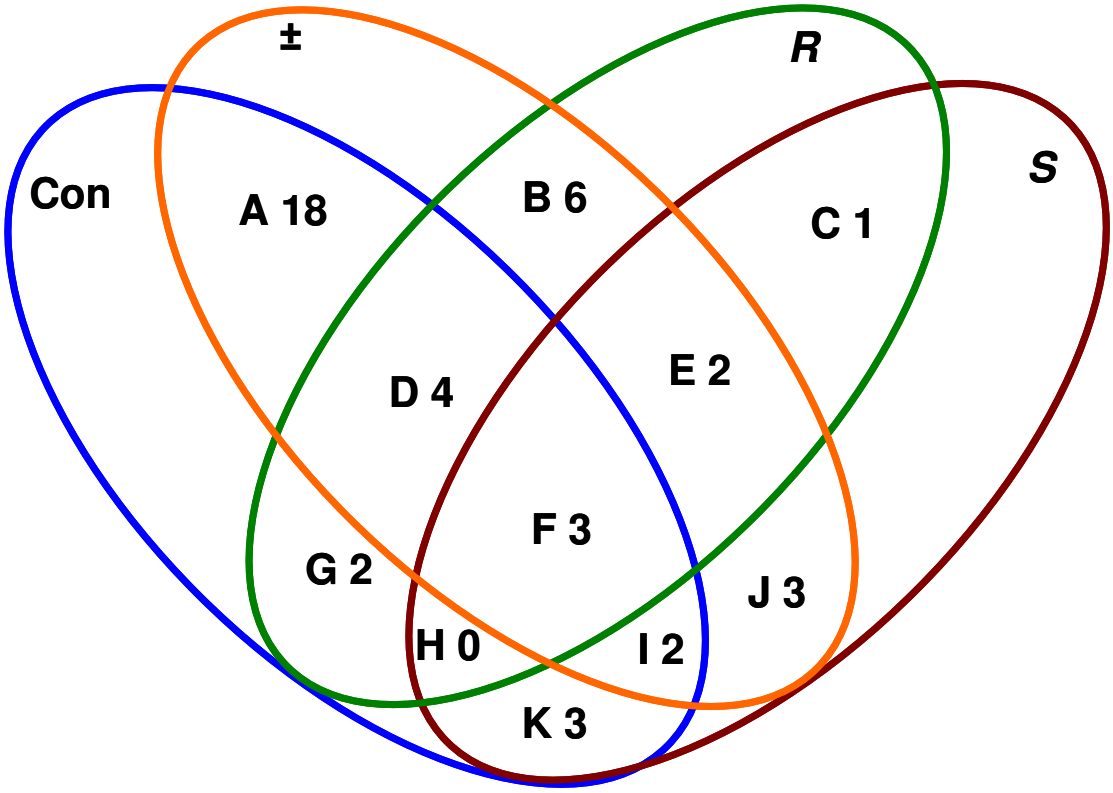


Figure 2.12: Venn diagram illustrating the number of metabolites identified as significantly different in *Pseudomonas aeruginosa* PA14 following exposure to (*R*)- , (*S*)- and (±)- Propranolol exposed cultures. Identification of the metabolites from each section is detailed in Table 2.6.

Table 2.6: Identification of the metabolites displaying significant differences in *Pseudomonas aeruginosa* PA14 following exposure to (*R*)- , (*S*)- and (±)- Propranolol, illustrated in the Venn diagram. Columns A-K refer to the different sections of the Venn diagram.

A	**C**	**I**
Tryptophan[d]	Unknown peak (1)	Amine[p]
2-hydroxyglutarate[d]	**D**	Unknown peak (1)
Ethanolamine[p]	Valine[d]	**J**
Citrulline[d]	Unknown peaks (3)	Lactose[d]
Nicotinamide[d]	**E**	Unknown peaks (2)
Ornithine[d]	Glycerate-3-phosphate / glycerol-3-phosphate[p]	**K**
Xanthine[p]	Amine[p]	Octadecanoic acid[d]
Unknown peaks (11)	**F**	Unknown peaks (2)
B	Unknown peaks (3)	
Gluconolactone[p]	**G**	
Fructose-6-phosphate[d]	Unknown peaks (2)	
Unknown peaks (4)		

[d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only.

Table 2.7: Metabolites found to be significantly altered in *Pseudomonas putida* KT2440 following exposure to (*R*)-, (*S*)- or (±)- Propranolol. Metabolites are listed in order of significance (highest first) according to the Friedman test. *

Metabolite	R	±	S
Glycerol-3-phosphate[d]			
Citric acid[d]			
Myristic acid[d]			
5-aminovaleric acid[d]			
N-acetylglucosamine[p]			
Leucine[d]			
Isoleucine[d]			
Tyrosine[d]			
Hexadecanoic acid[d]			
Citrulline[d]			
Lysine[d]			
Amine[p]			
Glutamic / pyroglutamic acid[d]			
Octadecanoic acid[d]			
Gluconolactone[p]			
Valine[d]			
Unknown peak (158)			
Unknown peak (181)			
Unknown peak (163)			
Unknown peak (182)			
Unknown peak (94)			
Unknown peak (91)			
Unknown peak (48)			
Unknown peak (28)			
Unknown peak (76)			
Unknown peak (116)			
Unknown peak (183)			
Unknown peak (200)			
Unknown peak (111)			

* Green points represent a significant increase and red points represent a significant decrease in the metabolites identified in the Friedman test ($p<0.02$).
[d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only. For the unknown peaks the database identifier is provided in parentheses.

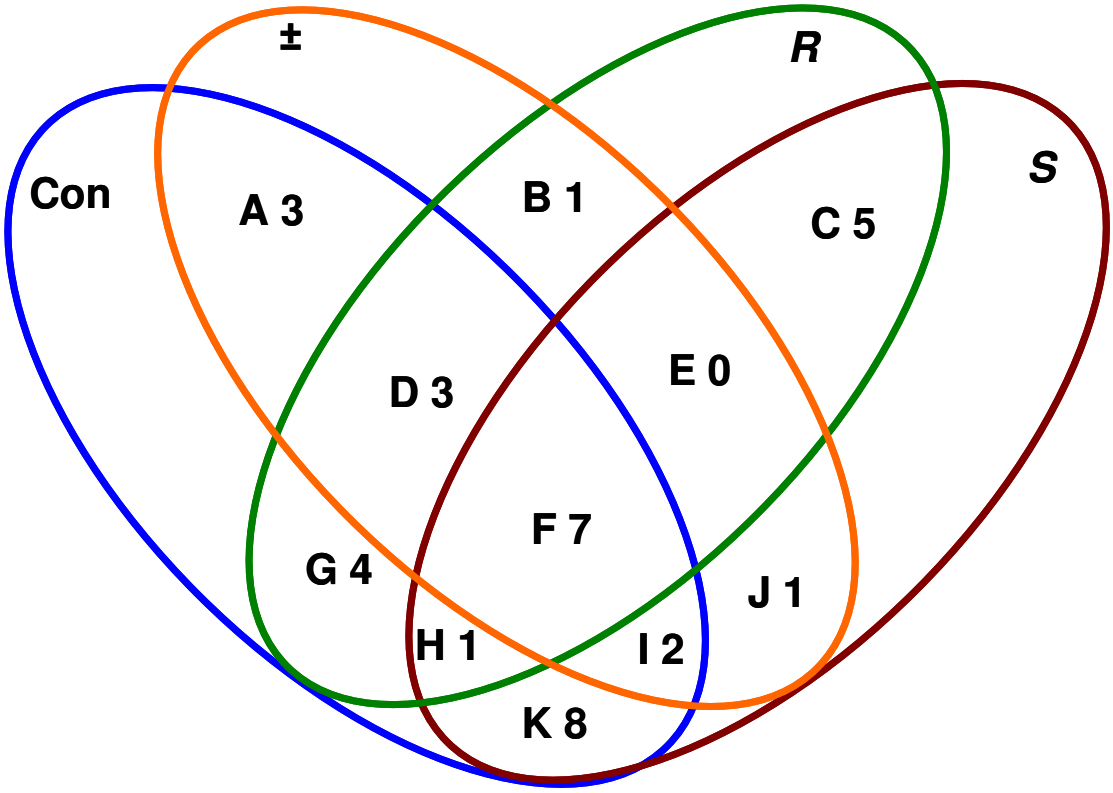


Figure 2.13: Venn diagram illustrating the number of metabolites identified as significantly different in *Pseudomonas putida* KT2440 following exposure to (*R*)- , (*S*)- and (±)- Propranolol. Identification of the metabolites from each section is detailed in Table 2.8.

Table 2.8: Identification of the metabolites displaying significant differences in *Pseudomonas putida* KT2440 following exposure to (*R*)- , (*S*)- and (±)- Propranolol, illustrated in the Venn diagram. Columns A-K refer to the different sections of the Venn diagram.

A	**F**	**I**
Unknown peaks (3)	Valine[d]	N-acetylglucosamine[p]
B	Glycerol-3-phosphate[d]	Octadecanoic acid[d]
Unknown peak (1)	Lysine[d]	**J**
C	Amine[p]	Glucose-6-phosphate[d]
Tyrosine[d]	Unknown peaks (3)	**K**
Tryptophan[d]	**G**	Leucine[d]
Trehalose[d]	Glutamic / pyroglutamic acid[d]	Isoleucine[d]
S-adenosylmethionine [p]	Gluconolactone[p]	Glycerol-3-phosphate[d]
Unknown peak (1)	Unknown peaks (2)	Citric acid[d]
D	**H**	Myristic acid[d]
5-aminovaleric acid[d]	Unknown peak (1)	Tyrosine[d]
Unknown peaks (2)		Hexadecanoic acid[d]
		Unknown peak (1)

[d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only.

2.5 Conclusions

It was clearly observed from the growth data that Propranolol had a biological effect on all the microorganisms studied. At the higher concentrations tested the growth was retarded and in most cases the death rate was increased, with associated changes observed in the metabolic fingerprints. The loadings plots from the PC-CVA of API dosed and unexposed *P. aeruginosa* PA14 cells (Figure 2.4b) indicated that Propranolol exerts a widespread effect on the bacterial cells and this finding was also observed in the other bacteria studied. The results from the HPLC analysis showed that the API was not degraded during the growth period and this suggests that the observed changes in the multivariate analysis of the metabolic fingerprints were not due to degradation of the API but more likely to be a secondary effect of the drug. Despite the genetic similarity of the two pseudomonads studied, our findings show that Propranolol exerted a different effect in the two species. In contrast, the growth data showed that an effect was not observed in the Atenolol exposed cultures and this finding was reflected in the multivariate analyses of the bacterial fingerprints (Figure 2.4e).

All four aquatic bacteria were exposed to the enantiomers of Propranolol at concentrations in which there was no difference observed in the growth rates. The FT-IR analysis revealed that Propranolol affected both the lipid and protein content of the bacterial cells. We hypothesise that this is likely due to the interaction of the APIs with the microbial cell walls. A more predictable effect on the metabolic fingerprints was noted during the exposure of *B. natatoria* 2.1 and *M. luteus* 2.13 to the Propranolol in which the racemate fell between the (*R*)- and (*S*)- enantiomers in the PC-CVA. Rather surprisingly the most significant effect on the two pseudomonads was with the racemate, whilst the enantiomers had identical effects on the phenotype of the cells. It is possible that the physical properties of the racemate were significantly different from the (*R*)- and (*S*)- enantiomers and this was reflected in how the cells responded to exposure to the API.

Multivariate analysis of the metabolite profile data from GC-MS of the two pseudomonads following exposure to a range of enantiomeric ratios of (*R*) : (*S*)-Propranolol revealed notable chiral specific effects in both *P. aeruginosa* PA14 and *P. putida* KT2440. The analysis revealed that it was possible to differentiate between all of the various enantiomeric ratios and indicated that differential metabolic effects were

exerted on the bacterial cells according to the enantiomeric ratios to which they had been exposed. In addition, significant reductions in a number of fatty acids were observed in *P. putida* KT2440 following exposure to the pure (*S*)- enantiomer and components from the bacterial cell membrane and cells wall were significantly reduced. The findings indicated that (*S*)- and (±)- Propranolol exerted an enhanced metabolic effect on *P. putida* KT2440 in comparison with the (*R*)- enantiomer. Furthermore, significant changes were observed in a number of intermediates of the urea cycle in *P. aeruginosa* PA14 cultures exposed to (±)- Propranolol. No significant changes were observed for these metabolites in *P. putida* KT2440, providing further evidence of a different metabolic effect in the two closely related pseudomonads.

In conclusion it has been shown that chiral specific effects do occur in bacteria which may have implications for environmental ecosystems as these APIs are regularly found in the aquatic environment. We believe FT-IR spectroscopy with appropriate chemometrics to be a very powerful method for investigating the phenotype and metabolic differences that APIs have when they interact with the bacterial cell.

3 Monitoring the phenotypic and spatial effect of chiral pharmaceuticals on green algae using FT-IR microspectroscopy

3.1 Abstract

Pharmaceuticals are an emerging class of environmental contaminants. Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment and a number of APIs have been detected in wastewaters due to excretion of the API following medicinal and veterinary use. Despite the low concentrations detected (ng - μg L^{-1}), the presence of APIs in the aquatic environment is of great concern due to the potential for causing undesirable ecological effects. In addition, the enantiomers of chiral APIs may possess differing toxicological and biological effects both from each other, and the racemate (an equal mixture of the two enantiomers). It is therefore necessary to enhance our understanding of the fate and biological effects of chiral pharmaceuticals in the aquatic environment. Of particular interest is the group of chiral APIs termed β-blockers as they all contain at least one chiral centre and are generally administered as the racemate. In this investigation FT-IR microspectroscopy was employed to investigate the chirality specific effects of the enantiomeric ratios of the β-blockers Atenolol and Propranolol on the total cellular fingerprints of the green algae *Micrasterias hardyi*. In addition, the spatially localised effects of the APIs and their enantiomer ratios on the biochemical components of the algal cells were also investigated. No notable difference was observed in *M. hardyi* exposed to Atenolol compared to the control samples. By contrast, a phenotypic effect on the algal cells was observed when exposed to Propranolol. A clear difference was observed in the FT-IR spectra of the algal cells exposed to 30 μg mL^{-1} Propranolol at all of the enantiomeric ratios and a dominant peak was observed in the FT-IR spectra at wavenumbers 1500 - 1200cm^{-1} which was not observed in cells exposed to (*R*)-Propranolol or the enantiomer. Inspection of the photomicrographs and functional group maps revealed that the algal cells exposed to (*R*)- and (±)- Propranolol were particularly affected by the API. The greatest reduction in the amide region of the FT-IR spectra was observed in cells exposed to the racemate. However, it was not possible to determine any specific localised effects in the algal cells relating to the cellular components. Analysis of both the lipid rich and amide I rich areas of the algal cells did not reveal significantly different responses from those observed throughout the cell. It is therefore likely that Propranolol exerts a general effect on the whole algal cell. This investigation illustrates the power of spatial metabolic fingerprinting with FT-IR microspectroscopy for the study of abiotic stresses on complex biological species.

3.2 Introduction

Pharmaceuticals are an emerging class of environmental contaminants (Fent *et al.*, 2006). Active pharmaceutical ingredients (APIs) and their metabolites are ubiquitous in the environment and a number of APIs have been detected in wastewaters due to excretion of the API following medicinal and veterinary use (Escher *et al.*, 2005). In addition, the incomplete removal of these compounds during the wastewater treatment processes has led to the presence of many APIs such as antibiotics, estrogens, non-steroidal anti-inflammatories, and β-blockers in aquatic environments such as rivers streams and estuaries (Fent *et al.*, 2006, Escher *et al.*, 2005, Ashton *et al.*, 2004, Carlsson *et al.*, 2006b). Despite the low concentrations detected (ng - μg L^{-1}), the presence of APIs in the aquatic environment is of great concern due to the potential for causing undesirable ecological effects (Huggett *et al.*, 2002, Kümmerer, 2009, Kümmerer *et al.*, 2000).

Although little is known about the effects of APIs in the environment, it must be taken into account that they are designed to elicit a specific mode of action in humans (Escher *et al.*, 2005) and it can be expected that any beneficial or adverse effect may also be observed in aquatic organisms with similar biological functions or receptors. It must also be noted that similar targets may control different metabolic processes in different species (Seiler, 2002), and therefore APIs and their metabolites may act through additional modes of action in aquatic organisms. The effects of APIs may be rather subtle due to the very low concentrations observed in the aquatic environment, and as a result these effects may go unnoticed (Escher *et al.*, 2005). In addition, it is likely that the effect of the API will impact on the local population dynamics throughout the whole ecosystem, from bacteria up to higher organisms.

The chirality of environmental contaminants such as APIs must be taken into consideration in order to understand fully the environmental fate and effects of these compounds. The enantiomers of a chiral API are able to interact differently with other chiral compounds such as enzymes, and therefore potentially exert different effects when released into the environment (Fono and Sedlak, 2005, Buser *et al.*, 1999, Nikolai *et al.*, 2006). In addition, the enantiomers of a chiral API may possess differing toxicological and biological effects both from each other, and the racemate (an equal mixture of the two enantiomers) (Lees *et al.*, 2003, Yang *et al.*, 2005). It is therefore necessary to enhance our understanding of the fate and biological effects of chiral

pharmaceuticals in the aquatic environment in order to appreciate the risks fully (Huggett *et al.*, 2002). Of particular interest is the group of APIs termed β-blockers as they all contain at least one chiral centre and are generally administered as the racemate (Mehvar and Brocks, 2001). In addition, they are widely used; with approximately 29 and 12 tonnes of Atenolol and Propranolol respectively taken each year in the UK (Ashton *et al.*, 2004, Carlsson *et al.*, 2006a, Carlsson *et al.*, 2006b).

It is difficult to determine the total effect of APIs in the environment as little information on the ecotoxicological effects of pharmaceuticals on both aquatic and terrestrial organisms is currently available. The effect of APIs on aquatic organisms is considered particularly important as they are exposed to these compounds via wastewater residues during their entire lifespan (Fent *et al.*, 2006). Microalgae account for approximately 50% of total planetary primary productivity and are the major primary producers in most aquatic ecosystems (Beardall *et al.*, 2001, Heraud *et al.*, 2005, Shelly *et al.*, 2002). These photosynthetic, unicellular organisms are known to produce hydrogen and lipids under stress conditions and have also been used as biomarkers of contamination in aquatic environments (Torres *et al.*, 2008). In addition, it is known that algae play an important role in the dispersal, chemical transformation and bioaccumulation of many APIs (Bopp and Lettieri, 2007, Kowalewska, 1999, Lei *et al.*, 2002, Murray *et al.*, 2003, Okay *et al.*, 2000, Todd *et al.*, 2002, Wang *et al.*, 1998). Microalgal cells have been reported to display reorganisation of the composition of macromolecular components in response to alterations in available nutrients within their environment. This effect has been accurately determined using FT-IR microspectroscopy (Beardall *et al.*, 2001, Giordano *et al.*, 2001, Heraud *et al.*, 2005).

As discussed above (Section 1.4.1, page 33) FT-IR spectroscopy is a rapid, reagentless and non-destructive technique, facilitating the identification of functional chemical groups and polar bonds within a biological sample through the generation of a spectral “fingerprint” (Stuart, 1996, Goodacre *et al.*, 1998, Winder *et al.*, 2004, Naumann *et al.*, 1994). FT-IR spectroscopy has previously been used to generate informative metabolic fingerprints from biological material (Johnson *et al.*, 2004, Winder *et al.*, 2007). Previous studies have also proved its applicability to biological systems by successfully demonstrating the discrimination of bacteria to sub-species level (Winder *et al.*, 2006, Timmins *et al.*, 1998, Naumann *et al.*, 1991a). In addition, the combination of FT-IR and trajectory analysis has proved successful in the identification of metabolic changes in natural multi-organism fermentations (Johnson *et al.*, 2004).

The integration of FT-IR spectroscopy with microscopic imaging techniques enables analysis of the spatial distribution of the functional groups which may be connected to specific components within the cell or tissue sample (Patel *et al.*, 2008, Lewis *et al.*, 1995, Beardall *et al.*, 2001, Lasch *et al.*, 2002).

Previous studies investigated the phenotypic effects of the β-blockers Atenolol and Propranolol on a number of prokaryotes commonly found in the aquatic environment (Chapter 2). The aim of the present study was to explore both the phenotypic and spatial effects of these APIs in a eukaryotic system. The green alga *Micrasterias hardyi* was selected for investigation following exposure to Atenolol and Propranolol; both of which have been found in effluent from wastewater treatment plants and in surface water of freshwater streams. The *Micrasterias* cells were exposed to a range of enantiomeric ratios at a fixed concentration of either Atenolol or Propranolol and analysed using FT-IR microspectroscopy. Through the comparison of both phenotypic and physiological changes, this investigation attempted to understand the effects exerted on the algal cells by these chiral APIs.

3.3 Materials and methods

3.3.1 Culture conditions

Micrasterias hardyi strain CCAP 649/15 was purchased from Culture Collection of Algae and Protozoa (Dunstaffnage Marine Laboratory, Oban, U.K.). *M. hardyi* was cultured in Jaworski's Medium (JM, (Thomkins *et al.*, 1995)) at 15 °C for 8 w, 60 r.p.m. in a Multitron (INFORS HT, Switzerland) orbital shaker, under photo periods of 12 h light and 12 h dark. Nine stock solutions (200 mL) were prepared as detailed in Table 3.1 and an aliquot (1 mL) of each stock solution was combined in 1 L of water and autoclaved (121 °C, 45 min, 15 psi).

M. hardyi was exposed to Propranolol and Atenolol individually at a variety of enantiomeric ratios (100:0, 75:25, 50:50, 25:75 and 0:100 (*R*):(*S*)) in triplicate at a concentration of 30 µg mL^{-1}, with additional control flasks that did not contain any API. The concentration selected for investigation has been shown previously to exert phenotypic effects in *M. hardyi* (Patel *et al.*, 2008); however, the phenotypic effects of Atenolol have not been reported for *M. hardyi*. Following growth, aliquots (20 mL) were taken from each flask and centrifuged at a low speed (3,080 *g*, 4 °C, 40 min) using a Jouan CR 322 (Thermo scientific) to harvest the algal cells without causing damage to the cellular structure. The biomass was resuspended in sterile distilled water (20 mL) and this washing process was carried out a total of 3 times in order to remove any residual API. A final aliquot (100 µL) of sterile distilled H_2O was added to the algal cells and this suspension was used for FT-IR microspectroscopic imaging.

Table 3.1: Stock solutions required for Jaworski's medium and their components.

Stock		g 200 mL^{-1}
1	$Ca(NO_3)_2 \cdot 4H_2O$	4.0
2	KH_2PO_4	2.48
3	$MgSO_4 \cdot 7H_2O$	10.0
4	$NaHCO_3$	3.18
5	EDTA FeNa	0.45
	EDTA Na_2	0.45
6	H_3BO_3	0.496
	$MnCl_2 \cdot 4H_2O$	0.278
	$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	0.20
7	Cyanocobalamin	0.008
	Thiamine HCl	0.008
	Biotin	0.008
8	$NaNO_3$	16.0
9	$NaHPO_4 \cdot 12H_2O$	7.2

3.3.2 FT-IR microspectroscopic imaging

Aliquots (20 µL) of each resuspended sample were applied onto CaF_2 disks (Crystran Ltd, Dorset, U.K.), and air dried prior to analysis. Drying was used to avoid absorption of water in the mid-IR region which may mask biologically important chemical information in the spectra. The spectra were collected in transmission mode using an Equinox 55 module step scan Fourier transform spectrometer coupled to a Hyperion 3000 microscope (Bruker optics Ltd). The microscope was equipped with a 64 × 64 liquid nitrogen cooled mercury cadmium telluride focal plane array (MCT FPA) detector, which allows the simultaneous acquisition of spectral data from a 267 µm × 267 µm area when employing a ×15 objective lens. The FT-IR spectral maps were collected over the wavelength range of 4000-900 cm^{-1} with a spectral resolution of 8 cm^{-1}. The data acquisition time was approximately 25 min for sample and 25 min for background measurements (Patel *et al.*, 2008, Lasch *et al.*, 2002).

Spectral data were collected using routines within the OPUS 4.0 IR imaging software (Bruker Optics). The ratio of the sample and background datasets was taken to generate transmittance images of each algal sample. The spectral images were converted to absorbance within the OPUS software and exported as ASCII files

3.3.3 Data analysis

The ASCII data were imported into Matlab version 7.1 (The MathWorks, Inc., Natick, MA, USA) and a number of in-house algorithms were used in order to explore the spectral data (see Figure 3.1 for workflow).

The pixels for each image were first selected as either “background” or those containing biological information using an automatic routine. Selection criteria were employed using a deterministic approach to set a threshold IR signal; this was compared with the optical image to ascertain if only those pixels of biological relevance were selected. The remaining pixels identified as “background” (or from sections of a second algal cell in the field of view) were set to zero in order to avoid any interference from these pixels which may lead to incorrect interpretation of the subsequent analyses. Following the removal of the background pixels two main approaches were employed for the analysis of the spectral data. Firstly, the spectra for each map were summed in order to provide a composite spectrum for each biological replicate, and

these composite spectra were used for subsequent multivariate analyses. Secondly, functional group mapping was performed on the spectral maps to investigate the spatial distribution of selected functional groups and to gain further insight into the effects of the APIs on the algal cells.

Once functional group mapping had been carried out, a threshold was set for the lipid:amide ratio maps. All of the pixels above the chosen threshold were considered to be areas rich in lipids and those below the threshold were considered to be pixels rich in amide I. The threshold was selected individually for each spectral map (biological replicate) in order to select the correct pixels for analysis (See Figure 3.1 for example). The spectral information for each of the pixels identified as rich in lipids or amide I were summed to form a composite spectrum for each of the biological replicates.

The composite spectra were pre-processed prior to multivariate analysis. Spectral regions which are dominated by CO_2 vibrations arising from the atmosphere (2403 – 2272 cm^{-1} and 683 – 656 cm^{-1}) were removed and filled with a linear trend and the spectra were row normalised (centred about the mean and scaled to unit variance). These pre-processed spectra were used for subsequent multivariate analyses. Principal components analysis (PCA; (Jolliffe, 1986)) is an unsupervised method for reducing the dimensionality of multivariate data whilst preserving the variance; this transformation was performed prior to canonical variates analysis (CVA). CVA is a supervised machine learning method that seeks to minimise within-group variance whilst maximising between-group variance, and can be used in conjunction with PCA to discriminate between groups on the basis of retained principal components (PCs), given *a priori* knowledge of group membership of the spectral replicates (Manly, 1994, Winder *et al.*, 2004). In this study, PC-CVA was performed on the composite spectra in order to investigate any phenotypic changes occurring in the algal cells when exposed to the APIs in their various enantiomeric ratios. The PC-CVA models were constructed with *a priori* knowledge of the biological replicates. In order to ensure these models were not over- or under-trained, validation was performed using the method of full cross-validation, where two of the biological replicates were used for model training with the third projected into the model for cluster validation purposes (Jarvis and Goodacre, 2004a). Finally, CVA also allows statistical significance to be displayed on the scores plots and circles were used to represent the 95% χ^2 confidence region constructed around each group mean based upon the χ^2 distribution with two degrees of freedom (Krzanowski, 1988).

Principal component regression (PCR, (Gemperline *et al.*, 1991)) is a supervised linear regression method and was used to ascertain if any linear relationships between the FT-IR spectra and the drug (*R*):(*S*) exposure existed. In this approach PCA is first conducted and the PC scores regressed against the %(*R*)- enantiomer. This process was again validated as detailed above as the optimal number of PCs (latent variables) against which to calibrate the model needs to be chosen.

Functional group mapping is widely used in the analysis of spectroscopic maps. It utilises the integrated peak area, intensity and position of FT-IR absorption bands attributed to specific functional groups and plots them as a function of X-Y position which can be laid over the optical image of the same sample area. Thus revealing the spatial distribution of chemical constituents within a biological sample and providing information regarding a single peak (functional group) or a ratio of two peaks (McIntosh *et al.*, 1999). Table 3.2 highlights the major functional groups of interest. The IR bands which are attributed to amide I and II are known to be sensitive to protein conformation (Fabian and Mäntele, 2002). The amide I band corresponds to a high fraction of α-helices compared to β-sheet secondary structure in the average protein content (Byler and Susi, 1986). The most intense of the lipid absorptions are found at wavenumbers 3000 – 2800 cm^{-1}. These bands are attributed to the symmetric and asymmetric stretching vibrations of CH_3 and CH_2 groups of the acyl chains.

Functional group mapping was employed in order to investigate localised biochemical effects in the *Micrasterias* cellular components following exposure to the APIs. The intensity and spatial distribution of amide I (1700 – 1600 cm^{-1}) and lipids (3000 – 2800 cm^{-1}) within the algal cells, were analysed in order to obtain information regarding the specific effects of the APIs in the nucleus and chloroplasts respectively. In addition, due to the high levels of amides observed throughout the algal cells, a ratio of lipid:amide mapping was employed in order to investigate those areas high in lipids.

Table 3.2: FT-IR band assignments for the major regions of biological interest of the *Micrasterias hardyi* algal cell.

Wavenumber (cm^{-1}) range	Dominant Compounds
3000-2800	CH_x stretches from Fatty acids
1700-1500	Proteins
(1700-1600)	C=O from Amide I
(1600-1500)	C-N and C-N-H from Amide II
1450-1200	Carboxylic groups of proteins, free amino acids, polysaccharides
(1250-1200)	P-O from RNA/DNA, phospholipids
1200-900	C-O or O-H from Polysaccharides

For each of the spectral maps the IR spectral intensity was converted into a colour code in which blue represents low IR absorbance and red represents high IR absorbance (see Figure 3.1 for example). Gouraud shading is an interpolated shading function and was applied to the maps in order to reduce the appearance of pixilation (Gouraud, 1971). It is piecewise bilinear, in which the value (i.e. colour) for each individual pixel varies linearly and interpolates with the values from the 8 pixels surrounding it.

As described above, PCA reduces the dimensionality of multivariate data whilst preserving the variance. PCA imaging transforms the data according to the variance within the dataset and constructs a map based on the PCA. PCA imaging was performed on the spectral maps and the first principal component (PC) was plotted for each sample; lower PCs were also inspected but contained little additional knowledge. This provides a visualisation of the areas of the algal cells which account for the greatest variance within the analysis. PCA imaging was employed in order to gain insight into the major spatial differences occurring within the cells when exposed to varying enantiomeric ratios of the APIs and the FT-IR bands related to these differences. Most PCA images are constructed on single images and do not allow direct comparison between different images. Therefore PCA was constructed from one alga and the FT-IR spectra from a different *Micrasterias* sample from the whole FPA was projected into its PC scores space. Thus allowing direct comparison of the PC scores from image 1 with PC projected scores from a second image. This method was performed in order to investigate the differences observed in algal cells exposed to differing enantiomeric ratios of the APIs and between the control and API exposed cells.

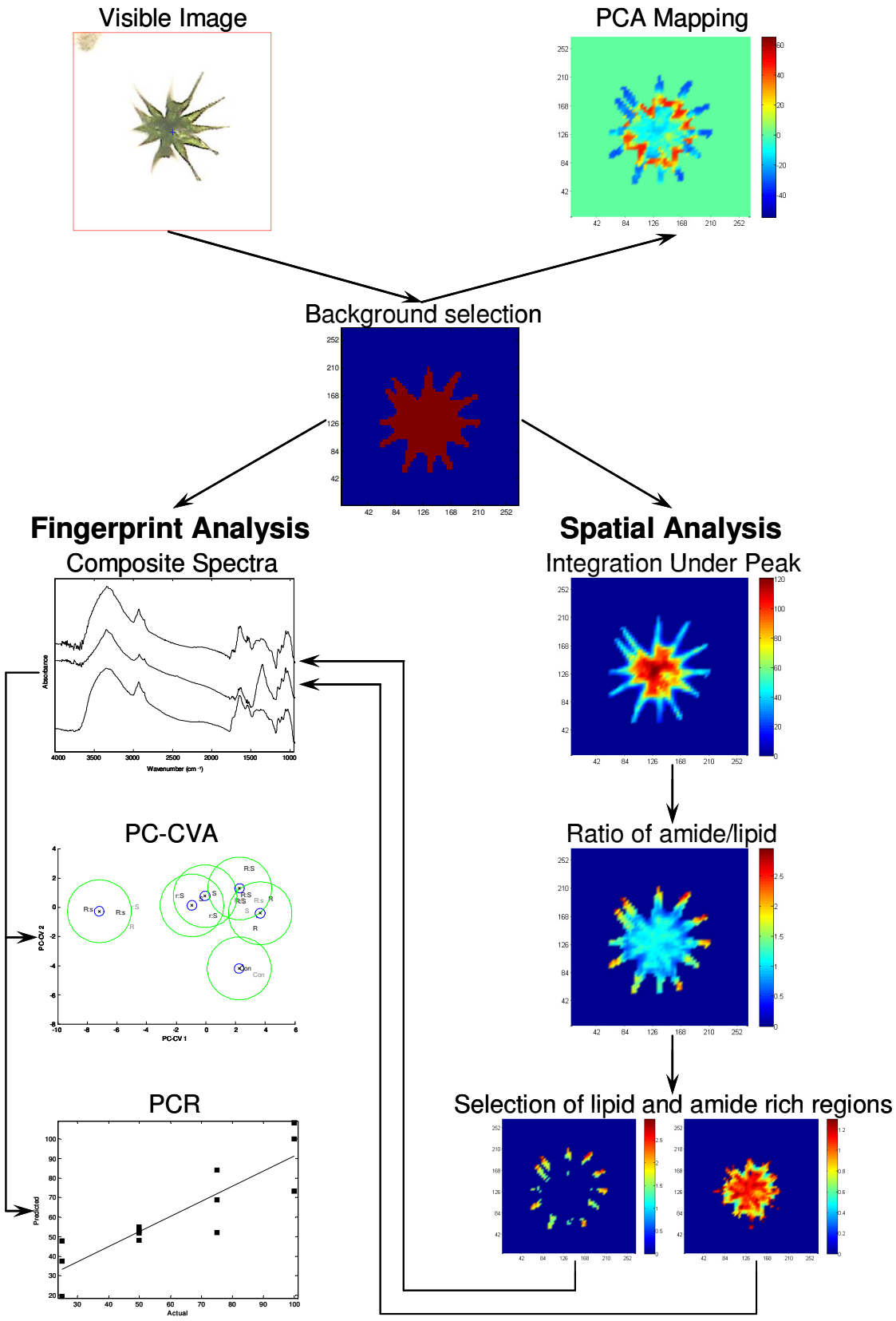


Figure 3.1: Workflow showing the different steps and methods involved in the analysis of the FT-IR spectroscopy data.

3.4 Results and discussion

3.4.1 *Micrasterias hardyi* 649/15

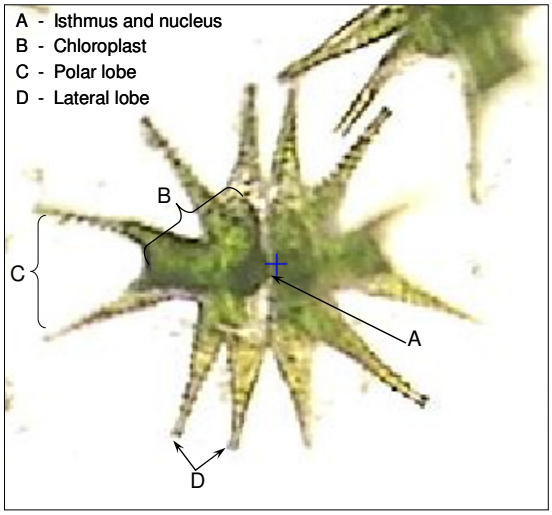


Figure 3.2: Photomicrograph image of *Micrasterias hardyi* 649/15 and the major cellular components.

FT-IR microspectroscopic imaging is well suited to the spectral analysis of *M. hardyi;* a single-celled organism which has a depth of <12 µm, allowing for the collection of spectra without issues regarding over absorbance of the infrared light (Heraud *et al.*, 2005). In addition *M. hardyi* has a large diameter (~80 – 200 µm), which permits imaging of the whole cell using a ×15 objective (Heraud *et al.*, 2005, Patel *et al.*, 2008). The algal cell consists of two semicells, each of which has four main lateral lobes and one larger polar lobe. The semicells are connected by a central constriction termed the isthmus, which accommodates the nucleus. Generally, one large chloroplast which extends into the lobes is found in each semicell (See Figure 3.2). The chloroplasts enable photosynthesis and pyrenoids are highly differentiated areas of the chloroplasts which are surrounded by a thick sheath composed of starch.

3.4.2 Investigating whether Atenolol or Propranolol exert a notable effect on the algal phenotype as judged by FT-IR spectroscopy

Figure 3.3 shows average composite FT-IR spectra of the whole algal cells exposed to Atenolol and Propranolol, and cells not exposed to the APIs (i.e. control); the averages were calculated as detailed above and show information about the whole algal cells

rather than any specific region. The dominant spectral regions typify the patterns commonly associated with the FT-IR spectra of (micro)organisms, and the major spectral bands of biological interest are highlighted in Table 3.2 (Winder *et al.*, 2004, Winder *et al.*, 2006, Zhao *et al.*, 2006, Naumann *et al.*, 1991a). There was no notable difference in the FT-IR spectrum of *M. hardyi* exposed to Atenolol compared to the control sample. However, for cells exposed to Propranolol there appeared to be a reduction in the level of lipids within the algal cells (3000 – 2800 cm^{-1}). In addition a dominant peak was observed at 1500 – 1200 cm^{-1}; a region typically associated with the carboxylic groups of proteins, free amino acids and polysaccharides (Winder *et al.*, 2004, Winder *et al.*, 2006, Zhao *et al.*, 2006, Naumann *et al.*, 1991a). The algal cells were washed during the harvesting stage and it is therefore unlikely that this peak is due to the presence of a contaminant from the Propranolol enantiomers. Further inspection of the FT-IR spectra revealed this peak was only present in samples exposed to 25:75 and 75:25 ratios of (*R*):(*S*)- Propranolol and the (*S*)- enantiomer. The intensity of the peak appeared to be higher in the algal cells exposed to the enantiomeric ratios of Propranolol compared to those exposed to the (*S*)- enantiomer; however there was no notable difference in the intensity of the peak between the enantiomeric ratios. This peak was not observed in the spectra of cells exposed to (*R*)- Propranolol and the racemate (Figure 3.4).

PC-CVA was performed on the composite average spectra from each alga individually in order to investigate any phenotypic changes occurring within the cells when exposed to the APIs and their various enantiomeric ratios. There was no notable phenotypic variation in the metabolic fingerprints of the samples exposed to the enantiomeric ratios of 30 µg mL^{-1} Atenolol (Figure 3.5a), because the different groups used by the CVA algorithm were not separated and the 95% χ^2 tolerance regions overlap with those of the control samples. This analysis clearly illustrates that no discrimination was observed between the control and Atenolol exposed conditions and suggests that Atenolol does not have a phenotypic effect on the algal cells detectable by FT-IR spectroscopy. By contrast, Propranolol exposed samples separate from the control samples across PC-CV 1. To further explore this separation, the Atenolol exposed samples were removed from the dataset and PC-CVA was carried out on the Propranolol exposed and control samples (Figure 3.5b). In this model the Propranolol exposed samples clearly separate from the control samples across PC-CV 1. As CVA uses *a priori* class knowledge (in this case pertaining to the different treatments) the model needs to be validated. Therefore as described above, 2 of the 3 biological

replicates were used for calibration (shown in black) and the 3rd replicate (grey font) was projected into the model. The majority of these projected data are grouped with the appropriate calibration samples indicating that the separation displayed in the model was valid. Moreover, the 95% confidence intervals for the groups are also plotted showing that for the control samples there was a distinct separation in CVA scores space from the Propranolol exposed cultures indicating that Propranolol exerts a phenotypic effect on the algal cells. It was perhaps unsurprising that the (*R*)- and (±)-exposed samples display some separation across PC-CV 2 as these samples are visibly distinct from the other enantiomeric ratios in the FT-IR spectra (Figure 3.4). However, the 95% confidence intervals for these samples indicated that the separation of these samples in this analysis was not wholly statistically significant. Inspection of the loadings for this analysis (data not shown) revealed that this partial separation was due to vibrations at wavenumbers 1509 – 1478 cm^{-1} , which correlate with the dominant peak observed in the average spectra for the algal cells exposed to the (*S*)-enantiomer, and the enantiomeric ratios (25:75 and 75:25 of (*R*) : (*S*)- Propranolol) (Figure 3.4). Previous studies have shown that (±)- Propranolol significantly altered the phenotype of *M. hardyi* during exposure to the API at 60 µg mL^{-1} (Patel *et al.*, 2008). However, the effects of the enantiomers on the algae have not been studied previously, and to our knowledge the phenotypic effect of Atenolol has not been investigated in algae.

3.4.3 Multivariate analysis of localised lipid and amide rich regions within the algae

The multivariate analysis of the composite spectra from the whole algae revealed a phenotypic difference in the algal cells exposed to all of the enantiomeric ratios of 30 µg mL^{-1} Propranolol. In order to investigate any localised effects exerted on the *M. hardyi* cells following exposure to the APIs multivariate analysis was carried out on composite spectra taken from areas of the algal cells which had been identified as being rich in either lipids or amide I relating to the lobes and the central cellular region respectively. Figure 3.6 shows the PC-CVA scores plots for the amide I and lipid rich regions of *Micrasterias* cells exposed to the enantiomeric ratios of 30 µg mL^{-1} Atenolol. It was clear from these plots that Atenolol exposed samples showed no separation from the control samples and this correlates with the PC-CVA results from the composite spectra of the whole algal cells. In addition no notable phenotypic variation was observed in the metabolic fingerprints of these samples for any of the enantiomeric

ratios. By contrast, Propranolol exposed samples showed a clear separation from the control samples across PC-CV 1 in both amide I and lipid regions (Figure 3.7). As observed in the PC-CVA carried out on the composite spectra from the whole algal cells, the (*R*)- and (±)- Propranolol exposed samples show some separation across PC-CV 2 but this separation was also not considered to be statistically significant. In addition, the loadings for these analyses demonstrated that the observed separation was attributable to vibrations at wavenumbers 1489 – 1478 cm^{-1}. The FT-IR spectra for both the amide I and lipid rich regions of the Propranolol exposed cells exhibited a similar phenotypic response, suggesting that the API does not exhibit a localised effect within the cellular components in the FT-IR spectra.

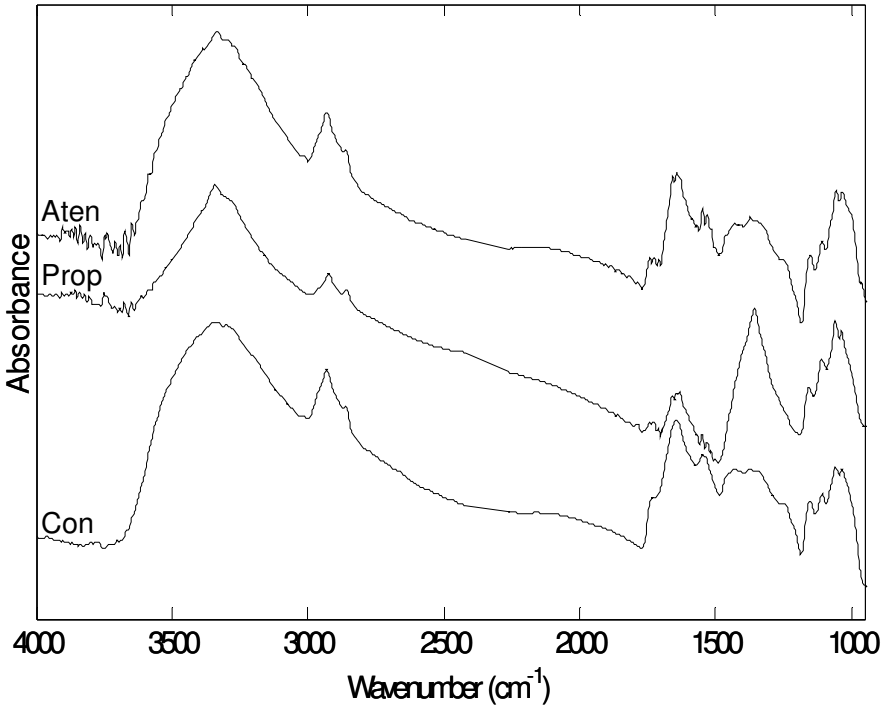


Figure 3.3: Average FT-IR spectra for *Micrasterias hardyi* cells exposed to 30 µg mL^{-1} Atenolol and Propranolol, and those grown in the absence of API (Con = control). The spectra are offset on the Y axis to allow easier visualisation, and the absorbance is relative.

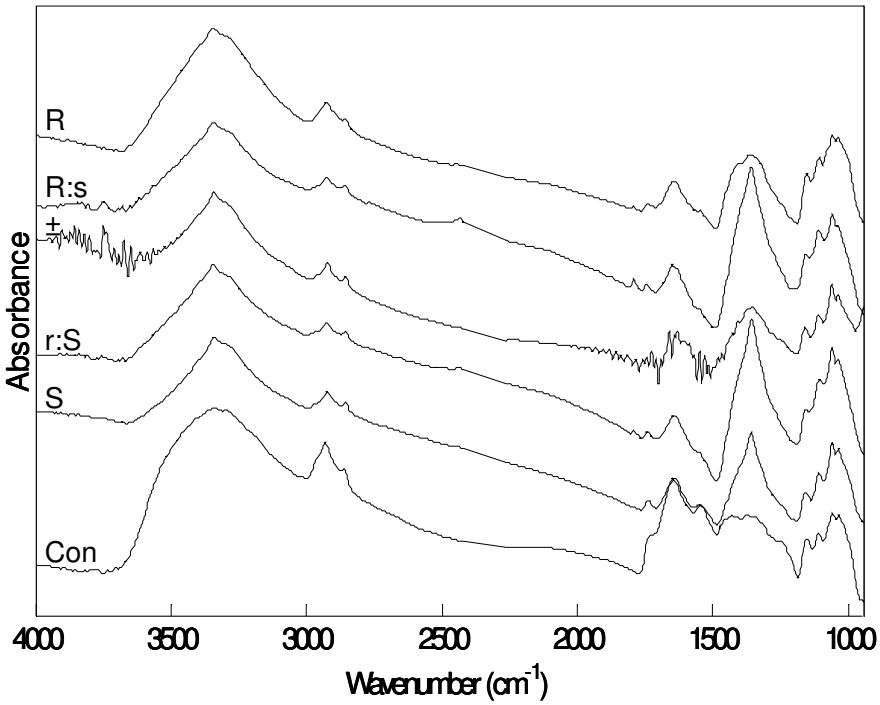


Figure 3.4: Average FT-IR spectra for *Micrasterias hardyi* cells exposed to varying ratios of 30 µg mL^{-1} Propranolol, and those grown in the absence of API (i.e. control). The spectra are offset on the Y axis to allow easier visualisation, and the absorbance is relative. Ratios 100:0, 75:25, 50:50, 25:75 and 0:100 (*R*) : (*S*) are represented by points R, R:s, ±, r:S and S respectively. Control samples are labelled "Con".

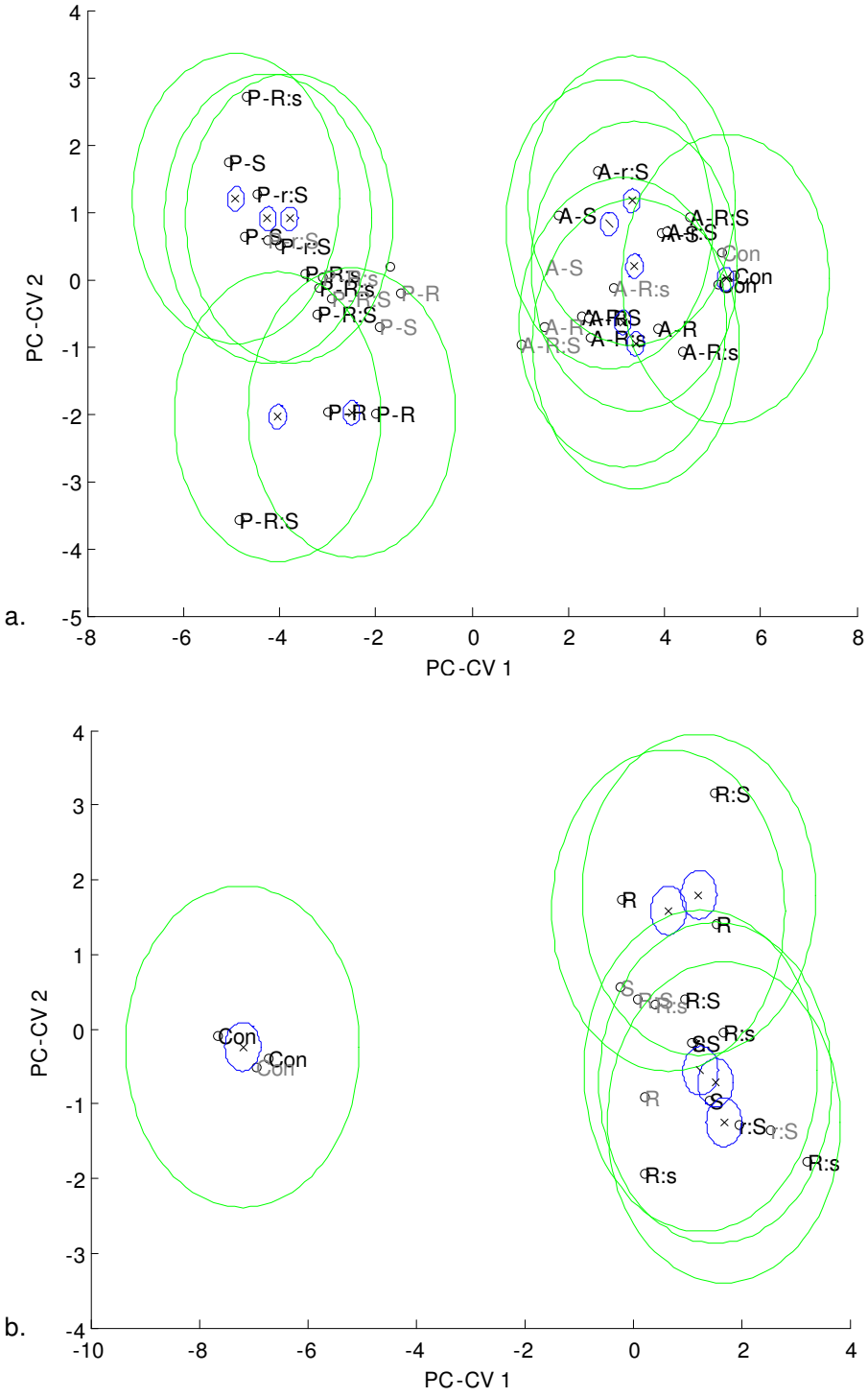


Figure 3.5: PC-CVA scores plots for *Micrasterias hardyi* exposed to enantiomeric ratios of 30 µg mL^{-1} Atenolol and Propranolol (a.) and Propranolol (b.). Plots were constructed using the composite spectra for the whole algal cells. Points shown in black represent the 2 biological replicates that were used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Blue circles represent the 95% confidence interval about the group centroid, and green circles the 95% confidence region about the group sample population. Plots a and b were constructed using 5 and 3 PCs and account for 94.1 and 77.4 % explained variance respectively. Points R, R:s, R:S, r:S, S and Con represent samples exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*) and control samples respectively.

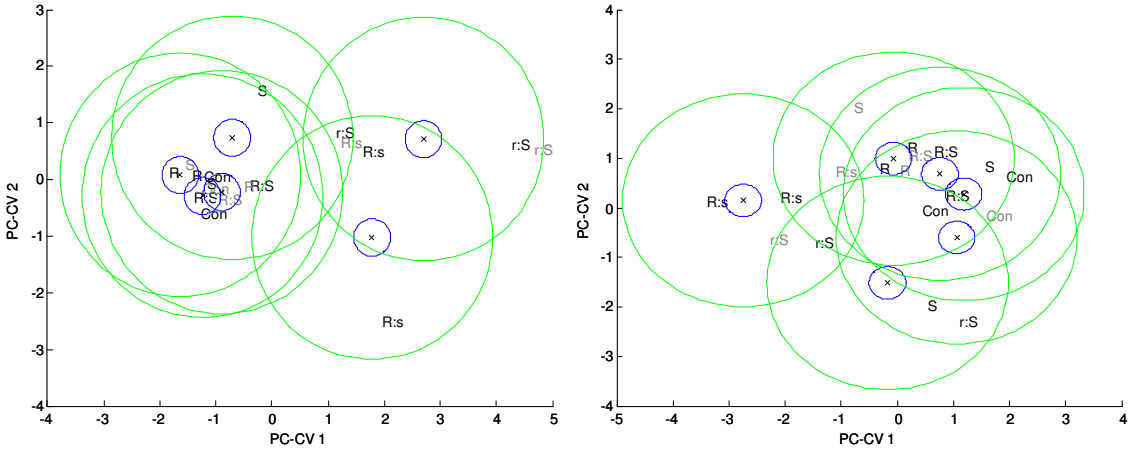


Figure 3.6: PC-CVA scores plots for amide (LHS) and lipid rich (RHS) regions of *Micrasterias hardyi* exposed to enantiomeric ratios of 30 µg mL^{-1} Atenolol. Points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Blue circles represent the 95% confidence interval about the group centroid, and green circles the 95% confidence region about the group sample population. The plots were constructed using 3 PCs and account for 86.2 and 92.7 % explained variance respectively (LHS and RHS). Points R, R:s, R:S, r:S, S and Con represent samples exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*) and control samples respectively.

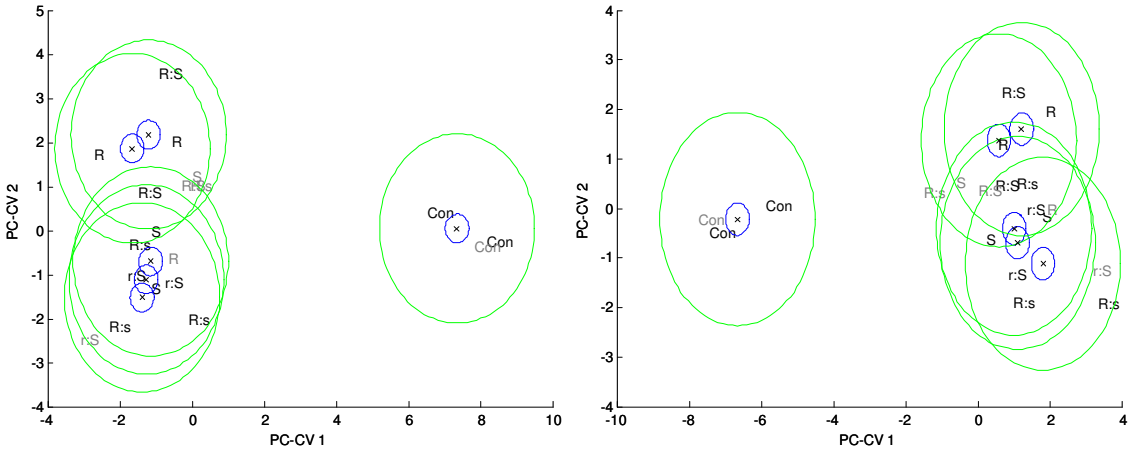


Figure 3.7: PC-CVA scores plots for amide (LHS) and lipid rich (RHS) regions of *Micrasterias hardyi* exposed to the enantiomeric ratios of 30 µg mL^{-1} Propranolol. Plots were constructed using the composite spectra for the whole algal cells. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Blue circles represent the 95% confidence interval about the group centroid, and green circles the 95% confidence region about the group sample population. The plots were constructed using 4 PCs and account for 93.6 and 88.0 % explained variance respectively (LHS and RHS). Points R, R:s, R:S, r:S, S and Con represent samples exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*) and control samples respectively.

3.4.4 Investigating supervised learning for quantification of the enantiomer effect of the APIs

As described above principal component regression (PCR) is a supervised learning method that effects quantitative analysis and was carried out in order to investigate the relationship between the percentage of enantiomer to which the algae had been exposed to and the phenotypic response. PCR was trained to quantify the percentage of (*R*)- enantiomer to which the samples had been exposed and the control samples were removed from the dataset prior to analysis in order to prevent any subsequent bias within the analyses. The PCR results (Figure 3.8) appeared to show some correlation, however the errors for these models were rather high. The root mean squared (RMS) errors for the Atenolol and Propranolol exposed cells were +/- 83.48 and +/- 99.54 respectively demonstrating that it was not possible to predict the ratio of enantiomers to which the algal cells had been exposed for either of the APIs from the FT-IR spectroscopy data. In addition PCR was also carried out on the composite spectra of the amide I and lipid rich regions of the cells exposed to Atenolol and Propranolol (data not shown). However, the results were consistent with those from the whole algal cells and it was not possible to model the relationship between the ratio of enantiomers to which the algal cells had been exposed and the phenotypic response for either of the APIs. PCR is capable of modelling linear relationships and therefore it was not perhaps surprising that a correlation was not observed in these analyses as this would rely on a mainly linear relationship between enantiomeric ratio and phenotypic response within the algal cells. Therefore a correlation would presume that the majority of the biological effect exerted on the algae was due to one of the enantiomers. Whilst it is known that the (*S*)- enantiomer of Propranolol exerts the majority of the β-blocking effect in man, the effects of this API and its enantiomers have not previously been studied in algae (Barrett and Cullum, 1968, Walle *et al.*, 1984). The multivariate data analyses indicated that both of the Propranolol enantiomers exert a phenotypic effect on *M. hardyi*, and appeared to exert differing phenotypic effects. Therefore, the PCR algorithm may not be able to model such complex phenotypic variations. In addition, the initial results from the PC-CVA suggested that Atenolol did not exert a noticeable phenotypic effect on the composite spectra of the algal cells and it is therefore unlikely that a linear relationship between %(*R*)- enantiomer and phenotypic response may be modelled for Atenolol.

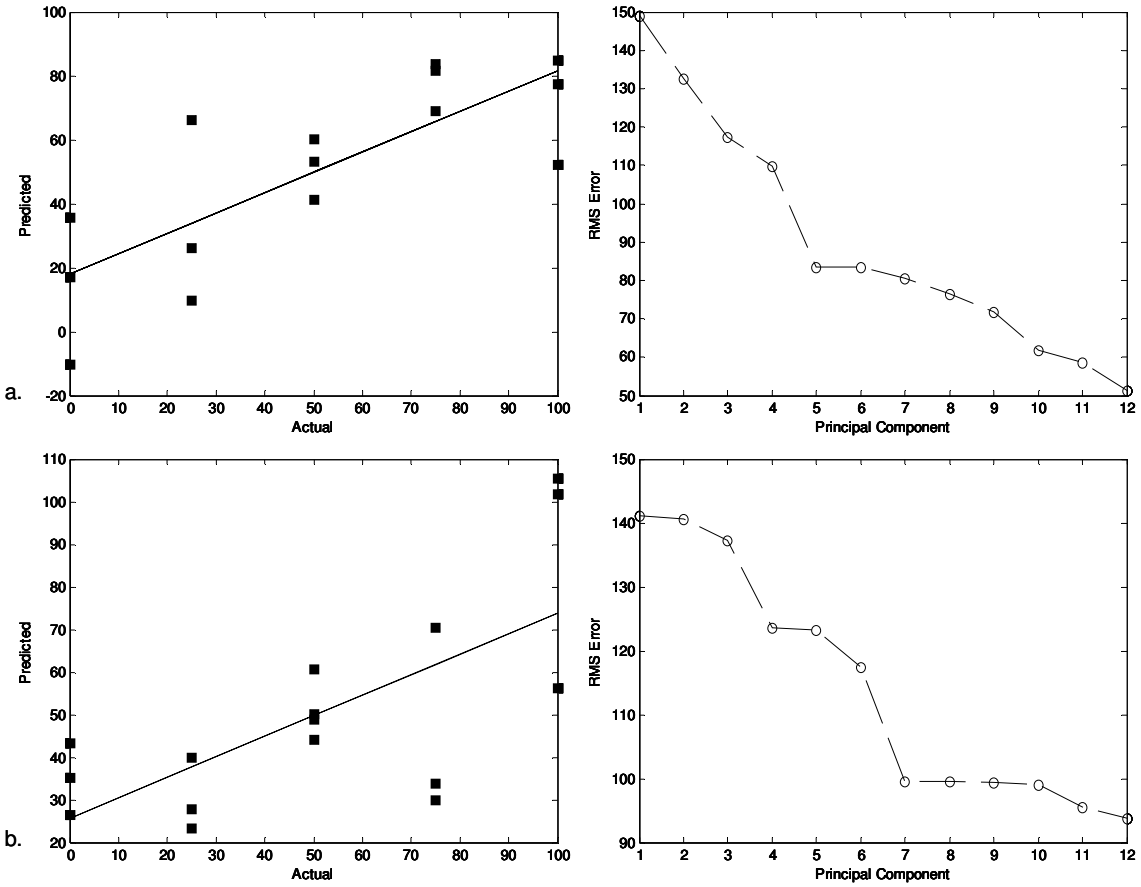


Figure 3.8: Principal component regression (PCR) carried out on composite spectra from whole *Micrasterias hardyi cells* exposed to varying ratios of 30 µg mL^{-1} Atenolol (a.) and Propranolol (b.). The models were built using two out of the three biological replicates and were given knowledge of the percentage of the (*R*)- enantiomer to which the algal cells had been exposed. The model for Atenolol and Propranolol were built using 5 and 7 PCs and had errors of +/- 83.48 and +/- 99.54 respectively.

3.4.5 Investigating whether the APIs effect the distribution and concentration of the chemical constituents within the algal cells

FT-IR microspectroscopy was carried out on whole algal cells exposed to the enantiomeric ratios of 30 µg mL^{-1} Atenolol or Propranolol in order to investigate any potential spatial effects on *M. hardyi*. Spectral maps were constructed for the total IR signal absorbed by the *Micrasterias* cell at wavenumbers 4000 – 900 cm^{-1}. In addition, functional group mapping was carried out on these spectral maps in order to provide a visualisation of the spatial distribution of the chemical constituents of the algal cells when exposed to the APIs and their enantiomeric ratios. Maps were constructed for the integrated peak areas under the amide I peak at wavenumbers 1700 – 1600 cm^{-1} and the lipid peak at wavenumbers 3000 – 2800 cm^{-1}. In addition, due to the high levels of amides observed throughout the algal cells, a ratio of lipid:amide I was employed in order to normalise to the same protein level. Finally, maps were constructed for the integrated peak areas under the peak observed at wavenumbers 1500 – 1200 cm^{-1} for the Propranolol exposed cells. Figure 3.9 and Figure 3.10 show the functional group maps for a selection of samples exposed to the enantiomeric ratios of 30 µg mL^{-1} Atenolol and Propranolol respectively and Figure 3.11 shows the functional group maps for the Propranolol exposed cells at wavenumbers 1500 – 1200 cm^{-1}.

The morphological components of the algal cells can be clearly observed in the total signal maps of the *Micrasterias* control cells (Figure 3.9a). These relate well to the relevant photomicrograph images and previous reports which have shown FT-IR microspectroscopy to be capable of resolving these cellular components (Heraud *et al.*, 2005, Patel *et al.*, 2008). Integration of the amide I band revealed a high concentration of amides throughout the centre of the cell and partially extending into the lobes. A particularly high intensity of the amide I band was observed in the isthmus, the region in which the nucleus is found (Meindl *et al.*, 1994, Heraud *et al.*, 2005). This correlates with the known composition of the nucleus, which has high levels of histones and low levels of lipids (Laskowski *et al.*, 1997, Heraud *et al.*, 2005). The presence of amides in the semi-cells extending into the lobes may be attributed to the presence of the chloroplasts, in particular the pyrenoids which are highly differentiated areas of the chloroplast responsible for the fixation of CO_2. Pyrenoids are known to contain high levels of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO); the key enzyme of the Calvin Cycle which is the main process for carbon fixation in plants,

algae and photosynthetic bacteria (Laskowski *et al.*, 1997, Venter *et al.*, 2004). Both RubisCO and histones are known to contain a high proportion of α-helices within their structure, and it is therefore likely that the intense amide I bands observed in the region of the chloroplasts and nucleus respectively was due to the presence of these proteins (Laskowski *et al.*, 1997). High levels of lipids are also observed throughout the control cells particularly within the centre of the semicells in the region in which the chloroplasts are observed. It has been suggested that this may be attributed to the closely packed membranes termed thylakoids found within the chloroplasts (Meindl, 1993, Heraud *et al.*, 2005). The spectral maps demonstrating the lipid:amide I ratio display highly differentiated areas in the end of the lobes which contain high concentrations of lipids, and have previously been suggested to be lipid storage bodies (Heraud *et al.*, 2005).

The photomicrograph images for *Micrasterias* exposed to the enantiomeric ratios of 30 μg mL^{-1} Atenolol do not reveal a notable difference in the appearance of the algal cells when compared to the control cells. This was reflected in the maps of the total FT-IR signal which appear similar to those of the control cells for all of the ratios. In addition the functional group maps for the integration under the amide I and lipid peaks do not display a difference from the control cells (Figure 3.9). This was not unexpected as the multivariate analysis did not show a difference between the Atenolol exposed and the control cells based on the composite spectra. Figure 3.10 shows the photomicrograph images and the functional group maps for *M. hardyi* control cells and a selection of those exposed to the enantiomeric ratios of 30 μg mL^{-1} Propranolol. It is clear from the photomicrograph images that Propranolol exerts a notable effect on the algal cells. The cells appear to be chlorotic following exposure to all of the selected enantiomeric ratios. In addition, the maps of the total FT-IR signal display a general reduction in intensity for all of the enantiomeric ratios, which was not surprising given the appearance of the algal cells. This trend was also observed for the amide I band which exhibits a dramatic reduction in amides throughout the Propranolol exposed cells compared to the control cells and corresponded to the chlorotic appearance of the algal cells. The *Micrasterias* cells exposed to the racemate (Figure 3.10d) display the greatest reduction in amide I. Although Figure 3.10 only displays one of the replicates from each of the experimental conditions, these trends are consistent throughout all of the replicates. The functional group mapping for the lipid:amide I ratio also reveals a notable effect in the Propranolol exposed cells. A general reduction of lipids was observed in the lipid storage bodies found at the end of the lobes in the algal cells

(Heraud *et al.*, 2005). Microalgae are well known for the production of lipids under stress conditions (Torres *et al.*, 2008), however Propranolol is a lipophilic compound (Lemmer *et al.*, 1985, Westerlund, 1985) and it was not surprising to note that it interacts with the lipid components of the *M. hardyi* cells. The reduction of lipids in these areas was particularly noticeable in the *Micrasterias* cells exposed to both (*R*)- and (±)- Propranolol (Figures Figure 3.10b and Figure 3.10d respectively). This observation was consistent with earlier analyses as these samples were notably different in the FT-IR spectra and showed some separation from the remaining samples in the PC-CVA carried out on the composite spectra for the whole algal cells. In addition, the photomicrograph images appear to reflect this enhanced effect on the algal cells by (*R*)- and (±)- Propranolol in comparison with the other enantiomeric ratios. The (*R*)- enantiomer is known to have a predominantly membrane stabilising effect in man, whilst (*S*)- Propranolol accounts for the majority of the β-blocking effect (Barrett and Cullum, 1968, Potter and Sweetland, 1967, Hanna and Evans, 2000, Walle *et al.*, 1984). It is therefore likely that the enhanced effect on the lipids may be due to interactions with the algal membrane and the lipid components. In addition the racemate possesses differing physical properties from the enantiomers and it may be expected that it could produce different effects in biological systems (Chickos *et al.*, 1981, Secor, 1963).

The FT-IR spectra of cells exposed to the (*S*)- enantiomer and the enantiomeric ratios (25:75 and 75:25 of (*R*) : (*S*)- Propranolol), compared to the (*R*)- enantiomer and the racemate, revealed an additional feature at 1500 – 1200 cm^{-1} (Figure 3.4) which was also observed from the PC-CVA (Figure 3.5b). Therefore in order to investigate this further, functional group maps for *M. hardyi* control cells and a selection of those exposed to the enantiomeric ratios of 30 µg mL^{-1} Propranolol for the total signal and the region at wavenumbers 1500 – 1200 cm^{-1} were constructed (Figure 3.11). The IR signal for this peak was most intense in the centre of the cells, the region in which the chloroplasts and the nucleus are located. The maps correspond to those for the amide I, lipid and lipid: amide I ratios maps (Figure 3.10) and show that Propranolol exerted a notable effect on the algal phenotype. Comparing the series of Propranolol exposures as the amount of (*S*)- enantiomer it was not readily possible to determine any differential effect on the algal cells when exposed to the pure enantiomers and the various enantiomeric ratios. However, on closer inspection it was possible to see that the overall signal strength for the integration between 1500 – 1200 cm^{-1} in the central region of the cells decreased as the (*S*) : (*R*) ratio increased, which may suggest that

the (*S*)- enantiomer had a more profound effect on the chloroplast or nucleus of the algae.

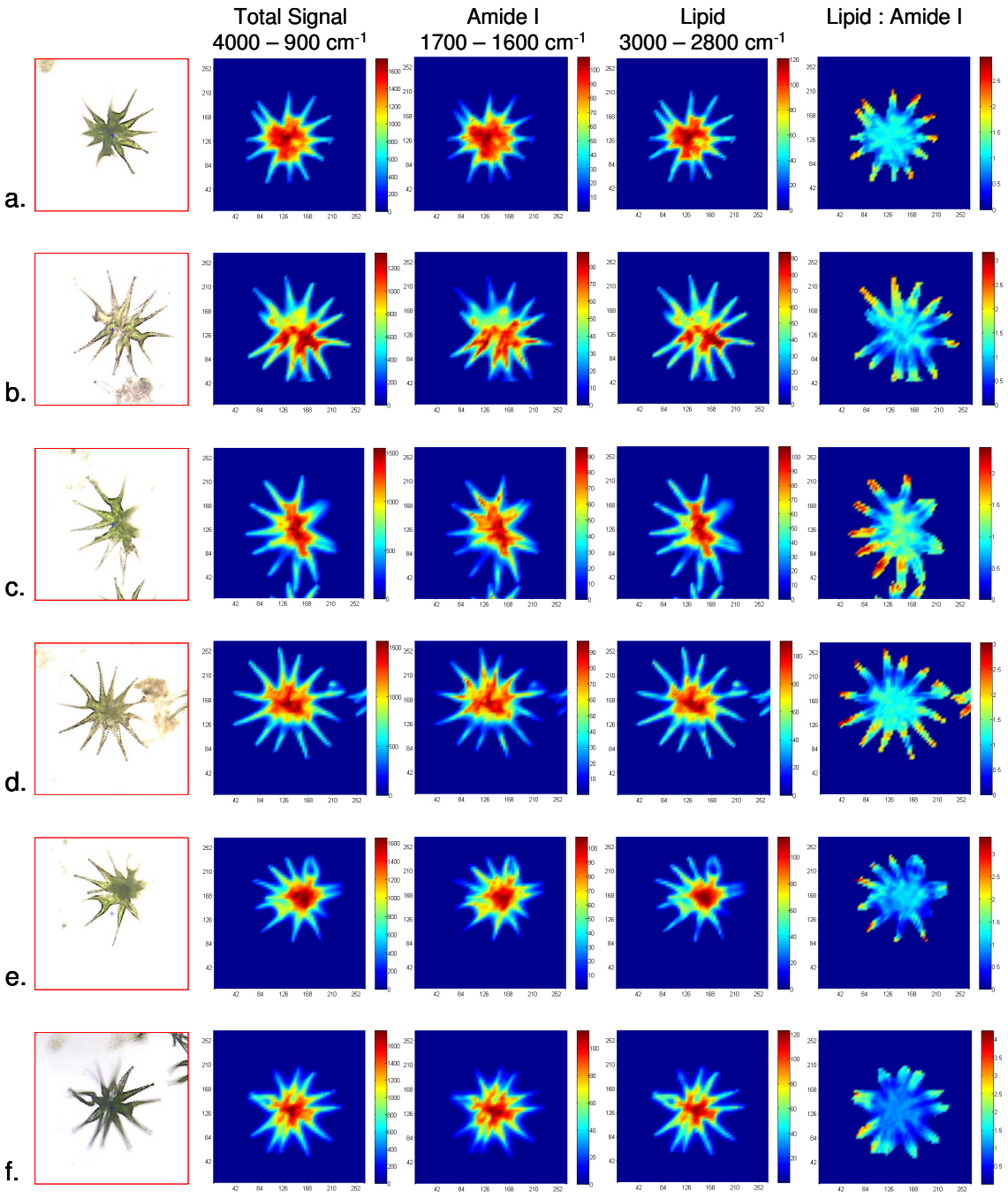


Figure 3.9: Photomicrograph images and functional group maps for *Micrasterias hardyi* cells exposed to varying ratios of 30 µg mL^{-1} (*R*) : (*S*)- Atenolol. Row a. shows results for control cells. Rows b-f represent cells exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*) respectively.

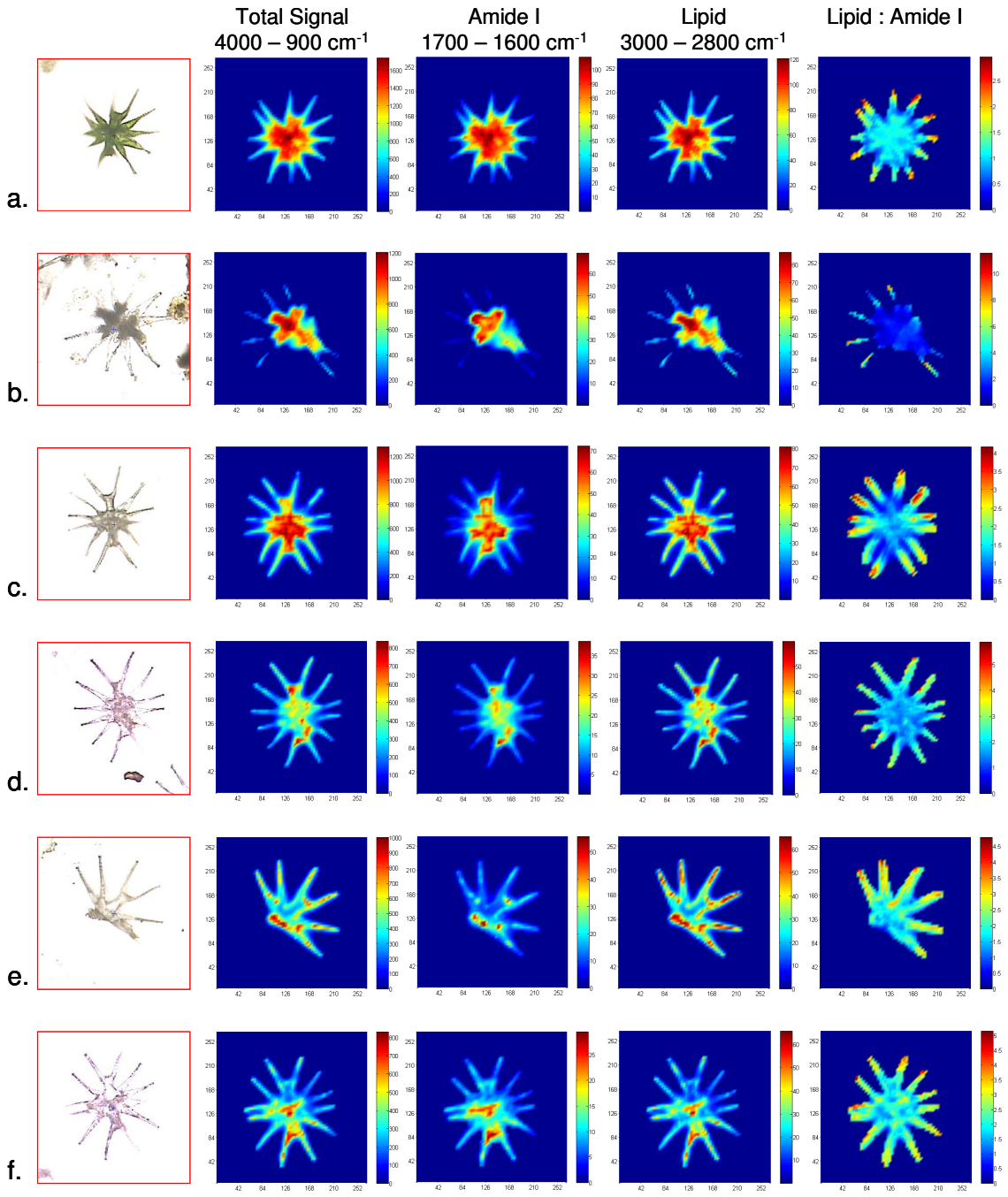


Figure 3.10: Photomicrograph images and functional group maps for *Micrasterias hardyi* cells exposed to varying ratios of 30 µg mL^{-1} (*R*) : (*S*)- Propranolol. Row a. shows results for control cells. Rows b-f represent cells exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*) respectively.

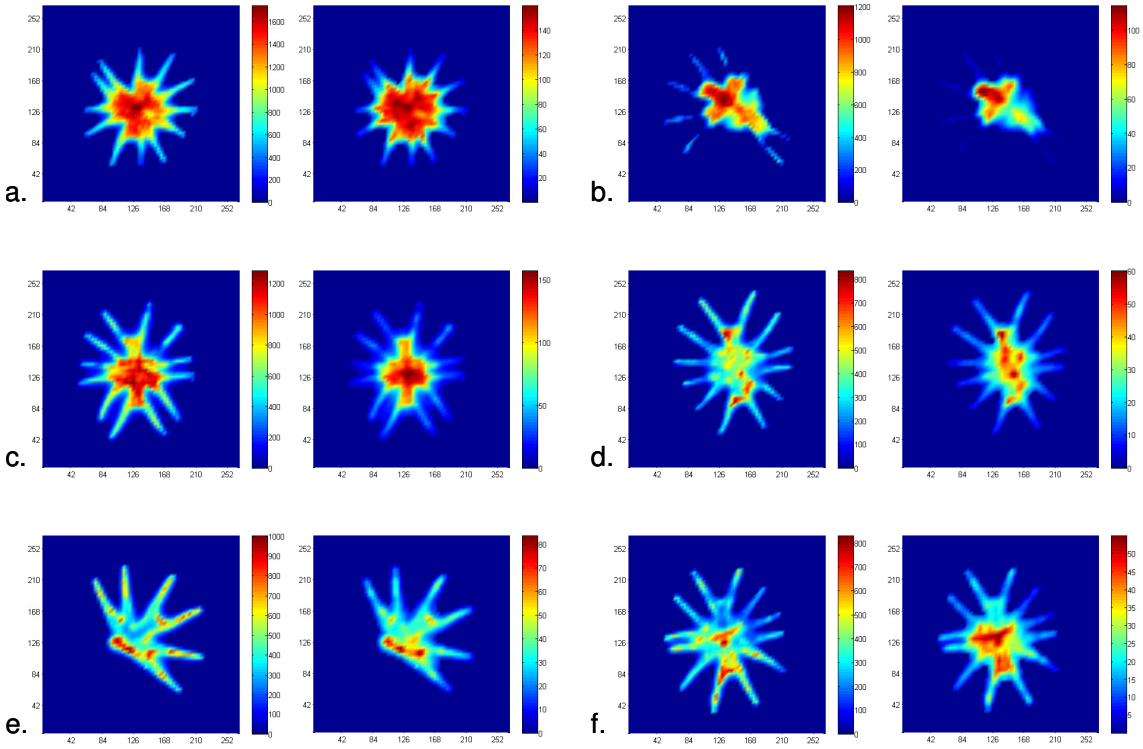


Figure 3.11: Functional group maps for *Micrasterias hardyi* cells exposed to varying ratios of 30 µg mL^{-1} (*R*) : (*S*)- Propranolol. Maps for the total IR signal are shown on the left and functional group maps for the peak found at wavenumbers 1500-1200 cm^{-1} are shown on the right hand side. Control samples are represented by (a.). Samples b-f represent cells exposed to 100:0, 75:25, 50:50, 25:75, 0:100 (*R*) : (*S*)- Propranolol respectively.

3.4.6 PCA Mapping

Functional group mapping was employed to investigate the phenotypic effects of the APIs on specific regions of the FT-IR spectra. Therefore, in order to investigate the whole spectrum rather than specific regions PCA mapping was performed where PCA was conducted on the whole algae and the first PC score plotted on the chemical map. As the scores may have a different scale these were all normalised to lie within the same range. However, on inspection of the PCA chemical maps it was clear that there was no obvious visual difference between the controls and any of the Atenolol exposed cells (Figure 3.12). By contrast, for the Propranolol treated cells the overall 'intensity' of the maps was generally lower in the treated cells compared to the controls, but no specific differences could be observed in the PC loadings matrices (Figure 3.13). Therefore, chemical mapping using PCA did not bring anything more to the investigation compared to functional group mapping.

Projections were used to highlight differences from controls and API exposed cells. In this process PCA was constructed on one alga and spectra from other algae were projected into this PCA space. The idea behind this is that if there is a phenotypic difference between the algae then the spectra will also be different and so when projected into the PCA space these projected scores should fall outside the PC scores range. This was tested within a replicate group; i.e. PCA was conducted on one of the control algae and the other two replicates projected in. The images (data not shown) were equivalent, showing that there was no phenotypic difference. However, when an alga treated with Atenolol or Propranolol was projected in it also looked very similar to the control. Thus the project routine also did not identify any enantiomeric specific changes.

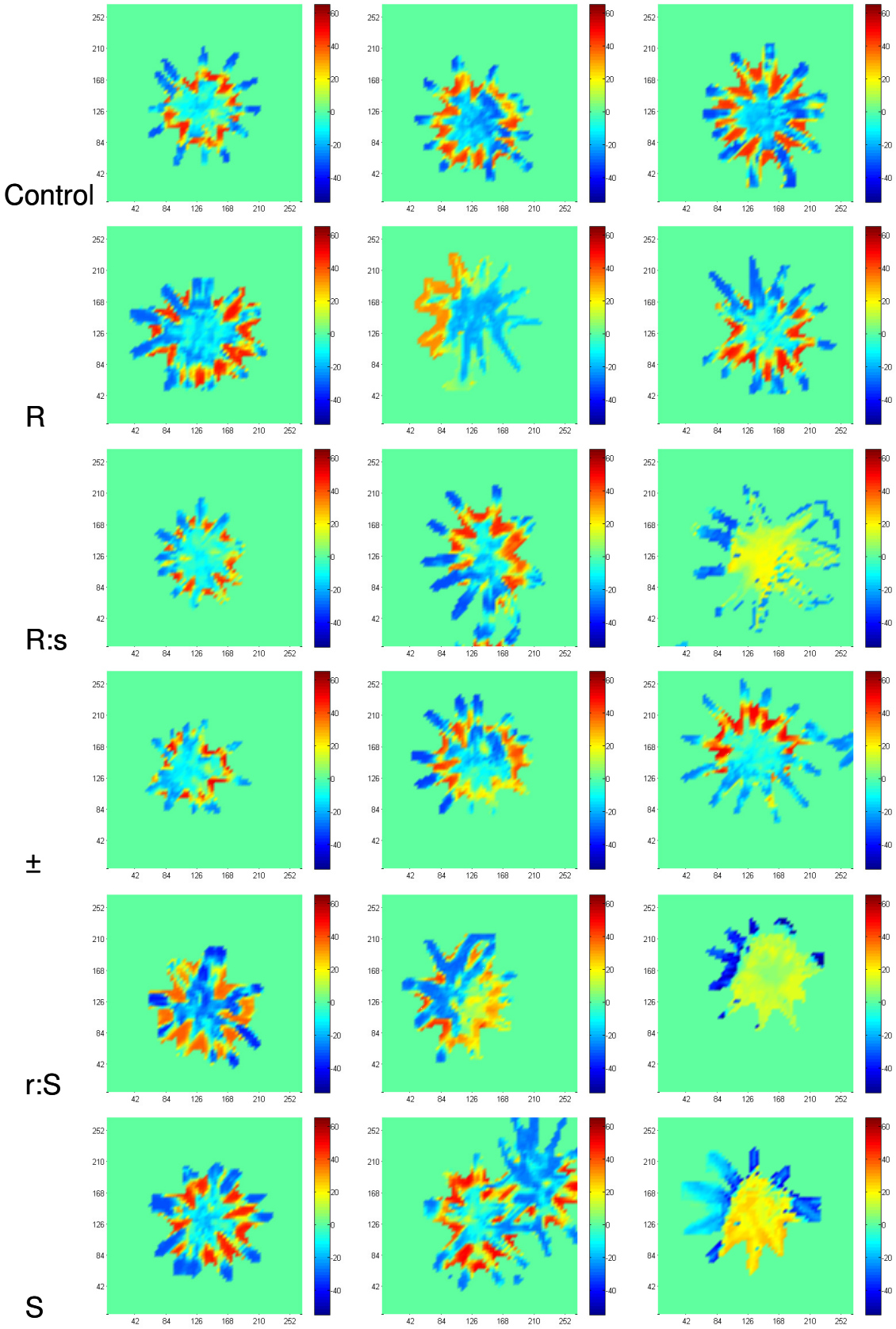


Figure 3.12: PCA maps for *Micrasterias hardyi* cells exposed to varying enantiomeric ratios of 30 µg mL^{-1} Atenolol. PCA was carried out on the whole algal cells and the first PC is plotted for each sample. Samples exposed to 100:0, 75:25, 50:50 25:75 and 0:100 (*R*) : (*S*)- Atenolol are labelled R, R:s, ±, r:S and S respectively.

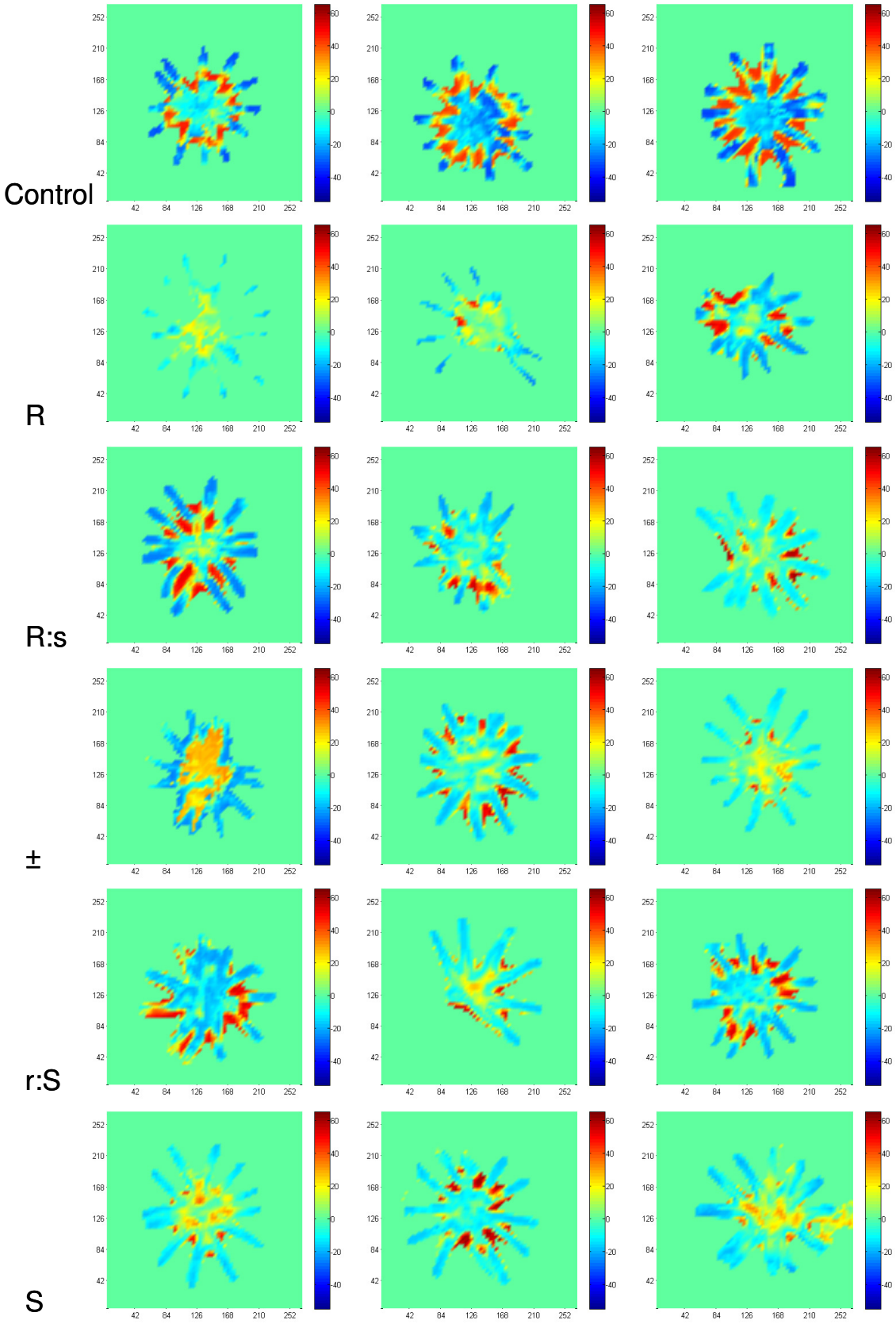


Figure 3.13: PCA maps for *Micrasterias hardyi* cells exposed to varying enantiomeric ratios of 30 µg mL^{-1} Propranolol. PCA was carried out on the whole algal cells and the first PC is plotted for each sample. Samples exposed to 100:0, 75:25, 50:50 25:75 and 0:100 (*R*) : (*S*)-Propranolol are labelled R, R:s, ±, r:S and S respectively.

3.5 Conclusions

The Propranolol racemate has previously been shown to exert phenotypic effects on *M. hardyi* at 30 µg mL^{-1}. In this investigation FT-IR microspectroscopy was employed to investigate the chirality specific effects of the enantiomeric ratios of the APIs Propranolol and Atenolol on the total cellular fingerprints of the algal cells. In addition the spatially localised effects of the APIs and their enantiomer ratios on the biochemical components of the algal cells were also investigated.

No notable difference was observed in *M. hardyi* exposed to Atenolol compared to the control samples as analysed by FT-IR spectroscopy. The spectra appeared to be indistinguishable from those of the control when subjected to multivariate analyses. In addition, the functional group mapping did not reveal any localised phenotypic differences in the Atenolol exposed cells for any of the enantiomeric ratios.

By contrast, as observed previously, a phenotypic effect on the algal cells was observed when exposed to Propranolol. A clear difference was observed in the FT-IR spectra of the algal cells exposed to 30 µg mL^{-1} Propranolol at all of the enantiomeric ratios and a dominant peak was observed in the FT-IR spectra at wavenumbers 1500 - 1200cm^{-1} which was not observed in cells exposed to (*R*)- Propranolol or the enantiomer. Inspection of the photomicrographs and functional group maps revealed that the algal cells exposed to (*R*)- and (±)- Propranolol were particularly affected by the API. However, these samples were not satisfactorily separated from the remaining Propranolol exposed samples in the multivariate analysis of the spectral data.

The greatest reduction in the amide region of the FT-IR spectra was observed in cells exposed to the racemate. However, it was not possible to determine any specific localised effects in the algal cells relating to the cellular components. Analysis of both the lipid rich (lobes) and amide I rich (centre of cell housing the nucleus and chloroplasts) areas did not reveal significantly different responses from those observed throughout the cell. It is therefore likely that Propranolol exerts a general effect on the whole algal cell.

4 FT-IR spectroscopy as a tool for monitoring the effects of environmental pollutants on complex biological communities

4.1 Abstract

A number of methods are employed in the removal of harmful substances both from domestic and industrial wastewaters prior to the release of effluents into the aquatic environment. These include chemical and physical remediation processes such as ozonation, and biological processes such as biofilm and membrane bioreactors and activated sludge systems. In this investigation we acquired activated sludge from an industrial bioreactor capable of degrading phenol. The sludge was incubated in the laboratory and monitored for the phenotypic effects exerted on the sludge by a range of environmental pollutants over a 48 h period. Multiple samples were taken across the time course and analysed by Fourier transform infrared (FT-IR) spectroscopy. FT-IR spectroscopy was employed as a whole organism fingerprinting approach to monitor biochemical changes in the bacterial cells during exposure to the chosen compounds. The ability of the activated sludge to degrade these compounds during the time course was also investigated. The microbial community was able to degrade 5 mM phenol within 36 h and this was accompanied by a detectable biochemical change in the FT-IR fingerprint related to the cellular phenotype of the microbial community. However, no notable phenotypic effect was observed in the microbial community when exposed to toluene, Atenolol and Propranolol. FT-IR spectroscopy, when combined with chemometric analysis is a very useful high-throughput screening approach for assessing the metabolic capability of complex microbial communities. The study demonstrates that the acquired ability of the microbial community is specific to phenol and it is not transferable to degrade compounds of a similar structure.

4.2 Introduction

The release of effluents containing harmful chemicals into receiving waters may cause long term effects in aquatic biota and the occurrence and fate of environmental contaminants such as aromatic hydrocarbons, personal care products and pharmaceuticals in the aquatic environment has drawn increasing attention over the last decade (Daughton and Ternes, 1999, Jjemba, 2006, Alvarez and Vogel, 1991, Nahar *et al.*, 2000, Ternes, 1998, Halling-Sørensen *et al.*, 1998, Kümmerer, 2009). However, little is currently known about the fate of these compounds following their intended use (Kolpin *et al.*, 2002) and there are a number of concerns regarding the presence of these compounds in the environment. These concerns include risks to physiological and reproductive processes (Huggett *et al.*, 2002, Kime and Nash, 1999, Purdom *et al.*, 1994, Larsson *et al.*, 1999), the development of antibiotic-resistance in bacteria (Khachatourians, 1998, Smith *et al.*, 1999, Gilliver *et al.*, 1999), and the likely increase in the toxicity of chemical mixtures (Abuhamed *et al.*, 2004, Sumpter and Jobling, 1995). For numerous substances, the potential effects on both man and aquatic ecosystems are not fully understood (Ayscough *et al.*, 2000, Halling-Sørensen *et al.*, 1998, Daughton and Ternes, 1999). A number of harmful compounds have been indentified as causing undesirable effects in aquatic organisms. For example the synthetic oestrogen ethinyloestradiol and the β-blocking agent Propranolol have been shown to exert a significant risk to reproduction in fish (Huggett *et al.*, 2002, Kime and Nash, 1999, Larsson *et al.*, 1999, Purdom *et al.*, 1994). In addition phenol has been shown to produce detrimental effects in aquatic organisms (Kühn *et al.*, 1989).

A number of methods are employed in the removal of harmful substances both from domestic and industrial wastewaters prior to the release of effluents into the aquatic environment. These include chemical and physical remediation processes such as ozonation (Huber *et al.*, 2003, Ternes *et al.*, 2003, Ahmaruzzaman and Sharma, 2005, Vázquez *et al.*, 2007, Qian *et al.*, 1994), and biological processes such as biofilm and membrane bioreactors (Jeong and Chung, 2006, Morper and Wildmoser, 1990, Zuehlke *et al.*, 2003, Kloepfer *et al.*, 2004) and activated sludge systems (Sutton *et al.*, 1999, Suzuki *et al.*, 2002, Fent *et al.*, 2006). Activated sludge systems are employed worldwide for the treatment of wastewaters and offer an effective alternative to the traditional chemical and physical methods which are known to produce hazardous by-products (Singleton, 1994, Suzuki *et al.*, 2002). A key characteristic of the activated sludge process is the recycling of a large proportion of the biomass, resulting in a

significant community of microorganisms capable of efficiently processing wastewaters to degrade hazardous contaminants (Bitton, 2005). The use of activated sludge has been shown to be more efficient than other methods for the removal of environmental pollutants such as pharmaceuticals (Drewes *et al.*, 2002, Quintana *et al.*, 2005), personal care products and aromatic hydrocarbons such phenols and toluene (Singleton, 1994, Sutton *et al.*, 1999, Tisler *et al.*, 1999) from wastewaters.

FT-IR spectroscopy is a phenotyping technique which has previously been used to generate informative metabolic fingerprints from biological material (Johnson *et al.*, 2004, Winder *et al.*, 2007). Previous studies have also proved its applicability to biological systems by successfully demonstrating the discrimination of bacteria to sub-species level (Winder *et al.*, 2006, Timmins *et al.*, 1998, Naumann *et al.*, 1991a) but only when combined with chemometrics (Goodacre *et al.*, 1998). In addition, the combination of FT-IR and trajectory analysis has proved successful in the identification of metabolic changes in natural multi-organism fermentations (Johnson *et al.*, 2004). FT-IR is particularly useful in gaining insight into complex biological problems as an initial screening method because it is high-throughput with analysis times of 10-60s per sample, requires minimal sample preparation, and is automated and relatively inexpensive.

In this investigation we aim to explore the phenotypic effects exerted on an activated sludge community when exposed to a range of environmental pollutants. The activated sludge community selected for investigation originated from an industrial steelworks where it was used in the degradation of phenol from coking effluents. The activated sludge was monitored for the phenotypic effects and potential degradation of phenol, the structurally similar aromatic hydrocarbon toluene, and the β-blocking agents Atenolol and Propranolol over a 48 h period using FT-IR spectroscopy and HPLC analysis. Through the comparison of the phenotypic changes with the degradation data, we aim to develop an understanding of the effects exerted on the activated sludge community by these environmental pollutants.

4.3 Materials and methods

4.3.1 Activated sludge

Activated sludge was collected from a bioreactor at an industrial wastewater treatment plant (Corus Steelworks, Scunthorpe, UK) and transported at room temperature immediately back to the laboratory where it was subsequently stored at 4°C until further use. This microbial community was stored for 1 day prior to supplementation with any of the chosen compounds.

4.3.2 Incubation conditions

Prior to the addition of the chosen pollutants the activated sludge samples (150 mL) were incubated aerobically for 2 h at 25 °C at 150 r.p.m. in a Multitron (INFORS HT, Switzerland) orbital shaker. Triplicate samples (so called biological replicates) were supplemented with the environmental pollutants 5 mM phenol, 1 mM toluene, 0.17 mM Propranolol, or 0.18 mM Atenolol, all of which have been shown to exert undesirable effects on aquatic organisms (*vide supra*). Additional triplicate control samples were incubated without the addition of any pollutant and were used to monitor the activity of the activated sludge during incubation under laboratory conditions. The activated sludge samples were maintained aerobically at 25 °C with 150 r.p.m. shaking throughout the incubation and sampling periods (Manefield *et al.*, 2002). Hourly samples were taken over a 10 h period (including t = 0 h and t = 15 min) and subsequently at 15, 20, 24, 36 and 48 h for analysis with FT-IR spectroscopy. An aliquot (1 mL) was centrifuged for 6 min at 16089 *g* and the supernatant was removed for quantitative analysis of the pollutant with HPLC or a colorimetric assay. The cell pellets were stored at -80 °C until required.

4.3.3 Phenol quantification

Quantification of phenol was performed using a colorimetric assay in which a red antipyrine dye is produced (King *et al.*, 1991). During this assay condensation of 4-aminoantipyrine with phenol in the presence of alkaline oxidising agents produces a red antipyrine dye (AAPPC) in what is known as the Emerson reaction (Emerson, 1943, Emerson and Kelly, 1948). It is known that this assay has a greater sensitivity than other colorimetric assays of phenol (Svobodová and Gasparič, 1971). The assay is most sensitive in the range of 0.002 – 0.02 mM. It was therefore necessary to dilute

the phenol samples from early time-point determinations with sterile H_2O in order to achieve the most sensitive assay. The colour develops within 2 min, and changes with time, but it does not significantly change within 1 h (Martin, 1949). The colorimetric assay is influenced by such factors as concentration of the reagents in relation to phenol concentration (Gasparic *et al.*, 1974, Svobodová and Gasparič, 1971) the pH at which the assay is performed (7.5 – 9.5) and the order in which the reagents are added. It is also considered necessary to have a least a 10-fold excess of 4-aminoantipyrine and the oxidising agent relative to phenol (Svobodová and Gasparič, 1971). For this reason, a blank was run with each series of determinations by mixing 4-aminoantipyrine, buffer, and the oxidising agent with sterile H_2O. The reagents and phenol standards were prepared just prior to their use.

The aliquots (1 mL) were filtered (0.22 µm, Millipore™) and treated with 25µL of 0.5 M NH_4OH and adjusted to pH 7.9 with 22.5 µL of phosphate buffer (0.5M KH_2PO_4, 0.6M K_2HPO_4). Samples were treated with 10 µL of 100mM 4-aminoantipyrine and 10 µL of 250 mM potassium ferricyanide. The samples were allowed to react for 5 min, and the reaction products were determined in triplicate against a blank for each replicate with a BioMate™ 5 (Thermo Electron Corporation) spectrophotometer at 500 nm. Phenol concentrations were calculated from standard curves prepared in parallel.

4.3.4 HPLC analysis

Concentrations of toluene, Atenolol and Propranolol were determined by HPLC (Agilent 1100 series). The supernatant samples were allowed to thaw at room temperature and were filtered (0.22 µm, Millipore™) in order to remove any microbial cells remaining in the medium. Aliquots (25 µL) were injected onto the HPLC column in a randomised order. Each sample was injected three times during the analysis, resulting in three analytical replicates for each biological sample. The HPLC system was equipped with a Chirobiotic V2 column (250 mm × 4.6 mm i.d.), particle size 5 µm (ASTEC, Whippany, NY, USA) and a UV detector operating at a wavelength of 230 nm. The column was eluted with an isocratic mixture of methanol and water (90:10, v/v) and 1.0% triethylamine acetate (TEAA) buffer, pH 5.0. The pH of the buffer was adjusted with acetic acid prior to the addition of methanol. The measurements were carried out a 25 ± 1 °C at a flow rate of 1 mL min^{-1} (Bosakova *et al.*, 2005).

4.3.5 FT-IR spectroscopy

A 96-well zinc selenide plate was rinsed with analytical grade 2-propanol and deionised water and allowed to dry at room temperature (Winder *et al.*, 2006). The dry cell pellets stored at -80 ℃ were allowed to thaw at room temperature and washed in order to remove any traces of phenol. Ice cold sterile water (2 mL) was added to each sample and vortexed. The samples were centrifuged for 10 mins (0 ˚C, 16089 *g*), and the supernatant was discarded. This cycle was repeated 3 times. A final aliquot of 100 µL sterile water was added to each sample and the solution was vortexed. Aliquots (20 µL) of each resuspended sample were applied to the ZnSe plate and oven dried at 50 ℃ for 10 min. Drying was used to avoid absorption of water in the mid-IR region which may mask biologically important chemical information in the spectra. Three replicates of each of the samples were randomly applied to the ZnSe plates and triplicate spectra were obtained from different positions of each well, a total of nine spectra (so called technical replicates) were collected per sample. The plate was loaded onto a motorised microplate module HTS-XT under the control of a computer programme with OPUS software version 4 (Winder *et al.*, 2006). Spectra were collected using an Equinox 55 FT-IR spectrometer (Bruker Optics Ltd), in transmission mode using the deuterium triglycine sulphate (DTGS) detector over the wavelength range of 4000-600 cm^{-1} and with a resolution of 4 cm^{-1}. In order to improve the signal to noise ratio, 64 spectra were co-added and averaged. The spectra are displayed in terms of absorbance (see Figure 4.2 for typical example spectra).

4.3.6 Data analysis

The ASCII data were imported into Matlab version 7.1 (The MathWorks, Inc., Natick, MA, USA) and as an initial step CO_2 vibrations arising from the atmosphere (2403 – 2272 cm^{-1} and 683 – 656 cm^{-1}) were removed and filled with a linear trend. The spectra were then corrected using extended multiplicative scatter correction (EMSC) which normalises and smoothes the spectra by application of a polynomial smoothing function (Martens *et al.*, 2003). These pre-processed spectra were used for subsequent multivariate analyses.

Principal components analysis (PCA; (Jolliffe, 1986)) is an unsupervised method for reducing the dimensionality of multivariate data whilst preserving the variance. This transformation was performed prior to canonical variates analysis (CVA). CVA is a supervised learning method that seeks to minimise within-group variance whilst

maximising between-group variance, and can be used in conjunction with PCA to discriminate between groups on the basis of retained principal components (PCs), given *a priori* knowledge of group membership of the spectral replicates (Manly, 1994, Winder *et al*., 2004). PC-CVA models were constructed with *a priori* knowledge of the biological replicates. In order to make sure these PC-CVA score plots were not over- or under-trained, these models were trained (constructed) with two of the biological replicates and cross-validated using the third biological replicate as detailed in (Jarvis and Goodacre, 2004a). Finally, CVA also allows statistical significance to be displayed on the scores plots and circles were used to represent the 95% χ^2 tolerance region constructed around each group mean by the χ^2 distribution of two degrees of freedom (Krzanowski, 1988).

4.4 Results and discussion

4.4.1 Quantification of the environmental pollutants

The concentration of the environmental pollutants remaining in the activated sludge samples was monitored during the 48 h incubation period. A colorimetric assay was used to determine levels of phenol in the supernatant for all of the experimental conditions. Figure 4.1 shows the quantification results for the phenol supplemented and control activated sludge samples. The activated sludge samples contained residual levels of approximately 0.22 mM phenol which was gradually degraded by the microbial community within 48 h. Activated sludge samples supplemented with 5 mM phenol displayed a notable degradation of the phenol within 2 h and complete degradation was achieved within 36 h.

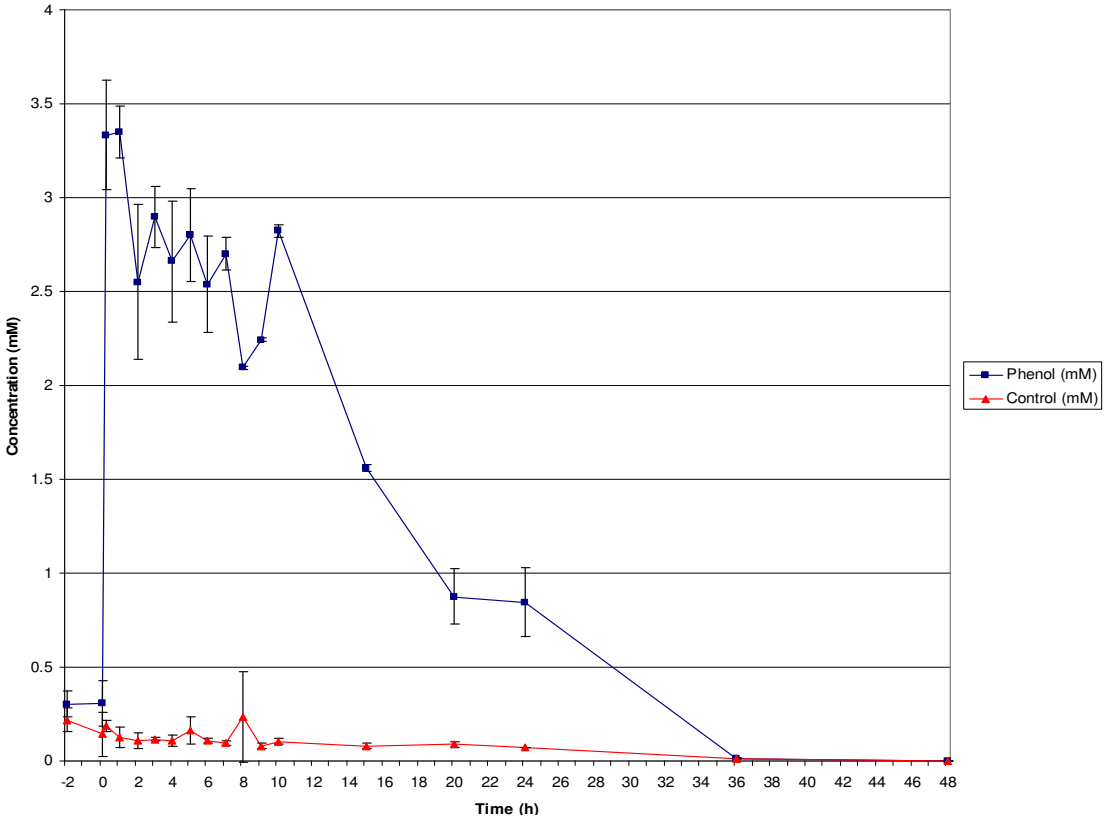


Figure 4.1: Quantification of phenol remaining in the supernatant of phenol supplemented and control samples during the 48 h incubation period. Data are averaged from 3 biological replicates and error bars show standard deviation.

Quantification of phenol in the activated sludge samples supplemented with toluene, Atenolol and Propranolol corresponded with the control data; containing a residual

concentration of approximately 0.22 mM phenol remaining in the activated sludge following collection from the industrial steelworks. This residual phenol was fully degraded by the activated sludge community within the 48 h incubation period and these samples were not supplemented with additional phenol. Unfortunately, due to the complex nature of the activated sludge samples, it was not possible to quantify the levels of the other environmental pollutants with HPLC. The samples produced high levels of background noise and as a result it was not possible to resolve the compounds using this method.

4.4.2 Effects of the environmental pollutants on the FT-IR metabolic fingerprints

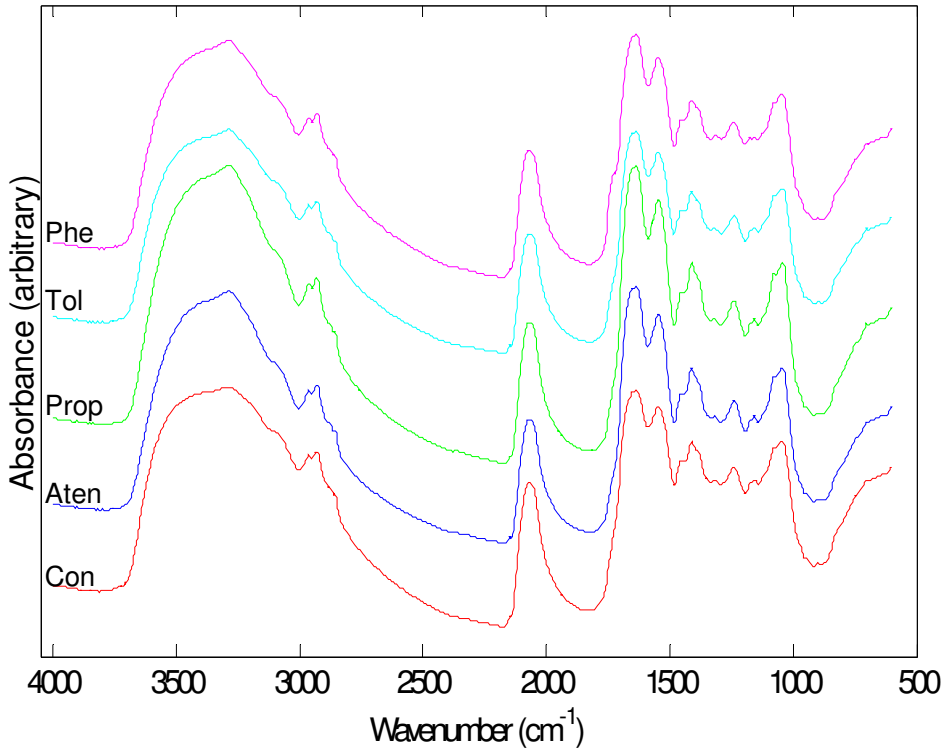


Figure 4.2: Average FT-IR spectra for activated sludge samples following 48 h incubation at 25 °C in the presence of the phenol, toluene, Propranolol and Atenolol. Samples are labelled Phe, Tol, Prop and Aten for samples exposed to phenol, toluene, Propranolol and Atenolol respectively. Control samples are labelled "Con". The spectra are offset on the Y axis to allow easier visualisation, and the absorbance is relative.

Typical FT-IR spectra of the activated sludge samples incubated with phenol, toluene, Propranolol or Atenolol and samples incubated without any supplementary compounds are shown in Figure 4.2. FT-IR spectroscopy results in the production of a composite spectrum that is representative of a wide range of endogenous and exogenous

metabolites. The resultant spectrum is therefore highly complex and not open to direct interpretation of individual compounds of interest (See section 1.4.1, page 33). On initial inspection a dominant peak was observed in all of the spectra at wavenumbers 2120 - 2000 cm^{-1}, which is not typically observed in the FT-IR spectra of bacterial samples (Winder *et al.*, 2006), and there are very few functional groups that vibrate in this area of the infrared spectrum. The most common are S-O from sulphates, C≡N from unsaturated N compounds, C-C stretching vibrations of alkynes and S=C=N from thiocyanates (Degen, 1997, Peoples *et al.*, 1987). It is known however, that phenol and thiocyanates are the two primary pollutants in coking effluent (Jeong and Chung, 2006, Prater and Fisher, 1991, Manefield *et al.*, 2005) and furthermore the effluent from the industrial plant in which the samples originate is known to contain thiocyanates at concentrations as high as 2.4 mM (Manefield *et al.*, 2005). It is therefore likely that this peak was attributable to the presence of thiocyanates within the activated sludge samples. The remaining spectral regions typify the patterns commonly associated with the FT-IR spectra of microorganisms. Many of the spectral bands have previously been assigned to functional groups of biological interest (Fabian and Mäntele, 2002, Lasch *et al.*, 2002, Schultz *et al.*, 1999, Winder *et al.*, 2006). The spectra for all of the experimental conditions appear to be dominated by the amide I and amide II functional group bands at wavenumbers 1700 – 1600 cm^{-1} and 1600 – 1500 cm^{-1} respectively. The band at 1200 – 900 cm^{-1} is attributed to C-O or O-H from polysaccharides. In addition, the absorptions at wavenumbers 3000 – 2800 cm^{-1} are attributed to lipids; in particular, the symmetric and asymmetric stretching vibrations of the CH, CH_2 and CH_3 groups of the acyl chains (Lorin-Latxague and Melin, 2005, Udelhoven *et al.*, 2000, Patel *et al.*, 2008).

Due to the complex nature of the FT-IR dataset multivariate analysis (MVA) methods were employed to investigate the relationship between the bacterial communities from their FT-IR data. PC-CVA was employed to investigate the phenotypic effect of the compounds on the activated sludge during incubation at 25 °C over a 48 h period. Due to the large number of groups within the FT-IR dataset (5 experimental conditions x 17 time points x 3 biological replicates x 9 analytical replicates) and in order to aid visualisation of the general trends within the data, a subset of samples were selected for multivariate analysis with PC-CVA. Figure 4.3 shows the PC-CVA scores and loadings plots for the FT-IR data of the activated sludge communities supplemented with phenol, toluene, Atenolol or Propranolol and those not supplemented with any additional compounds (i.e., the control samples) at 0, 4, 8, 15, 20, 24 and 48 h. The 0

h samples were collected just prior to the addition of any compounds for all of the experimental conditions and therefore as expected, all of the samples for this time point clustered with the control. A clear separation was observed across the first canonical variate (PC-CV 1) which correlated to the time of sampling. The activated sludge samples supplemented with phenol displayed a distinct separation from the remaining samples across PC-CV 2 following the 15 h time point. This separation demonstrates a clear phenotypic difference in the activated sludge samples when supplemented with phenol; a finding which was not perhaps surprising, as the activated sludge community selected for investigation was employed for the degradation of phenols in the coking effluent (Manefield *et al.*, 2005). It is probable that the observed change in the phenotype of the sludge was directly related to the time at which the community was actively degrading phenol, and this will be investigated further. By contrast, samples supplemented with toluene, Atenolol and Propranolol displayed no notable separation from the control samples, as the 95% χ^2 tolerance regions around the group means overlap. Visual inspection of the PC-CVA loadings plot for this analysis (Figure 4.3b) provides information relating to which spectral features are important for this separation. In these analyses the 95% χ^2 confidence band is represented as a green ellipse, such that any spectral features which are positioned outside of this region are considered significant in the PC-CV loadings plot. The separation across PC-CV 1 appeared to be due to variation in the FT-IR spectra at wavenumbers 2125 – 2055 cm^{-1}. As described above this peak was likely to be due to the presence of thiocyanates within the samples. A number of studies have investigated the degradation of thiocyanates in coking effluent by activated sludge communities (Paruchuri *et al.*, 1990, Shieh and Richards, 1988, Hung and Pavlostathis, 1997, du Plessis *et al.*, 2001, Staib and Lant, 2007). Thiocyanates are employed as a growth substrate by the activated sludge and whilst it is known that phenol does not affect this process, the presence of cyanides within the effluent have been shown to significantly inhibit thiocyanate degradation (Staib and Lant, 2007). Furthermore the degradation of thiocyanates within activated sludge communities has been reported to be much slower than that of phenol and cyanide (Staib and Lant, 2007). Thus it is likely that the changes in this peak of the FT-IR spectra for the control samples may be due to the degradation of thiocyanates by the activated sludge during the incubation period.

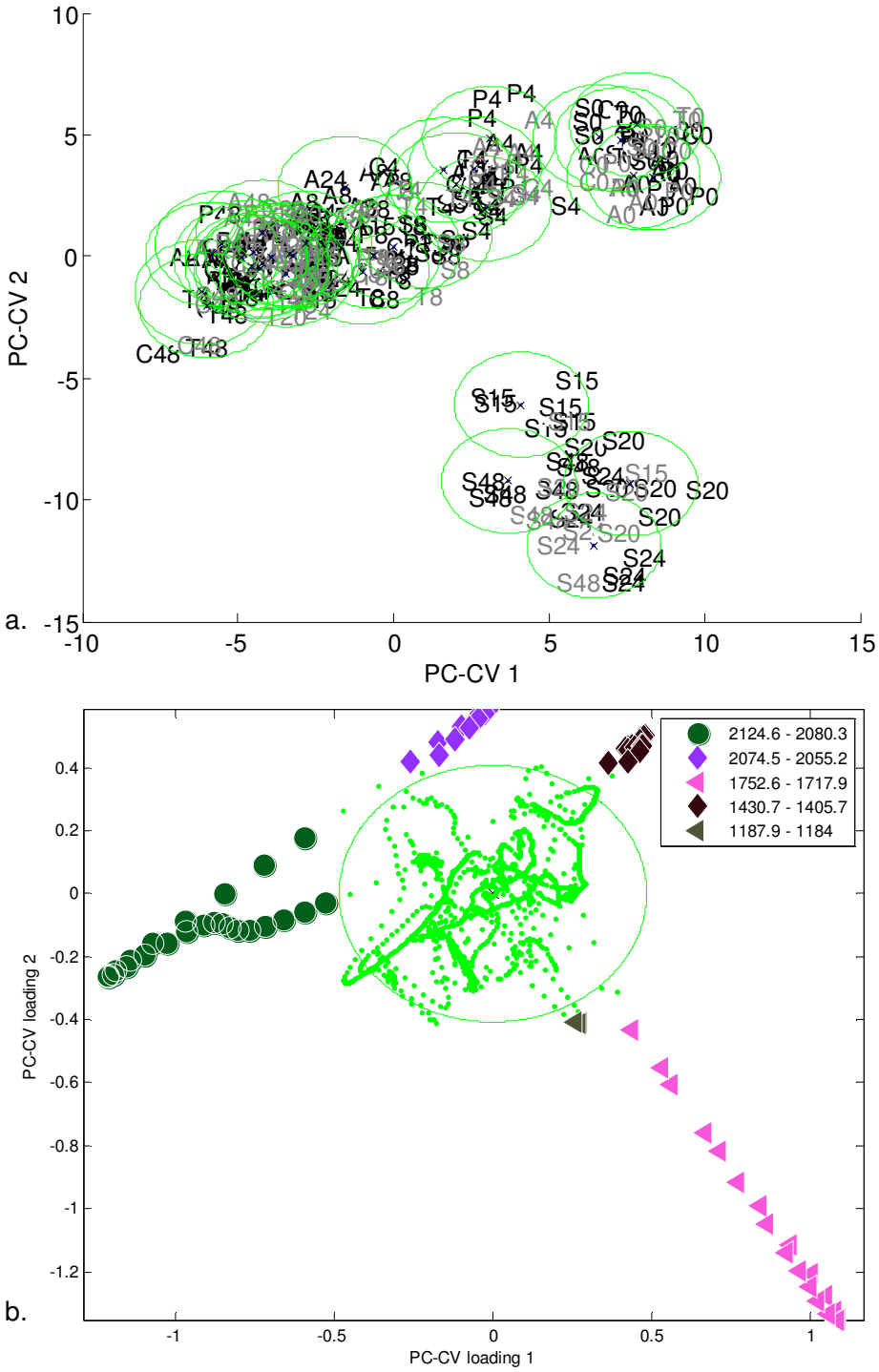


Figure 4.3: PC-CVA scores (a.) and loadings (b.) plots for FT-IR spectra of activated sludge samples supplemented with phenol, toluene, Atenolol and Propranolol for 0, 4, 8, 15, 24 and 48 h at 25 ℃. Samples supplemented with phenol, toluene, Atenolol and Propranolol are labelled "S", "T", "A" and "P" respectively. Control samples are represented by "C". PCs 1-8 were employed by the CVA algorithm with the *a priori* knowledge of the biological replicates and accounts for 98.9% explained variance. In the scores plot points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Black circles represent the 95% confidence interval about the group centroid, and grey circles the 95% confidence region about the group sample population. For the loadings biplot (b.), contiguous spectral regions falling beyond 2 standard deviations from the mean are encoded by colours and symbols detailed in the legend.

The greatest variance observed in the PC-CVA of all of the experimental conditions was between the samples with / without supplementary phenol. Therefore, the data for each of the experimental conditions was analysed individually with the control samples in order to investigate any more subtle phenotypic changes in these communities when exposed to the chosen compounds. As described above, a subset of samples was selected for analysis in order to allow visualisation of the strongest trends within the dataset. The full datasets were also analysed (data not shown) to ensure that the trends observed in the selected samples represent the trends observed in the entire dataset. Figure 4.4a and Figure 4.4b show the PC-CVA scores and loadings plots for the activated sludge samples at 0, 24 and 48 h following the addition of phenol, and the control samples. As observed in the analysis of the whole dataset (Figure 4.3a) the samples for the 0 h time point clustered together and those supplemented with phenol at 24 and 48 h show a distinct separation from the control samples across PC-CV 1. The PC-CVA loadings plot (Figure 4.4b) demonstrates that this separation was due to changes in the FT-IR spectra at wavenumbers 1755 – 1718 cm^{-1} which is most commonly associated with the C=O stretching vibrations of esters, aldehydes or ketones. In addition the separation across PC-CV 1 was also shown to be due to changes at wavenumbers 1199 – 1178 cm^{-1}, most often associated with the C-N stretching vibrations of amines and the C-O stretching vibrations of alcohols, ethers and esters. Some separation of the 24 and 48 h phenol supplemented samples is observed across PC-CV 1, however, the 95% confidence regions overlap and therefore the separation is not considered significant for this analysis. The control communities display separation across PC-CV 2 according to the time of sampling. Inspection of the PC-CVA loadings plot for this analysis indicated that this separation was due to variation at wavenumbers 2123 – 2047 cm^{-1} corresponding with the analysis carried out on the whole dataset (Figure 4.3b). The observed separation of the control samples according to the time of sampling indicates a phenotypic change within the activated sludge community during the incubation period.

The PC-CVA scores and loadings plots for activated sludge samples for 48 h with and without the addition of toluene are shown in Figure 4.4c and Figure 4.4d respectively. As described above, the 0 h samples displayed no separation in this analysis, whilst a clear separation between the 24 and 48 h time points from the 0 h samples was observed across PC-CV 1. However, no notable difference was observed between the control and toluene supplemented samples indicating that there was no phenotypic difference in the toluene supplemented and non-supplemented sludge samples,

despite toluene's somewhat close structural similarity to phenol. The loadings plot for this analysis demonstrates that the separation of samples across PC-CV 1 was due to changes at wavenumbers 2121 – 2084 cm^{-1}, 1036 – 1018 cm^{-1} and 1651 – 1618 cm^{-1} which are associated with the S=C=N from thiocyanates, the C-O stretching vibrations of alcohols and the C=C stretch of aromatic rings respectively. The presence of aromatic rings within the activated sludge samples was not unexpected as the chemical structures of phenol (residual levels of phenol in the activated sludge samples) and toluene both contain an aromatic ring (Figure 4.5). Unfortunately, due to the high levels of background noise from the supernatant it was not possible to use HPLC to quantify levels of toluene remaining in the medium following the incubation period. However, the phenotypic response appears dramatically different in samples actively degrading phenol and that of the activated sludge community supplemented with toluene appears to mimic the control communities suggesting that it did not exert a great phenotypic response in the activated sludge community. This was rather surprising as toluene is very similar in structure to phenol (Figure 4.5). In addition it is known that a number of microorganisms are capable of degrading both toluene and phenol, however; there have been no reports on the ability of the activated sludge used in this investigation to degrade toluene (Kahng *et al.*, 2001, Kukor and Olsen, 1990, Nelson *et al.*, 1987). The first step in the degradation pathways for both of these compounds is a monooxygenase and it has been shown that the monooxygenase employed in the degradation of toluene has close homology with that of the phenol degradation pathway (Shingler *et al.*, 1992, Yen *et al.*, 1991, Nordlund *et al.*, 1990). A number of the same enzymes are employed for the degradation of both of these compounds in *P. putida* F1 (Reardon *et al.*, 2000, Spain *et al.*, 1989). Furthermore, Rogers and Reardon demonstrated the simultaneous degradation of toluene and phenol showing that the same pathway was utilised in the degradation of both of the compounds in *Burkholderia* sp. (Rogers and Reardon, 2000). The activated sludge was exposed to a relatively low concentration of toluene (1mM) in comparison with the concentration of phenol (5mM) to which it was exposed. It is therefore possible that some degradation of toluene did occur within the activated sludge community that may have not been detectable with FT-IR spectroscopy due to the low levels of toluene with which is was exposed. However, it is not possible to investigate this further without the relevant toluene quantification data.

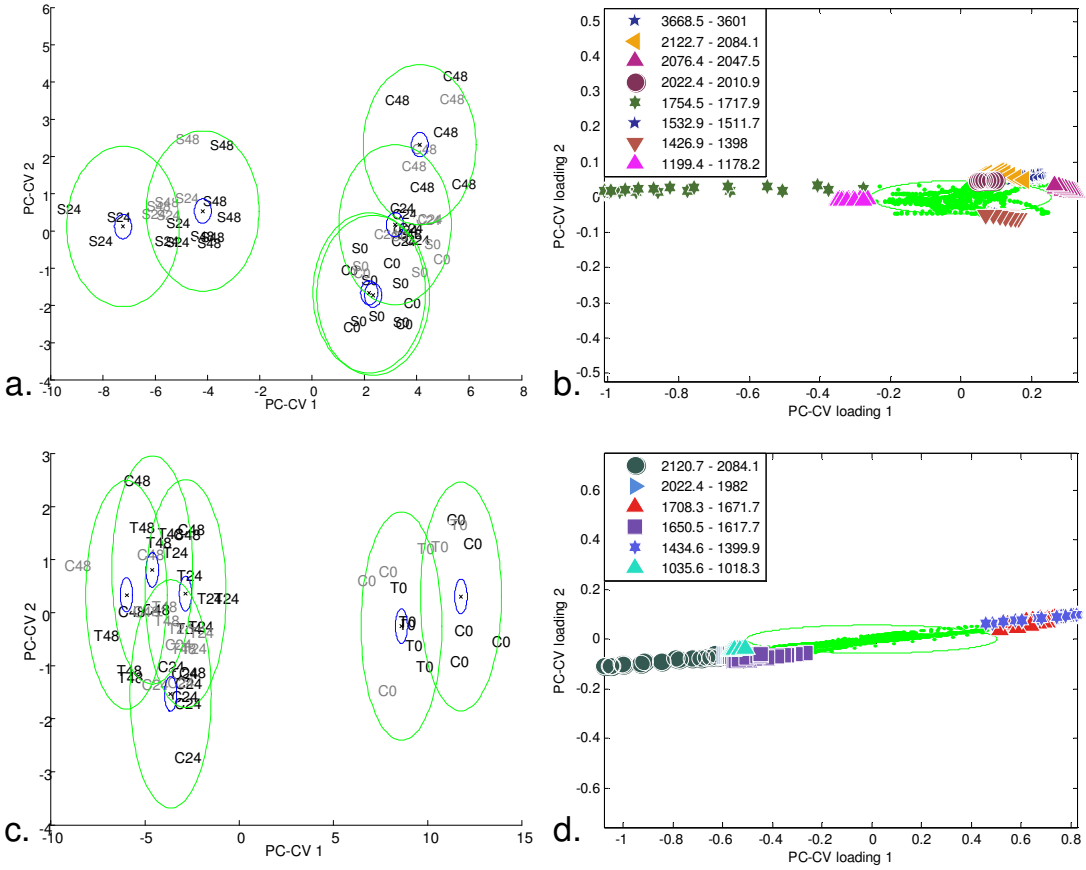
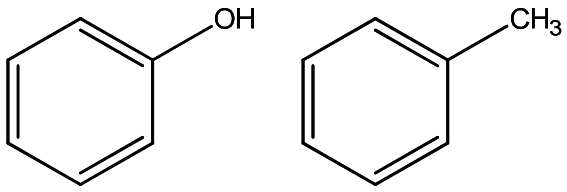


Figure 4.4: PC-CVA scores (LHS) and loadings (RHS) plots for FT-IR spectra of activated sludge samples supplemented with phenol (a and b) and toluene (c and d) for 0, 24 and 48 h at 25 °C. Samples supplemented with phenol and toluene are labelled "S" and "T" respectively. Control samples are represented by "C". PCs 1-5 were employed by the CVA algorithm with the *a priori* knowledge of the biological replicates and account for 97.7% and 98% explained variance respectively. In the scores plots points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Black circles represent the 95% confidence interval about the group centroid, and grey circles the 95% confidence region about the group sample population. For the loadings biplots (RHS), contiguous spectral regions falling beyond 2 standard deviations from the mean are encoded by colours and symbols detailed in the legend.



Phenol Toluene

Figure 4.5: The chemical structures of the aromatic hydrocarbons phenol and toluene.

PC-CVA was also carried out on the FT-IR spectra of activated sludge samples incubated in the presence of the APIs Atenolol and Propranolol. Figure 4.6 shows the PC-CVA scores and loadings plots for this analysis. As noted for the toluene 'supplementation' no notable difference was observed in samples supplemented with Atenolol from the control samples (Figure 4.6a). As described above, the samples separate across PC-CV 1 according to the time point in the incubation period. However, the samples supplemented with Atenolol co-cluster with the control samples, indicating that the API did not exert a phenotypic effect on the activated sludge community. In addition, Propranolol did not appear to exert an effect on the phenotypic fingerprint of the activated sludge community (Figure 4.6c). A separation of the 0 h samples from the 24 and 48 h samples was observed across PC-CV 1, but there was no observable separation between the control and API supplemented samples for both Atenolol and Propranolol. The PC-CVA loadings for these analyses (Figure 4.6b and Figure 4.6d) show similarities in the wavenumbers responsible for the separation observed in both of the PC-CVA scores plots. For example the majority of the separation observed across PC-CV 1 for samples supplemented with Atenolol was due to wavenumbers 2115 – 2080 cm^{-1}. The differences observed between control samples and those supplemented with Propranolol is due to changes in the FT-IR spectra at wavenumbers 2124 – 2078 cm^{-1} and as described above, this falls within the region which is associated with the S=C=N vibrations from thiocyanates (Peoples *et al.*, 1987). Again it was not possible to quantify the concentration of Atenolol and Propranolol remaining in the supernatant following the 48 h time course. This was due to the complex nature of the samples which masked the presence of the APIs in the HPLC analysis. However, since the activated sludge samples supplemented with Atenolol or Propranolol did not display any phenotypic variation from the control samples in the analysis of the FT-IR data it was unlikely that any degradation of the APIs occurred.

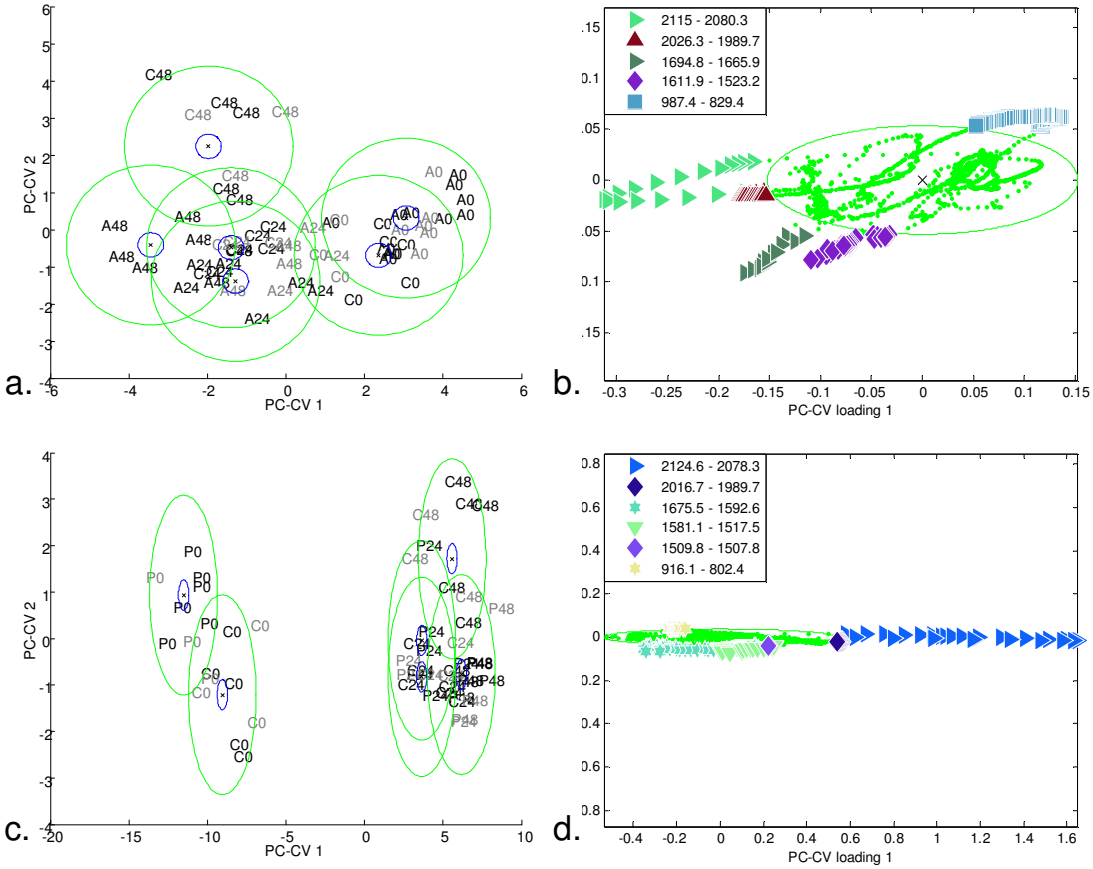


Figure 4.6: PC-CVA scores (LHS) and loadings (RHS) plots for FT-IR spectra of activated sludge samples supplemented with Atenolol (a and b) and Propranolol (c and d) for 0, 24 and 48 h at 25 ℃. Samples supplemented with Atenolol and Propranolol are labelled "A" and "P" respectively. Control samples are represented by "C". PCs 1-5 were employed by the CVA algorithm with the *a priori* knowledge of the biological replicates and account for 97.5% and 98.4% explained variance respectively. In the scores plots points shown in black represent the 2 biological replicates used to train the PC-CVA models. Points shown in grey represent the third biological replicate which was used to validate the PC-CVA model. Black circles represent the 95% confidence interval about the group centroid, and grey circles the 95% confidence region about the group sample population. For the loadings biplots (RHS), contiguous spectral regions falling beyond 2 standard deviations from the mean are encoded by colours and symbols detailed in the legend.

4.5 Conclusions

In this investigation activated sludge was acquired from an industrial bioreactor capable of degrading phenol. The sludge was incubated in the laboratory and monitored for the phenotypic effects exerted on the sludge by a range of environmental pollutants over a 48 h period. Multiple samples were taken across the time course and analysed by FT-IR spectroscopy. FT-IR was employed as a whole organism fingerprinting approach to monitor biochemical changes in the bacterial cells during exposure to the chosen compounds. We also investigated the ability of the activated sludge to degrade these compounds during the time course. The phenol quantification data obtained from samples supplemented with 5 mM phenol showed that the activated sludge community was capable of degrading 5 mM phenol in 36 h, and it was evident from the FT-IR data that a metabolic effect was exerted on the microbial cells. However, no notable phenotypic effect was observed in the microbial community when exposed to toluene, Atenolol or Propranolol and this was likely to be because this complex microbial community had been selected to degrade phenol rather than its close analogue toluene or the two structurally diverse APIs.

Phenotypic differences were observed in the control cells during the time course and we hypothesise that this effect may be attributed to the presence of thiocyanates. However, further investigation is required to confirm these observations. In conclusion, high-throughput metabolic fingerprinting with FT-IR spectroscopy has been used to monitor phenotypic shifts within complex microbial communities when exposed to a range of environmental pollutants. This investigation demonstrates that FT-IR spectroscopy when combined with chemometric analysis is a very useful high-throughput screening approach for assessing the metabolic capability of complex microbial communities.

5 Monitoring the phenotypic changes in complex bacterial communities capable of degrading phenol

5.1 Abstract

The coking process produces great volumes of wastewater contaminated with pollutants such as cyanides, sulphides and phenolics. Chemical and physical remediation of this wastewater removes the majority of these pollutants; however, these processes do not remove phenol and thiocyanate. The removal of these compounds has been effected during bioremediation with activated sludge containing a complex microbial community. In this investigation we acquired activated sludge from an industrial bioreactor capable of degrading phenol. The sludge was incubated in our laboratory and monitored for its ability to degrade phenol over a 48 h period. Multiple samples were taken across the time course and analysed by Fourier transform infrared (FT-IR) spectroscopy. FT-IR spectroscopy was employed as a whole organism fingerprinting approach to monitor biochemical changes in the bacterial cells during the degradation of phenol. We also investigated the ability of the activated sludge to degrade phenol following extended periods (2-131 days) of storage in the absence of phenol. A reduction was observed in the ability of the microbial community to degrade phenol and this was accompanied by a detectable biochemical change in the FT-IR fingerprint related to cellular phenotype of the microbial community. In the absence of phenol a decrease in thiocyanate vibrations was observed, reflecting the ability of these communities to degrade this substrate. Actively degrading communities showed an additional new band in their FT-IR spectra that could be attributed to phenol degradation products from the ortho- and meta-cleavage of the aromatic ring. We believe that FT-IR spectroscopy when combined with chemometric analysis is a very useful high-throughput screening approach for assessing the metabolic capability of complex microbial communities.

5.2 Introduction

Phenolic compounds are ubiquitous in the environment due to their release from both industrial and natural processes. The accumulation of phenolic compounds comprises an environmental hazard, and the removal of these pollutants is required prior to the release of wastewater from such industrial processes (Environment Agency, 1995, Whiteley and Bailey, 2000). The coking process produces great volumes (approximately 1,800 L min^{-1}) of wastewater contaminated with pollutants such as cyanides, sulphides, thiocyanates and phenolics (Patterson, 1985, Manefield *et al.*, 2002, Neufeld and Valiknac, 1979, Philp *et al.*, 2003, Vázquez *et al.*, 2006). Chemical and physical remediation of this wastewater removes the majority of these pollutants (Vázquez *et al.*, 2007, Ahmaruzzaman and Sharma, 2005, Minhalma and de Pinho, 2002, Minhalma and de Pinho, 2004, Qian *et al.*, 1994, Vázquez *et al.*, 2006). However, these processes do not remove phenol and thiocyanate, and the removal of these compounds takes place during bioremediation with activated sludge (Manefield *et al.*, 2005).

The use of microorganisms in the bioremediation of phenols produced during industrial processes is widespread (Sutton *et al.*, 1999, Tisler *et al.*, 1999). However, until the development of RNA stable isotope probing (RNA-SIP), the composition of the microbial communities specifically degrading phenolics within activated sludge was largely speculative and based upon the presence of phenolic degradation pathways within isolates cultured axenically from the system under study (Whiteley and Bailey, 2000). In order to target the key degraders directly, Manefield *et al.* (2002) applied RNA-SIP on the activated sludge utilised in this investigation, to identify the dominant members of this community responsible for the degradation of phenol. Bacterial genera found to dominate the acquisition of carbon from phenol, through cellular isotopic labelling after ^{13}C labelled phenol degradation, included members of the β-Proteobacteria, including the *Thauera* and *Acidovorax* genera, and the α-proteobacterial *Rhodopseudomonas* genus. All of the phenol degrading microorganisms highlighted by this study have previously been isolated when aromatic compounds such as phenol, benzoate, or toluene were used as the sole source of carbon (Tschech and Fuchs, 1987, Harwood and Gibson, 1988). However, whilst members of the *Thiobacillus* and *Acidobacterium* genera were also found to be dominant within this activated sludge community, these species do not assimilate carbon from phenol (Manefield *et al.*, 2002).

Whilst there has been considerable interest in characterising the members of the complex microbial community involved in biodegradation, little attention has been directed towards the metabolic pathways utilised in the degradation of phenol. In this investigation we have employed Fourier transform infrared (FT-IR) spectroscopy as a phenotypic typing technique which has previously been used to generate informative metabolic fingerprints from biological material (Johnson *et al.*, 2004, Winder *et al.*, 2006). Previous studies have also proved its applicability to biological systems by successfully demonstrating the discrimination of bacteria to sub-species level (Winder *et al.*, 2006, Timmins *et al.*, 1998, Naumann *et al.*, 1991a) but only when combined with chemometrics (Goodacre et al., 1998). In addition, the combination of FT-IR and trajectory analysis has proved successful in the identification of metabolic changes in natural multi-organism fermentations (Johnson *et al.*, 2004). FT-IR spectroscopy is particularly useful in gaining insight into complex biological problems as an initial screening method because it is high-throughput with analysis times of 10-60s per sample, requires minimal sample preparation, is automated and relatively inexpensive.

In this study three objectives were investigated to provide insight into the process of phenol degradation by an activated sludge community previously developed for the degradation of phenol in coking effluent. In the first instance, biochemical changes in the activated sludge community were monitored during the degradation of phenol over a 48 h period through the use of FT-IR spectroscopy. In addition we determined the effect of storing the activated sludge for extended periods of time, (without added phenol) to understand the ability of this microbial community to degrade phenol and investigated the phenotypic changes by FT-IR analysis. In addition, GC-MS was employed for metabolite profiling of the activated sludge community in order to monitor the changes in the meta-metabolome during the active degradation of phenol.

5.3 Materials and methods

5.3.1 Sample Information

The overall workflow for this series of experiments is detailed in Figure 5.1. Activated sludge was obtained from a bioremediation system at an industrial steelworks where it was employed in the degradation of phenol in coking effluent (Manefield *et al.*, 2005, Manefield *et al.*, 2002, Whiteley and Bailey, 2000). As described above, coking effluents also typically contain pollutants such as sulphides, cyanides and thiocyanates. The activated sludge was collected and transported immediately back to the laboratory (Manchester University) at room temperature where it was stored at 4°C until further use. Activated sludge was stored for 2, 9, 16 and 131 days prior to any supplementation with phenol. This ageing of the samples was carried out in order to assess the phenol degrading capabilities of the activated sludge following extended storage periods, during which no additional phenol was added to the samples.

5.3.2 Laboratory Microcosms

Prior to the addition of phenol the activated sludge samples (150 mL) were incubated aerobically for 2 h at 25 °C at 150 r.p.m. in a Multitron (INFORS HT, Switzerland) orbital shaker. Triplicate samples (so called biological replicates) were supplemented with 5 mM phenol as the sole carbon source. Additional triplicate control samples were incubated without added phenol, and were used to monitor the activity of the activated sludge during incubation under laboratory conditions. All of the activated sludge samples were maintained aerobically at 25 °C with 150 r.p.m. shaking throughout the incubation and sampling periods (Manefield *et al.*, 2002). Hourly samples were taken over a 10 h period (including t = 0 h and t = 15 min) and subsequently at 15, 20, 24, 36 and 48 h for the quantification of phenol and analysis with FT-IR spectroscopy. For FT-IR spectroscopy a 1 mL aliquot was centrifuged for 6 min at 16089 *g*, the supernatant was removed and the cell pellets were stored at -80 °C until required. Cell pellets were stored so that all of the activated sludge samples may be analysed by FT-IR spectroscopy at the same time, in order to reduce any minor instrument variation within the dataset.

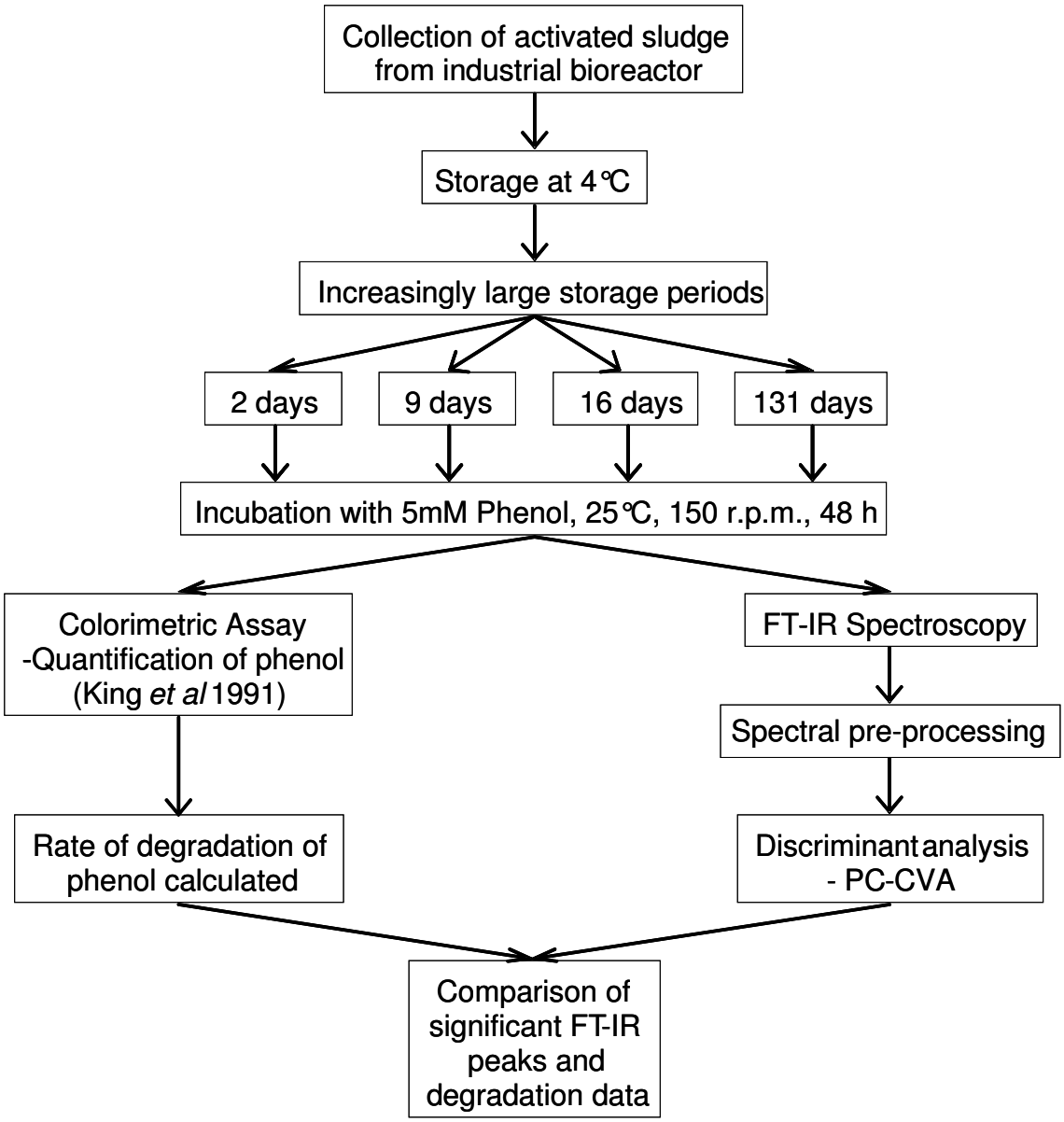


Figure 5.1: Experimental workflow showing the different steps between collection of the activated sludge samples, incubation with phenol and phenol quantification and FT-IR spectroscopy.

5.3.3 Assay for Phenol Quantification

The phenol concentration at each sampling point was determined by the Emerson reaction; a colorimetric assay in which a red coloured antipyrine dye (AAPPC) is produced through the condensation of 4-aminoantipyrine with phenol in the presence of alkaline oxidising agents (King *et al.*, 1991). The assay has a greater sensitivity than other colorimetric assays of phenol (Svobodová and Gasparič, 1971) and is most sensitive in the range of 0.002 – 0.02 mM. The colour develops within 2 min, and changes with time, but it does not significantly change within 1 h (Martin, 1949, Gasparic *et al.*, 1974, Svobodová and Gasparič, 1971). The reagents and phenol standards were prepared just prior to their use.

Sample supernatants generated by centrifugation as detailed above (1 mL) were filtered (0.22 µm, Millipore^{TM}) and treated with 25 µL of 0.5 M NH_4OH and adjusted to pH 7.9 with 22.5 µL of phosphate buffer (0.5 M KH_2PO_4, 0.6 M K_2HPO_4). Samples were treated with 10 µL of 100 mM 4-aminoantipyrine and 10 µL of 250 mM potassium ferricyanide. The samples were allowed to react for 5 min, and the reaction products were determined in triplicate against a blank (4-aminoantipyrine, buffer, and the oxidising agent was mixed with sterile H_2O) for each replicate with a BioMate™ 5 (Thermo Electron Corporation) spectrophotometer at 500 nm. Phenol concentrations were calculated from standard curves prepared in parallel. The samples were diluted with sterile deionised water when not in the sensitive range of the assay (i.e. 0.002 – 0.02 mM).

5.3.4 Analysis of microbial cells by FT-IR spectroscopy

Fingerprint analysis using FT-IR spectroscopy (Maquelin *et al.*, 2002) was performed on all of the activated sludge samples, taken from each of the four time delayed experiments at the same time in order to reduce any minor instrument variation within the dataset. FT-IR spectroscopy is a whole organism fingerprint analysis approach and analyses the total complement of the (bio)chemical components from bacterial cells (Goodacre *et al.*, 1998, Magee, 1993); this includes lipids, sugars, nucleic acids as well as proteins (Helm and Naumann, 1995). A 96-well zinc selenide plate was rinsed with 2-propanol and deionised water and allowed to dry at room temperature (Winder *et al.*, 2006). The dry cell pellets stored at -80 °C were allowed to thaw at room temperature and washed in order to remove any traces of phenol. Ice cold sterile water (2 mL) was

added to each sample and vortexed. The samples were centrifuged for 10 mins (0 °C, 16089 *g*), and the supernatant was discarded. This cycle was repeated 3 times. A final aliquot of 100 µL sterile water was added to each sample and the solution was vortexed. Aliquots (20 µL) of each resuspended sample were applied to the ZnSe plate and oven dried at 50 °C for 10 min. Drying was used to avoid absorption of water in the mid-IR region which may mask biologically important chemical information in the spectra. Three analytical replicates of each of the samples were randomly applied to the ZnSe plates and triplicate spectra were obtained from different positions within each well, a total of nine spectra (so called analytical replicates) were collected per biological sample. For subsequent multivariate models one biological sample (i.e. biological replicate) is represented by nine FT-IR spectra. The plate was loaded onto a motorised microplate module HTS-XT under the control of a computer programme with OPUS software version 4 (Winder *et al.*, 2006). Spectra were collected using an Equinox 55 FT-IR spectrometer (Bruker Optics Ltd), in transmission mode using the deuterated triglycine sulphate (DTGS) detector over the wavelength range of 4000-600 cm^{-1} and with a resolution of 4 cm^{-1}. In order to improve the signal to noise ratio, 64 spectra were co-added and averaged. The spectra are displayed in terms of absorbance.

5.3.5 Metabolite profiling with GC-MS

In order to quench metabolism, activated sludge samples (2 mL) were plunged rapidly into an equal volume of 60% aqueous methanol solution (–48 °C). The quenched biomass was centrifuged for 10 min (3,000 *g*, –9 °C) and the supernatant was removed. The biomass pellets were stored at –80 °C for further analysis (Winder *et al.*, 2008).

For metabolite extraction, the biomass pellets were suspended in 550 µL of 80% aqueous methanol solution (–48 °C), frozen in liquid nitrogen, and allowed to thaw on dry ice. The freeze thaw cycle was carried out three times in order to permeabilise the cells and allow leakage of the metabolites into the supernatant. The suspensions were centrifuged for 7 min (16,060*g*, –9 °C) and the supernatant was retained and stored on dry ice. An additional aliquot (550 µL) of 80% aqueous methanol solution was added to the biomass pellet and the procedure was repeated. The second aliquot of the supernatant was combined with the first and the sample was stored on dry ice for further analysis (Winder *et al.*, 2008). In order to prepare the metabolite extracts for

GC-MS analysis, aliquots (1 mL) of each extract was spiked with 100 µL of internal standard solution (0.19 mg mL^{-1} succinic-d_4 acid, 0.27 mg mL^{-1} malonic-d_2 acid, 0.22 mg mL^{-1} glycine-d_5 in HPLC-grade water) and lyophilised (HETO VR MAXI vacuum centrifuge attached to a HETO CT/DW 60E cooling trap; Thermo Life Sciences, Basingstoke, U.K.) (Winder *et al.*, 2008).

Samples were subsequently derivatised in two stages. An aliquot (40 µL) 20 mg mL^{-1} *O*-methylhydroxylamine solution in pyridine was added and heated at 40 °C for 80 min followed by addition of 40 µL of MSTFA (*N*-acetyl-*N*-(trimethylsilyl)-trifluoroacetamide) and heating at 40 °C for 80 min. A retention index solution was added for chromatographic alignment (20 µL, 0.6 mg mL^{-1} $C_{10}/C_{12}/C_{15}/C_{19}/C_{22}$ *n*-alkanes).

The gas chromatography time-of-flight mass spectrometry (GC-TOF-MS) method used to analyse the metabolites is suitable for detection and semi-quantification of a wide range of metabolite classes as detailed in available mass spectral / retention index libraries (Kopka *et al.*, 2005) or genome-scale reconstructions of metabolic networks (Feist *et al.*, 2007) Samples were analysed in a random order employing GC-TOF-MS (Agilent 6890 GC coupled to a LECO Pegasus III TOF mass spectrometer) and using the optimal settings determined previously for yeast analysis (O'Hagan *et al.*, 2004). The raw data were processed using LECO ChromaTof v2.12 and its associated chromatographic deconvolution algorithm, with the baseline set at 1.0, data point averaging of 3, and average peak width of 2.5. A reference database was prepared which incorporated the mass spectrum and retention index of all metabolite peaks detected in a random selection of samples in order to allow the detection of all metabolites present. Each metabolite peak in the reference database was searched for in each sample, and if matched, the peak area was reported and the response ratio relative to the internal standard was calculated (Winder *et al.*, 2008).

5.3.6 Data analysis

The IR data were imported into Matlab version 7.1 (The MathWorks, Inc., Natick, MA, USA) and as an initial step CO_2 vibrations arising from the atmosphere (2403 – 2272 cm^{-1} and 683 – 656 cm^{-1}) were removed and filled with a linear trend using a protocol we have developed previously (Alsberg *et al.*, 1998). The spectra were then corrected using extended multiplicative scatter correction (EMSC) which normalises and smoothes the spectra by application of a polynomial smoothing function (Martens *et al.*,

2003). EMSC was employed as a baseline correction technique and was carried out on the entire dataset such that unavoidable baseline variations were removed; visible inspection of this processing step was used to check for the inclusion of any artefacts and these were not seen. Other potential pre-processing methods were assessed empirically and we found EMSC to correct for the non-linear background most effectively, when compared against basic scaling and normalisation functions. The pre-processed spectra were used for subsequent multivariate analyses.

Principal components analysis (PCA; (Jolliffe, 1986)) is an unsupervised multivariate projection method employed as a data reduction strategy in which principal components (PCs) are derived that explain the majority of the variance within a complex dataset. The aim of the data compression is to obtain a set of novel PCs which describe as much of the variance in the original dataset as possible. This transformation was performed prior to canonical variates analysis (CVA). CVA is a supervised learning method that seeks to minimise within-group variance whilst maximising between-group variance, and can be used in conjunction with PCA to discriminate between groups on the basis of retained principal components (PCs), given *a priori* knowledge of group membership of the spectral replicates (Manly, 1994, Winder *et al.*, 2004). In these analyses, *a priori* knowledge of class structure of the biological replicates was employed to achieve the supervision of the algorithm (Windig *et al.*, 1983, MacFie *et al.*, 1978, Goodacre *et al.*, 1998). In order to make sure these PC-CVA score plots were not over- or under-trained, the models were trained (constructed) with two of the biological replicates and cross-validated using the third biological replicate as detailed in (Jarvis and Goodacre, 2004a).

5.4 Results and discussion

Phenol quantification data (Figure 5.2) indicated that the fresh activated sludge was capable of degrading 5 mM phenol within 48 h. When aged for 2 d, a notable degradation of phenol by the microbial community was observed within 2 h, and complete degradation of 5 mM phenol was achieved within 36 h. By contrast, the microbial community aged for 9 d did not exhibit a noticeable degradation of phenol until after an 8 h lag period and complete degradation occurred within 48 h. In contrast, activated sludge samples which were aged for 16 and 131 d were not able to degrade 5 mM phenol within the 48 h period, indicating a gradual loss in the ability of the activated sludge community to degrade phenol effectively during the ageing process; a finding which has not previously been noted.

Typical FT-IR spectra of the activated sludge samples incubated with / without 5 mM phenol as the sole carbon source are shown in Figure 5.3. FT-IR spectroscopy results in the production of a composite spectrum that is representative of a wide range of endogenous and exogenous metabolites. The resultant spectrum is therefore highly complex and not open to direct interpretation of individual compounds of interest (See section 1.4.1, page 33). On initial inspection of the spectra a dominant peak is observed in both spectra at wavenumbers 2124.6 - 2082.2 cm^{-1}, this is not observed in the FT-IR spectra of bacterial samples (e.g., see Winder et al., 2006), and there are very few functional groups that vibrate in this area of the infrared spectrum. The most common are S-O from sulphates, C≡N from unsaturated N compounds, C-C stretching vibrations of alkynes and S=C=N from thiocyanates (Degen, 1997, Peoples *et al.*, 1987). It is known however, that phenol and thiocyanates are the two primary pollutants in coking effluent (Neufeld and Valiknac, 1979, Patterson, 1985, Vázquez *et al.*, 2006, Jeong and Chung, 2006, Prater and Fisher, 1991, Manefield *et al.*, 2005) and furthermore the effluent from the industrial plant in which our samples originate is known to contain thiocyanates at concentrations as high as 2.4 mM (Manefield *et al.*, 2005). It is therefore likely that this peak is attributable to the presence of thiocyanates within the activated sludge samples. In order to confirm this we obtained the FT-IR spectrum from potassium thiocyanate (see Figure 5.4) and the 2050 cm^{-1} vibration from S=C=N is clearly observed. The remaining spectral regions typify the patterns commonly associated with the FT-IR spectra of bacteria, and the major spectral bands of biological interest are highlighted in Table 5.1 (Winder *et al.*, 2004, Winder *et al.*, 2006, Zhao *et al.*, 2006, Naumann *et al.*, 1991a).

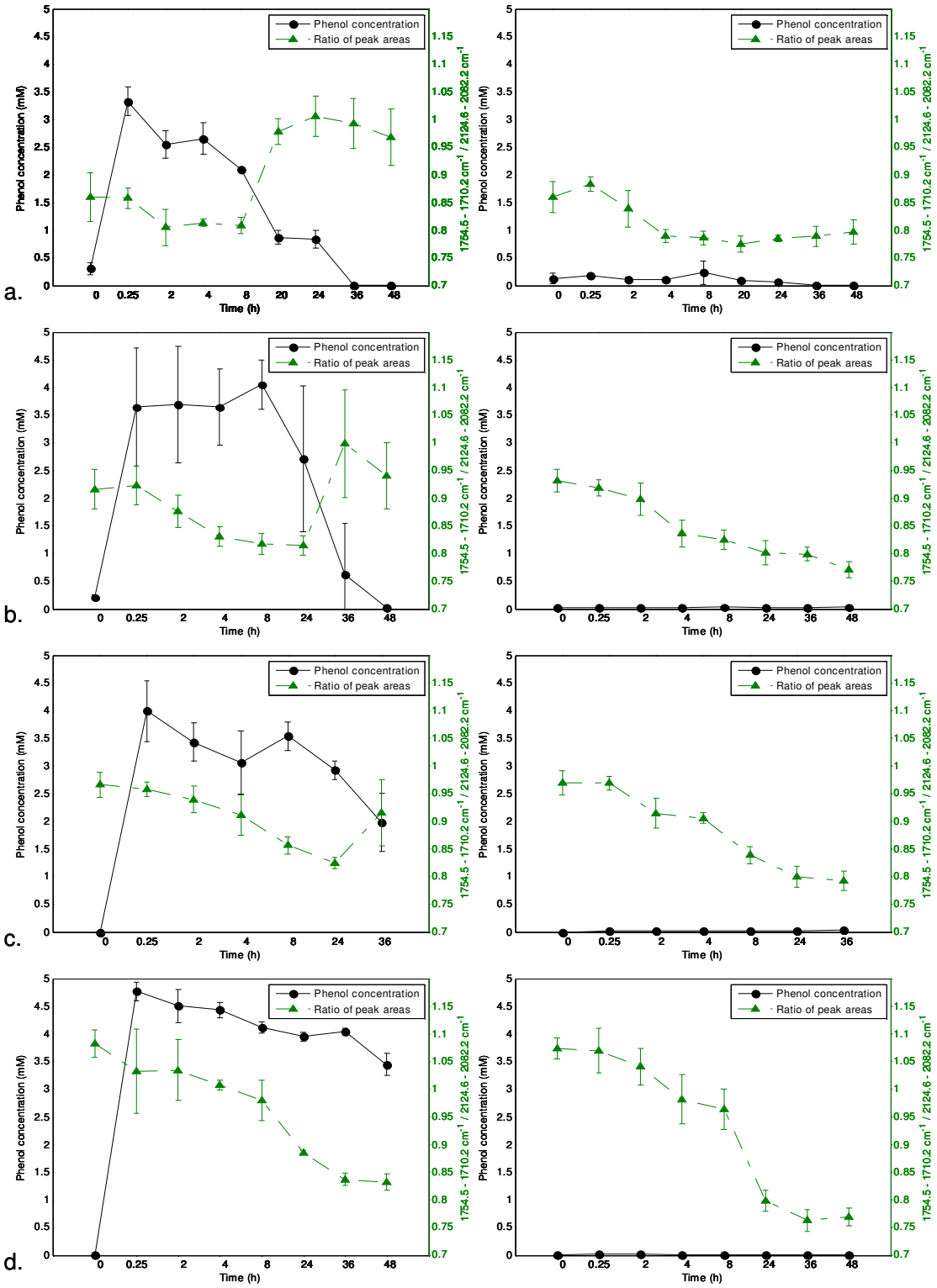


Figure 5.2: Phenol quantification data (black symbols and lines) with the ratio of COOH and thiocyanate peak areas (1754.5 - 1710.2 cm^{-1} / 2124.6 - 2082.2 cm^{-1}) shown in green. Data for all ageing experiments are shown (2, 9, 16 and 131 days; a, b, c and d respectively). Samples supplemented with 5 mM phenol are depicted on the left hand side and samples incubated in the absence of phenol (control) are on the right hand side.

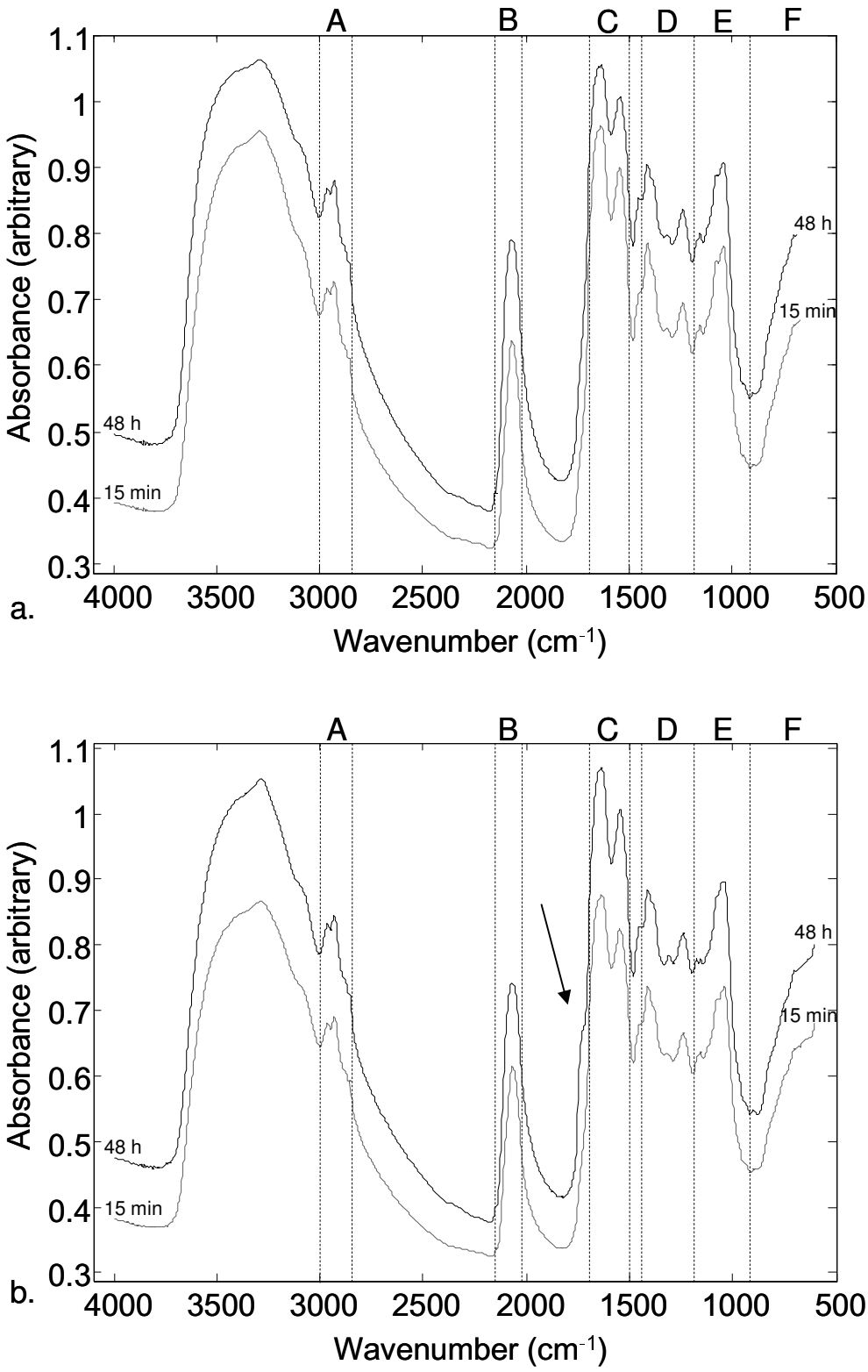


Figure 5.3: Typical FT-IR spectra of activated sludge samples incubated (a) without phenol and (b) in the presence of 5 mM phenol, both maintained at 25 °C for 48 h. Grey spectra indicate samples taken after 15 min of incubation and black offset spectra incubated for 48 h. The major spectral regions of biological interest are highlighted for both plots (See Table 5.1 for details). The arrow highlights shoulder due to carbonyl vibration from COOH.

Table 5.1: Major spectral regions of biological interest.

	Wavenumber (cm^{-1}) range	Dominant Compounds
A	3000-2800	CH_x stretches from Fatty acids
B	2130-2085	S-O C≡N from unsaturated N compounds C=N=S from isothiocyanates
C	1700-1500 (1700-1600) (1600-1500)	Proteins C=O from Amide I C-N and C-N-H from Amide II
D	1450-1200 (1250-1200)	Carboxylic groups of proteins, free amino acids, polysaccharides P-O from RNA/DNA, phospholipids
E	1200-900	C-O or O-H from Polysaccharides
F	<900	Unassigned

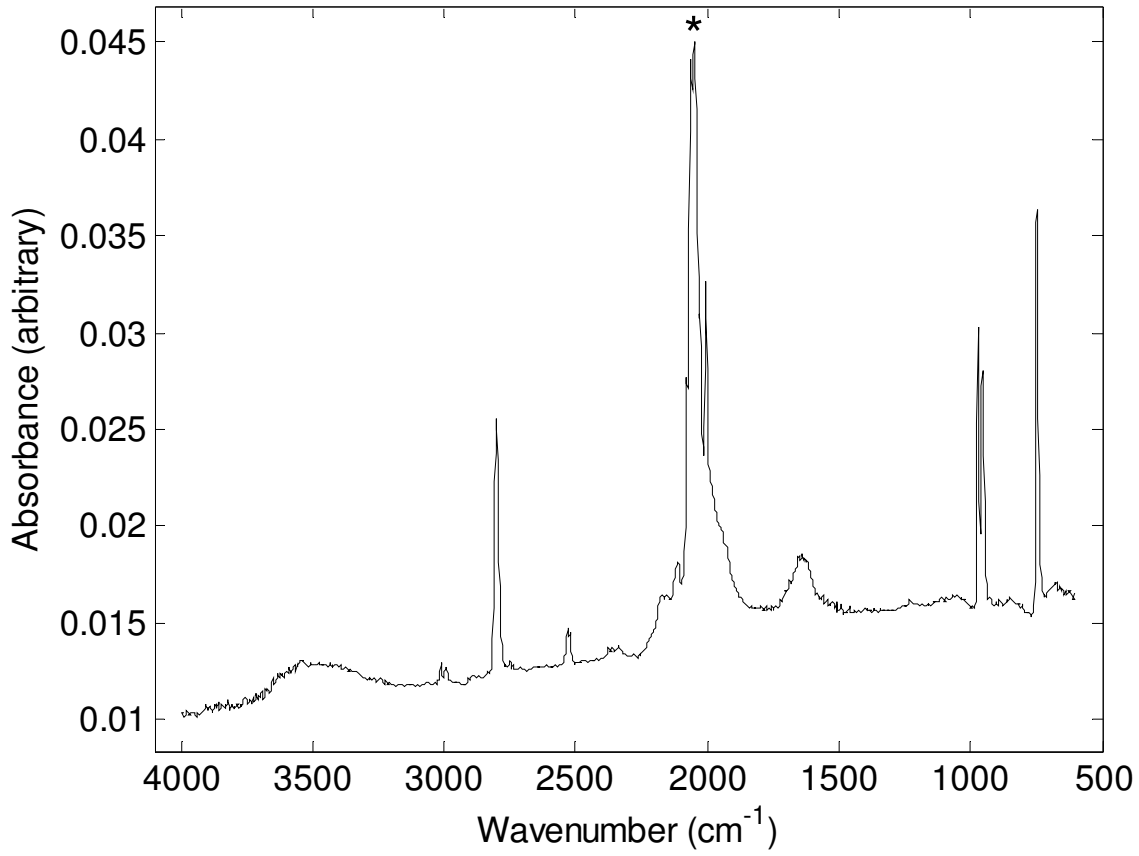


Figure 5.4: FT-IR absorbance spectrum of potassium thiocyanate. An asterisk is used to highlight the main vibration feature at 2050 cm^{-1} which arises from S=C=N.

Due to the large dimensionality and qualitative similarity observed in these spectra, multivariate analysis (MVA) methods were employed to interrogate the data, as described above. The full dataset comprised 1836 spectra (4 experiments x 17 time points x 3 biological replicates x 9 analytical replicates) and PC-CVA on this whilst possible would be very difficult to visualise. Therefore to allow a general visualisation of the phenotypic changes between the control and phenol supplemented samples across the total experimental time course, the analytical replicates from each sample condition were averaged (so that the 3 biological replicates for every time point within the ageing conditions each represented 9 FT-IR spectra) prior to PC-CVA. The PC-CVA scores plot (Figure 5.5) was constructed using the first 9 PCs (which accounts for 99.95 % of the total explained variance) with the *a priori* knowledge of the biological replicates (i.e., 68 (4 ageing experiments x 17 time points) groups). In Figure 5.5 the arrows indicate the general trends observed within the FT-IR spectra with respect to time; both in terms of storage time and experimental sampling time supplemented (or not) with phenol. It can be observed that all of the samples exhibit a strong trend with relation to the time of sampling across the first canonical variate (PC-CV 1). It can also be observed that the two experimental conditions cluster together to left of the origin which corresponds to samples either not being exposed to phenol or being exposed but not yet degrading phenol. By contrast, samples to the right of the origin bifurcate in PC-CV 2; those that continue on a horizontal trajectory are samples not supplemented with phenol, and those that increase in PC-CV 2 are now metabolically active with respect to phenol and are able to degrade it. This separation suggests that there is a strong phenotypic difference between the samples supplemented with 5 mM phenol as the sole carbon source and those maintained without the addition of phenol. Moreover, this PC-CVA scores plot indicates when the bacteria become metabolically active and able to degrade phenol. This activity appeared different and reflected the various storage conditions.

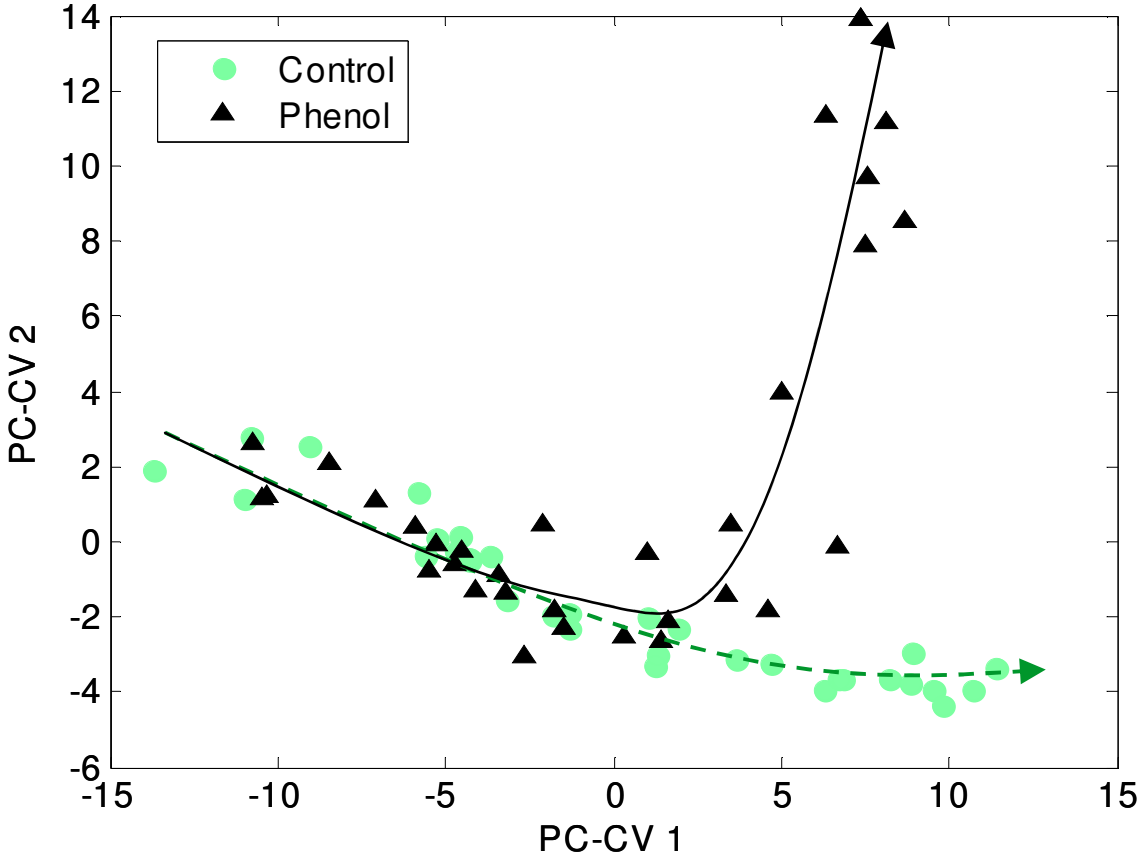


Figure 5.5: PC-CVA scores plot performed on averaged FT-IR data (from the 3 biological replicates at each of the 17 time points) from all of the "time-delay" experiments (2, 9, 16 and 131 days). PCs 1-9 were employed by the CVA algorithm with the *a priori* knowledge of the biological replicates (64 classes; 3 experiments (biological replicates) x 17 time points). The black triangles represent the samples incubated in the presence of 5 mM phenol and the green circles represent the samples incubated in the absence of phenol. The arrows illustrate a *general* trend with respect to time (including storage and time during incubation) and hold no statistical significance. (Note that unfortunately the 48 h time point for the 16 day storage sample was lost during storage).

Since a strong phenotypic difference was observed in the PC-CVA for the entire dataset, further PC-CVA was carried out for each of the time-separated investigations (Figure 5.6). This strategy was used in order to observe the effects of extended storage periods or “ageing” in the absence of phenol on this activated sludge, which may affect the metabolic potential of the community.

The analysis of the spectral data obtained from the investigation carried out on activated sludge samples aged for 2 d shows a strong trend with respect to time across PC-CV 1 (Figure 5.6a). Both the control samples and those supplemented with 5 mM phenol as the sole carbon source cluster together until 15 h, after which point the samples supplemented with phenol separate from the control samples across PC-CV 2. The trends observed here are also observed in the PC-CVA of the samples aged for 9 d (Figure 5.6b). The separation of the samples after 15 h suggests that the activated sludge communities phenotype changes significantly during the degradation of phenol. Moreover, these data support the general trends observed in the PC-CVA scores plots for the complete FT-IR dataset (Figure 5.5).

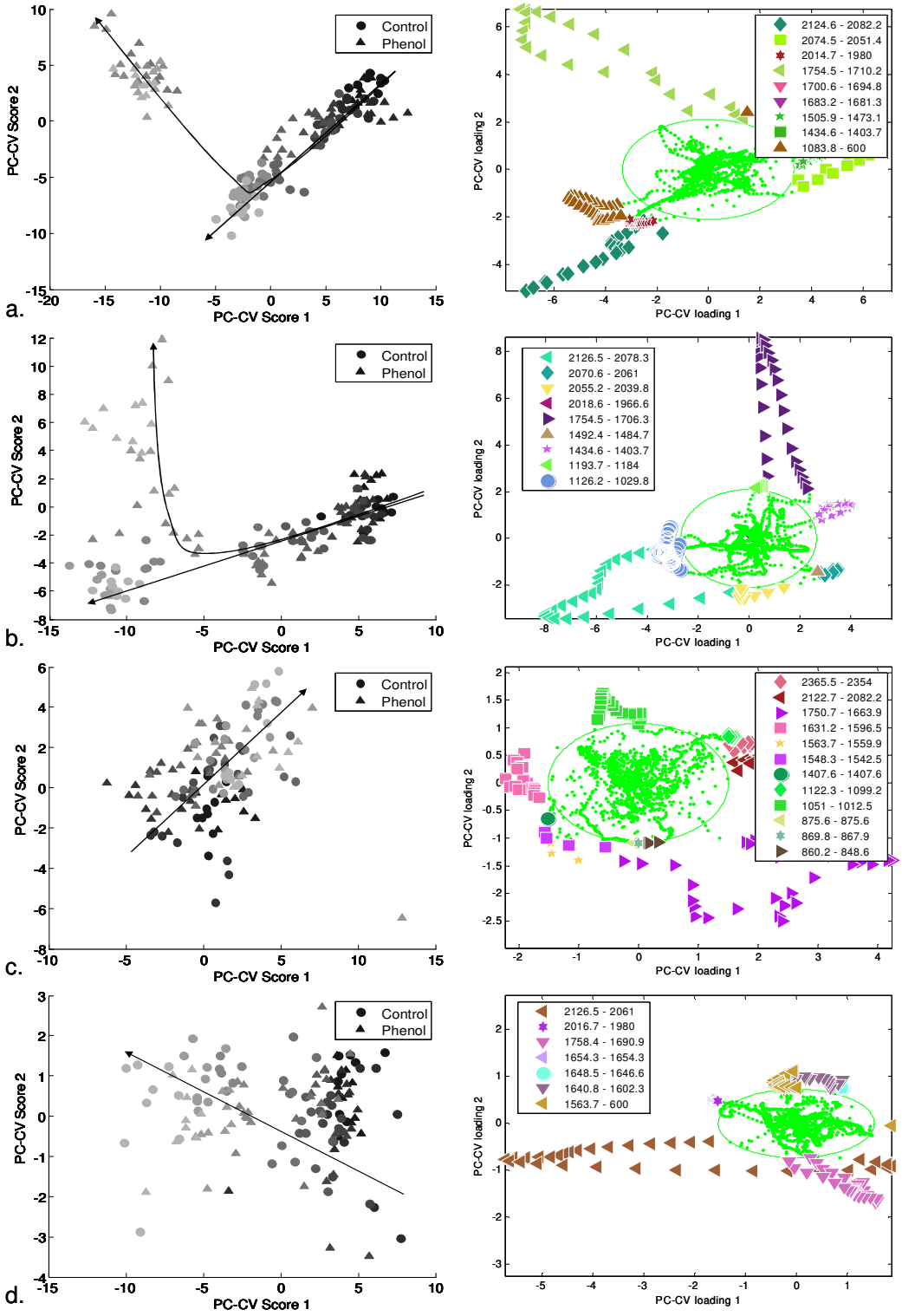


Figure 5.6: PC-CVA scores (left hand side) and loadings plots (right hand side) performed on FT-IR data from each of the "time-delay" experiments individually (2-131 days; (a) – (d) respectively). PCs 1-7 were employed by the CVA algorithm with the *a priori* knowledge of the biological replicates (17 classes for each experiment, representing the 17 time points). In the scores plots (LHS) the time at which each sample was taken following the addition of phenol is represented by a grey-scale in which the intensity decreases as the time of sampling increases. The arrows illustrate the trend with respect to time and hold no statistical significance. For the loadings biplots (RHS), contiguous spectral regions falling beyond 2 standard deviations from the mean are encoded by colours and symbols detailed in the legend.

Visual inspection of the PC-CVA loadings plots for these analyses (Figure 5.6a and Figure 5.6b, RHS) provides information relating to which spectral features are important for this separation. In this plot a boundary of 2 standard deviations from the loadings centroid was used to give a non-statistical approximation of a 95% confidence interval (this aid to visualisation highlights bands which are changing the most). The PC-CVA loadings plot for the activated sludge samples aged for 2 d (Figure 5.6a) shows that the separation of the samples supplemented with phenol at 15 h is due to a new peak occurring at 1754.5 - 1710.2 cm^{-1} in the samples supplemented with phenol (highlighted in Figure 5.3 by an arrow), and a concomitant increase in the thiocyanate peak (2124.6 - 2082.2 cm^{-1}) in the control samples. In order to highlight this further we generated an FT-IR difference spectrum by subtracting the (average control spectrum at time=48) from (average phenol spectrum at time=48), and this also highlighted the band at 1754.5 - 1710.2 cm^{-1} as being important (see Figure 5.7). Note that vibrations from phenol itself did not appear to be of importance upon inspection of the PC-CVA loadings. Closer inspection of the FT-IR spectra for the control samples showed that the thiocyanate peak broadens over time, accounting for the observed increase in the peak area during the incubation period. The PC-CVA loadings plot for the activated sludge aged for 9 d (Figure 5.6b) shows a very similar effect as described above. The fact that those activated sludge samples supplemented with 5 mM phenol are clearly separated across PC-CV 2 in both the PC-CVA scores plots (Figure 5.6a and Figure 5.6b) when these activated sludge communities are actively degrading phenol (Figure 5.2a and Figure 5.2b) supports the hypothesis that a strong phenotypic change is observed in the FT-IR spectra of the whole bacteria (i.e. the total biological components of the bacterial cells), and FT-IR spectroscopy has been used to follow metabolic events in complex microbial communities (Huang *et al.*, 2005, Huang *et al.*, 2006).

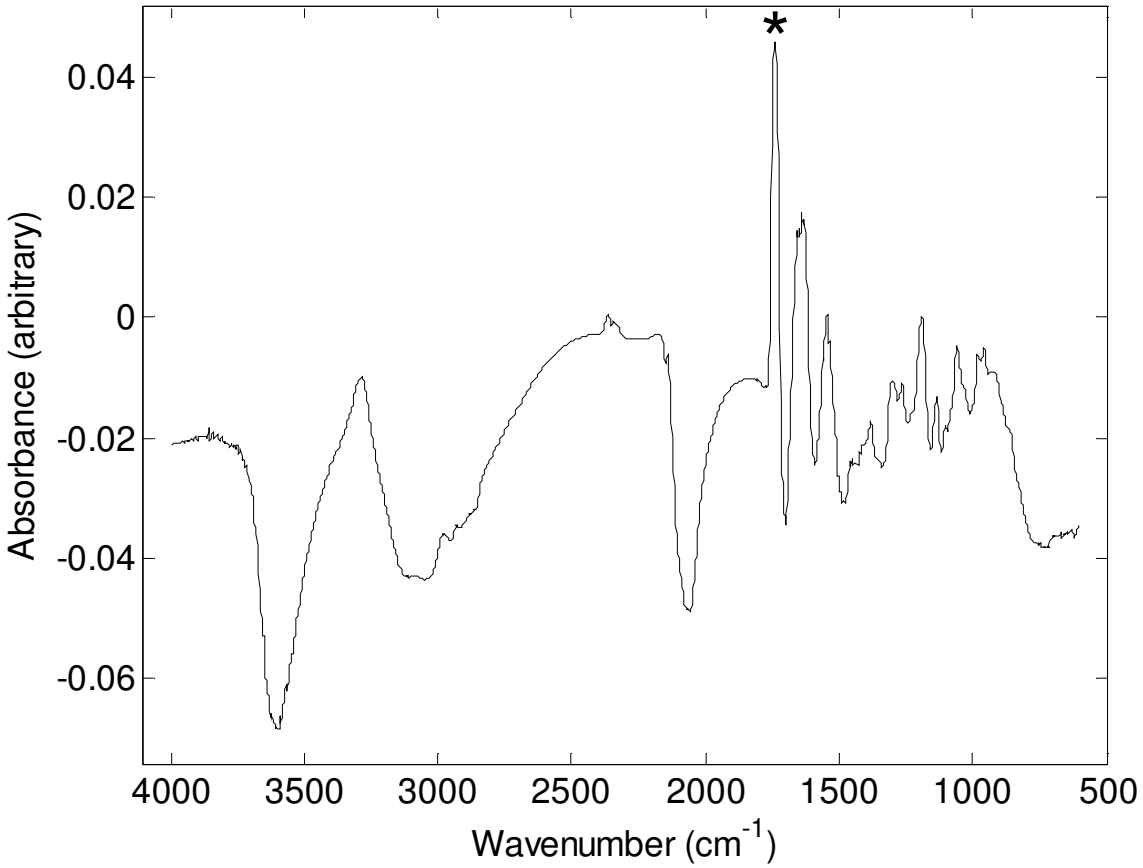


Figure 5.7: FT-IR difference spectrum of: (average phenol IR spectrum at t=48) minus (average control IR spectrum at t=48). The peak highlighted with an asterisk is the carbonyl stretch at 1754.5 - 1710.2 cm^{-1} which increases following supplementation with phenol.

This change in the FT-IR spectra for communities actively degrading phenol (Figure 5.2) is likely to be due to the community altering its metabolic potential as it induces gene transcription to produce enzymes capable of catabolising phenol and its byproducts (Agarry *et al.*, 2008). The degradation process is generally initiated by a monooxygenase phenol hydroxylase which produces catechol. Subsequently, one of two metabolic pathways may be employed. The ortho pathway results in the formation of succinyl Co-A and acetyl Co-A, and the meta pathway produces pyruvate and acetaldehyde (Schlegel, 1993, Agarry *et al.*, 2008, Leonard and Lindley, 1998, Paller *et al.*, 1995, Zhao and Ward, 1999). The initial intermediates of these pathways are detailed in Figure 5.8 and Figure 5.9.

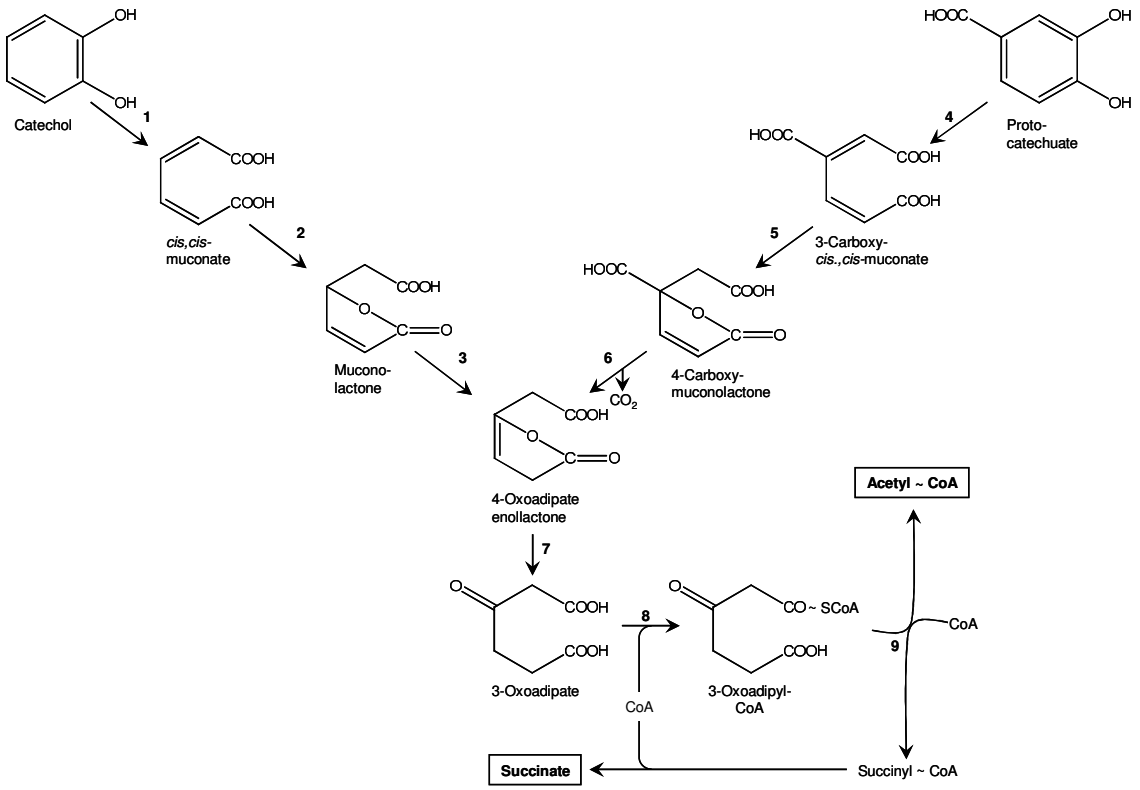


Figure 5.8: Ortho cleavage of the aromatic ring and the 3-oxoadipate pathway. Adapted from Schlegel (1993).
Enzymes involved: (1) catechol 1,2-dioxygenase; (2) muconate cycloisomerase; (3) muconolactone isomerase; (4) protocatecuate-3,4-dioxygenase; (5) 3-carboxymuconate cycloisomerase; (6) 4-carboxymuconolactone decarboxylase; (7) 4-oxoadipatenollactone hydrolase; (8) 3-oxoadipate-succinyl-CoA transferase; (9) 3-oxoadipyl-CoA thiolase.

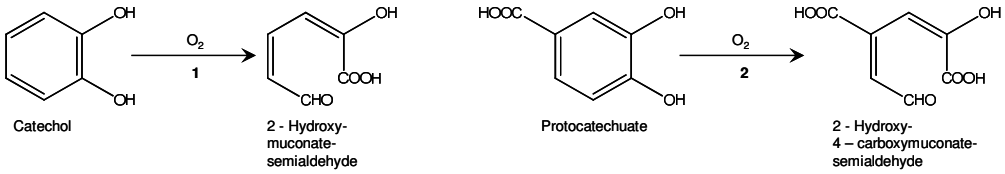


Figure 5.9: Meta cleavage of the aromatic ring. Adapted from Schlegel (1993).
Enzymes involved: (1) metapyrocatechae (catcechol-2,3-dioxygenase); (2) protocatcechuate-4,5-dioxygenase.

Unfortunately these metabolite intermediates from the ortho- or meta-cleavage of the aromatic ring in catechol are not available for FT-IR analysis. However, immediately downstream of catechol these products contain carboxylic acid groups (-COOH) and in addition to these moieties mucono-lactone and 4-oxoadipate enol-lactone also have a lactone ring that may constrain the carbonyl group (C=O) vibration. The region highlighted at 1754-1710 cm^{-1} increased after 15 h only in activated sludge samples exposed to phenol that are actively degrading it. This region is due to the carbonyl in

the COOH group. Moreover, we have also observed previously constrained carbonyl vibrations in penicillin due to its β-lactam ring that when cleaved by the enzyme β-lactamase resulted in a shift from 1767 cm^{-1} to *ca*. 1740 cm^{-1} (Winson *et al.*, 1997) providing further evidence for the detection of phenol degradation products by FT-IR spectroscopy. If the sequences of these genes are known then a targeted transcriptomics and proteomics analysis, while the sludge community is actively degrading phenol, may provide further information. This type of analysis is rather difficult in complex microbial communities particularly in communities which contain bacteria with unsequenced genomes. However, a number of methods have been developed in recent years for the proteomic analysis of complex communities (Lacerda *et al.*, 2007, Lacerda and Reardon, 2009, Ram *et al.*, 2005, Valenzuela *et al.*, 2005).

The PC-CVA scores plots for the samples aged for 16 d and 131 d can be seen in Figure 5.6c and Figure 5.6d. Whilst a trend with respect to time can be observed across PC-CV 1 and PC-CV 2 in both plots, there is no noticeable separation of the control samples and those supplemented with phenol. This observation is supported by the phenol quantification data (Figure 5.2c and Figure 5.2d) which show a visible reduction in the ability of the bacterial community to degrade phenol. Interestingly, the PC-CVA loadings plots (Figure 5.6c and Figure 5.6d) show the separation of samples according to time, irrespective of phenol degrading ability, is due to an increase in the thiocyanate peak (2122.7 – 2082.2 cm^{-1} for 16 d and 2126.5 – 2061.0 cm^{-1} for 131 d), an effect which was observed in the control samples for earlier ageing periods. This indicates that whilst the activated sludge community has lost the ability to degrade phenol during the extended storage periods, it appears to still be equivalent at the phenotypic level to the control samples stored for 2 and 9 d; a result that is also clear from the analysis of the whole dataset (Figure 5.5). This suggests that the bacterial community is still metabolically active despite the loss of phenol degrading ability.

As detailed above, the PC-CVA highlighted two major differences in the FT-IR spectra between actively degrading phenol communities and non-supplemented controls or non-degrading communities. These differences were due to changes in the peaks at wavenumbers 1754.5 - 1710.2 cm^{-1} (from the -COOH on metabolites in the phenol degradation pathways) and 2124.6 - 2082.2 cm^{-1} (from thiocyanates). In order to investigate this further and determine any time-related or phenol-degradation related trends, the phenol quantification data were plotted with a ratio of these peak areas

(1754.5 - 1710.2 cm^{-1} / 2124.6 - 2082.2 cm^{-1}) calculated from the raw spectra against time (Figure 5.2).

The phenol quantification and FT-IR peak area ratios for activated sludge samples aged for 2, 9, 16 and 131 d can be seen in Figure 5.2 (and also see Figure 5.10 for the areas under carbonyl peak). The control samples did not contain phenol (with the exception of a possible residual amount (~0.1 mM) in the 2 d old samples) and the ratio of the COOH / thiocyanate vibrations more-or-less follows the same trend and decreases with respect to time (clearly with different rates). By contrast the ratio increases in samples that are actively degrading phenol; for 2 d old activated sludge samples this occurs after 8 h and for 9 d is after 24 h, a possible trend was also observed at the 36 h time point in the 16 d aged sample, although this was less obvious than in the previous samples. For the 131 d aged activated sludge the ratio follows the same trend as the control experiments, and with the exception of the 36 h time point so do the 16 d stored samples. Inspection of the peak areas alone (Figure 5.10 and Figure 5.11) indicate that the COOH vibration also increases when the sludge is actively degrading phenol. This is not surprising given the complex biological processes involved in the degradation of phenol as discussed above (Agarry *et al.*, 2008).

It is likely that the community structure of the activated sludge changes during the degradation of phenol (Manefield *et al.*, 2005, Manefield *et al.*, 2002, Whiteley and Bailey, 2000). Changes in the structure of the bacterial community may be investigated with RNA stable isotope probing (SIP) (Manefield *et al.*, 2002, Madsen, 2006, Dumont and Murrell, 2005, Whiteley *et al.*, 2006). This data combined with metabolomic analyses may provide valuable insight into the principal bacterial species and metabolic pathways employed by the activated sludge community during the degradation of phenol, and this will be an area of future work.

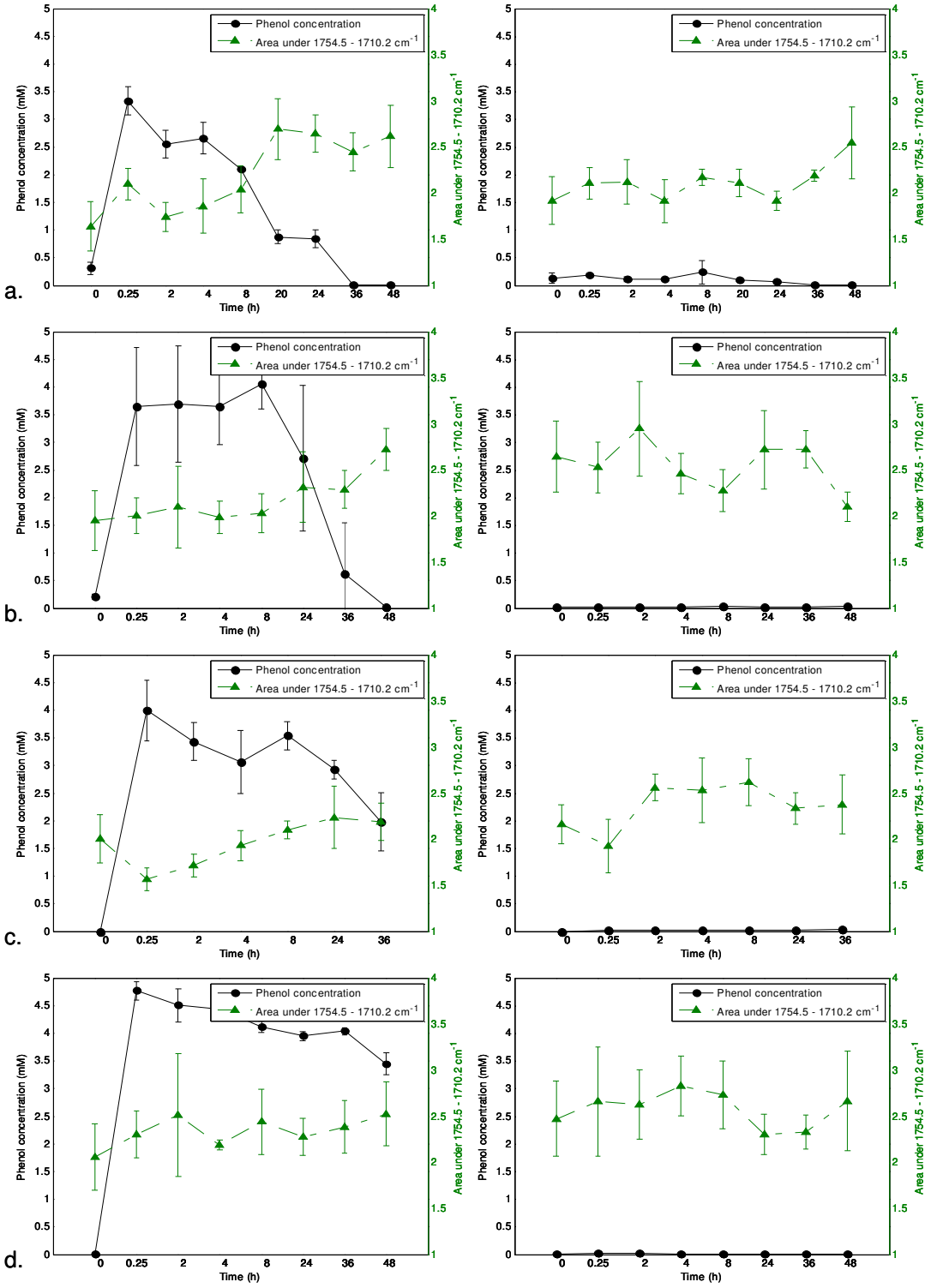


Figure 5.10: Phenol quantification data (black symbols and lines) plotted with the COOH band (1754.5 - 1710.2 cm^{-1}); FT-IR peak areas shown in green. Data for all ageing experiments are shown (2, 9, 16 and 131 days; a, b, c and d respectively). Samples supplemented with 5 mM phenol are shown on the left hand side and samples incubated in the absence of phenol (control) are on the right hand side.

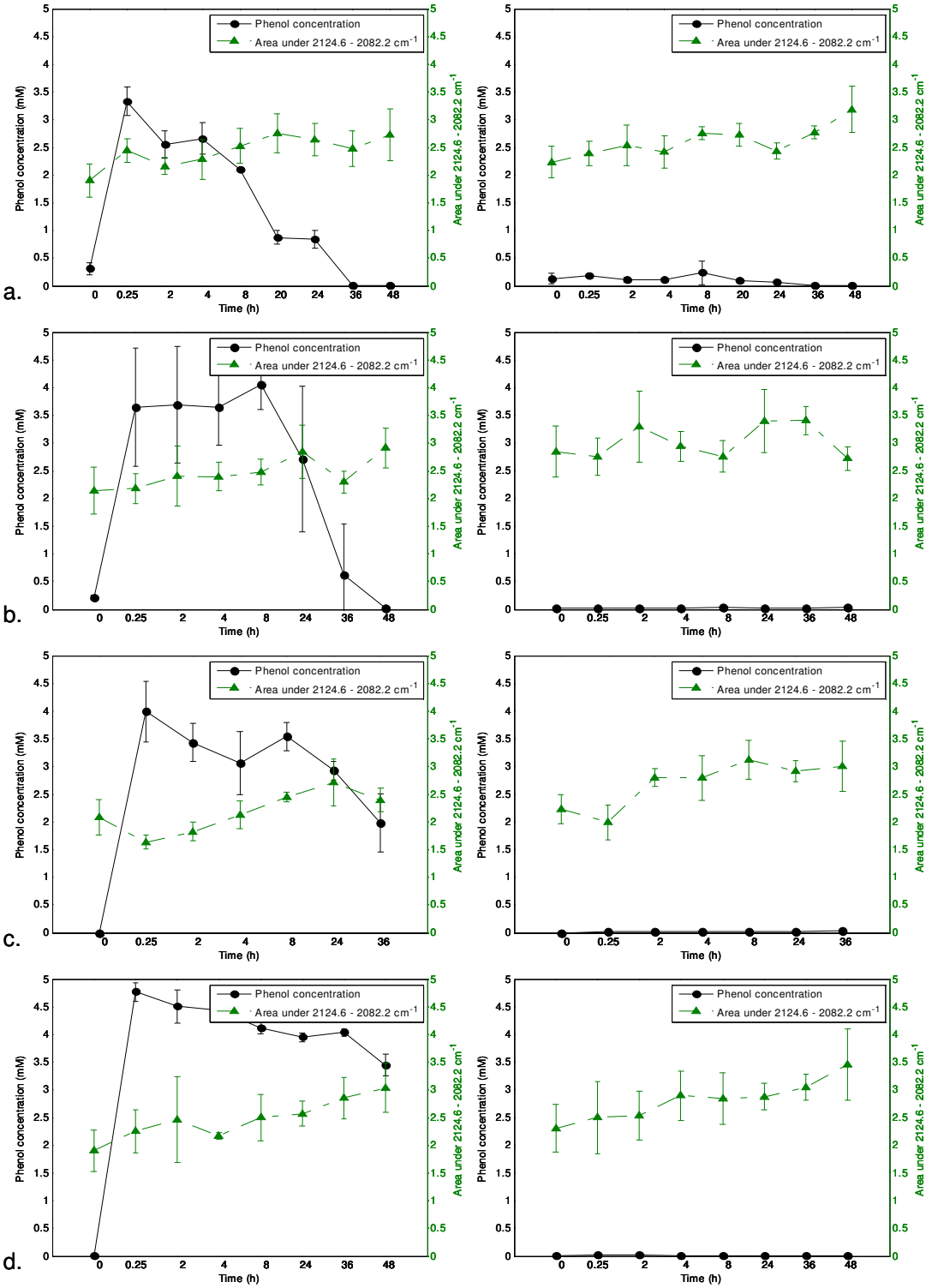


Figure 5.11: Phenol quantification data (black symbols and lines) plotted with the thiocyanate band (2124.6 - 2082.2 cm^{-1}); FT-IR peak areas shown in green. Data for all ageing experiments are shown (2, 9, 16 and 131 days; a, b, c and d respectively). Samples supplemented with 5 mM phenol are shown on the left hand side and samples incubated in the absence of phenol (control) are on the right hand side.

Metabolite profiling by GC-MS was employed in order to investigate the metabolic changes occurring in the activated sludge community during the degradation of phenol. Parallel samples were supplemented with 5 mM of either ^{12}C-phenol or $^{13}C_6$-phenol and monitored during a 48 h period, for the degradation and possible degradation products of phenol within the metabolite profiles. From the GC-MS data two pairs of metabolite peaks were identified as a mix of ^{12}C and ^{13}C analogues, thus indicating uptake of ^{13}C from the labelled phenol. The metabolite peaks were identified as the amino acids glycine and glutamic / pyroglutamic acid.

The multivariate analysis method PC-DFA was employed in order to provide a visualisation of the general trends in the metabolic profiles during the time course. Figure 5.12 shows the PC-DFA scores plot (PC-DF 1 *vs*. PC-DF 2) for the ^{12}C and $^{13}C_6$-phenol supplemented activated sludge samples during the 48 h time course. A general trend was observed across PC-DF 1 according to the time of sampling. These findings corresponded with those of the FT-IR spectroscopy analyses (Figure 5.6) and indicated a strong metabolic response to the phenol supplement within the activated sludge community. There was a clear separation across PC-DF 2 and this appeared to be due to the presence of an outlier (labelled P24). Therefore, the scores for PC-DF 1 *vs.* PC-DF 3 were plotted (Figure 5.13) in order to investigate any further changes within the metabolic profiles of the activated sludge communities during the degradation of phenol. The samples displayed a clear separation across PC-DF 1 according to the time of sampling until the 24 h time point, after which, the samples at 36 and 48 h appeared to co-cluster. It is likely that this was due to a reduction in the metabolic activity of the activated sludge following the depletion of phenol, as it was shown above that following phenol supplementation it was significantly removed by the community in the first 20 h. Some separation between the ^{12}C-phenol and $^{13}C_6$-phenol supplemented samples between 1-10 h was observed across PC-DF 3.

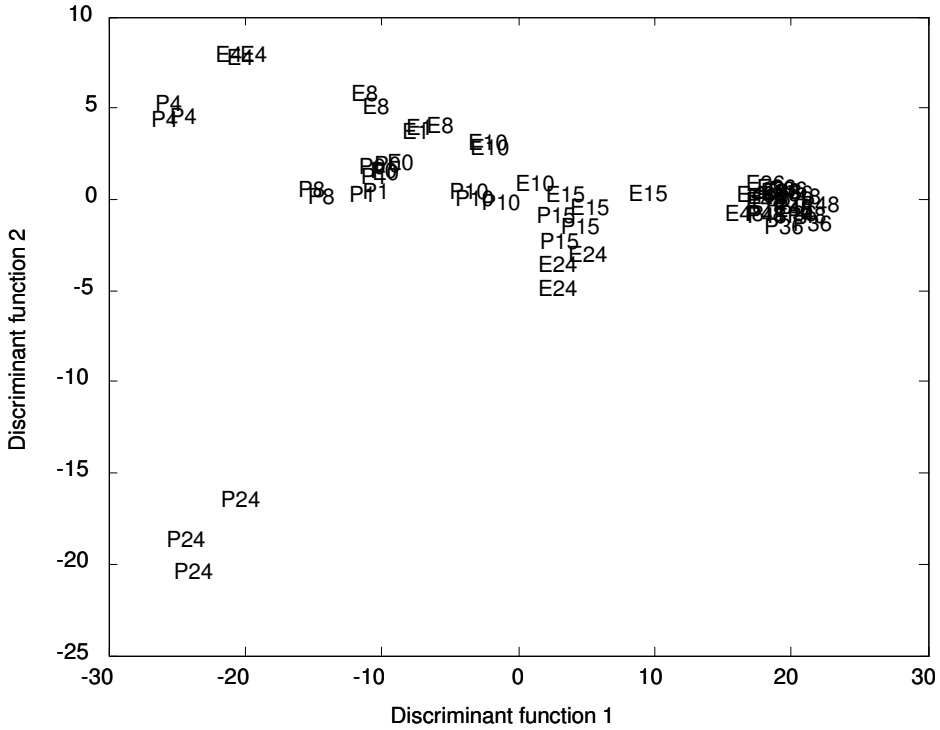


Figure 5.12: PC-DFA scores plot for the metabolic profile data of activated sludge supplemented with 5 mM phenol (^{12}C or $^{13}C_6$) over 48 h. PCs 1-20 were employed by the DFA algorithm with the *a priori* knowledge of the biological replicates. Points labelled P are those supplemented with ^{12}C phenol, and those labelled E were supplemented with ^{13}C labelled phenol.

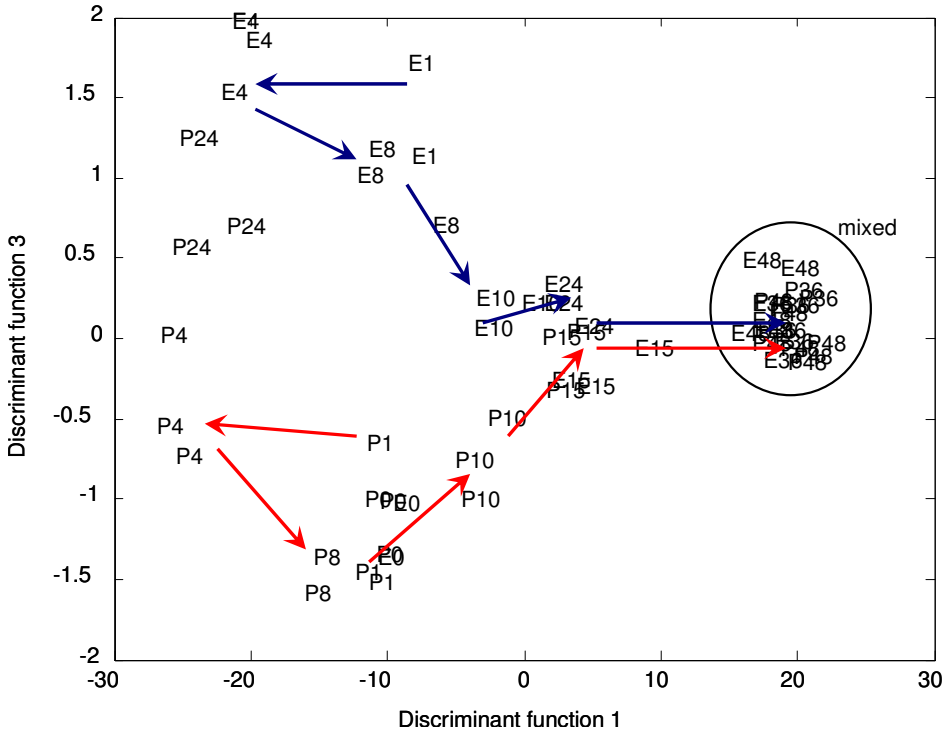


Figure 5.13: PC-DFA scores plot for the metabolic profile data of activated sludge supplemented with 5 mM phenol (^{12}C or $^{13}C_6$) over 48 h. PCs 1-20 were employed by the DFA algorithm with the *a priori* knowledge of the biological replicates. Points labelled P are those supplemented with ^{12}C phenol, and those labelled E were supplemented with ^{13}C labelled phenol. The trends are highlighted with red and blue arrows, which represent ^{12}C-phenol and $^{13}C_6$-phenol respectively.

Kruskal-Wallis analysis was carried out in order to determine any significant changes in the levels of metabolites during the time course. A total of 23 metabolite peaks were found to display significant changes of which 16 could be identified by means of matching to authentic standards by retention time and mass spectrum, or by the mass spectrum only (Winder *et al.*, 2008). Comparison of the ^{12}C-phenol and $^{13}C_6$-phenol supplemented samples with Kruskal-Wallis analysis did not find any significant changes in the metabolic profiles and it is therefore unlikely that the observed separation across PC-DF 3 in the PC-DFA scores plot (Figure 5.13) was due to a small contaminant within the activated sludge samples and may be due to the quantification ion used for defining the chromatographic peak.

Phenol was not detected by the GC-MS analysis and it is likely that it was lost during the sample preparation due to its high volatility. However a number of aromatic hydrocarbons (including fluoranthene and pyrene, see Figure 5.14) were detected and it is likely that these were hydrocarbons present in the activated sludge from the coking effluent. In addition, the presence of aromatic hydrocarbons within the activated sludge samples may be due to degradation products of phenol. It was observed that the levels of these aromatic hydrocarbons reduced during the time course, thus indicating degradation by the activated sludge. In addition, a number of sugars were detected in the metabolic profiles of the activated sludge, some of which decreased significantly during the incubation period and others displayed a significant increase in concentration during the time course. It was not possible to provide definitive identification for these sugars. The amino acids cysteine and glutamine were found to increase significantly in the 36 and 48 h time points (Figure 5.15). In addition, the carboxylic acid malonic acid was found to gradually reduce in concentration during the time course.

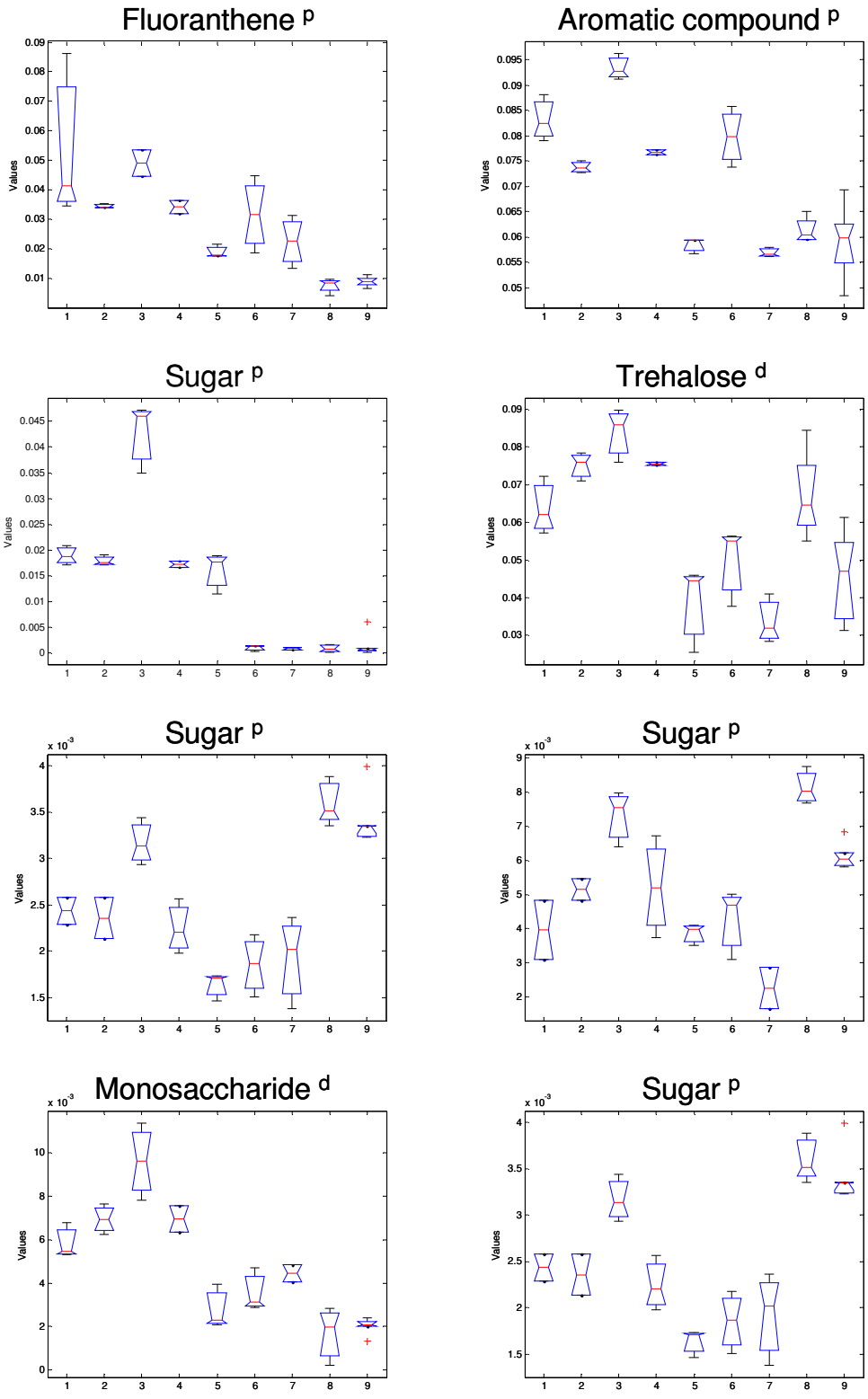


Figure 5.14: Box and whisker plots showing the changes in metabolite levels during the 48 h time course. [d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only. Classes 1-9 represent the sampling time points during the incubation period at 0, 1, 4, 8, 10, 15, 24, 36 and 48 h.

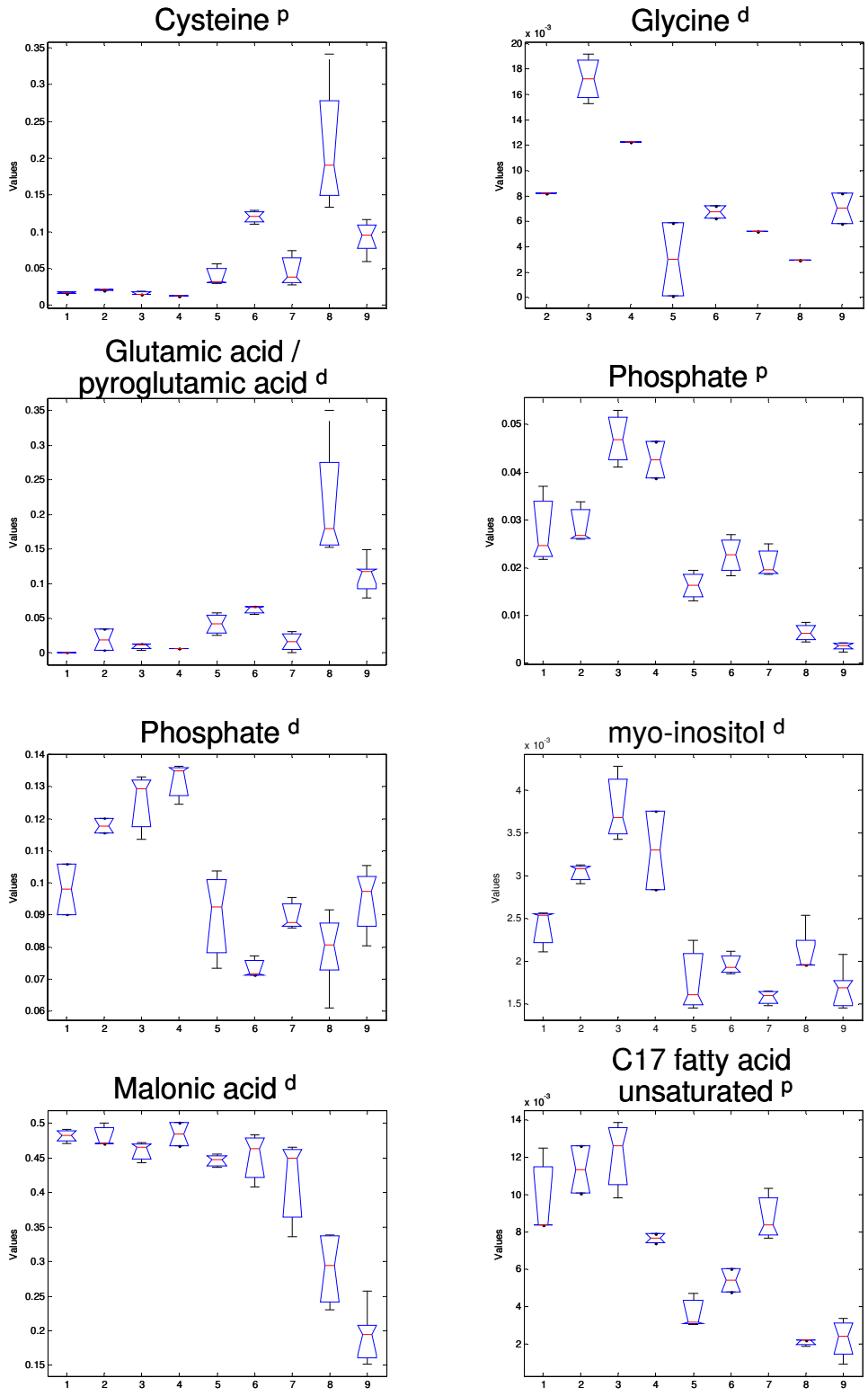


Figure 5.15: Box and whisker plots showing the changes in metabolite levels during the 48 h time course. [d] represents definitive identification by means of matching to authentic standard by retention time and mass spectrum. [p] represents putative identification by mass spectrum only. Classes 1-9 represent the sampling time points during the incubation period at 0, 1, 4, 8, 10, 15, 24, 36 and 48 h.

5.5 Conclusions

The phenol quantification data obtained from samples aged for 2 and 9 d show that the activated sludge community is capable of degrading 5 mM phenol in 48 h, and it is evident from the FT-IR data that this has a metabolic effect on the microbial cells since vibrations that are non-phenol related are clearly important. The results also show that the activated sludge community loses its ability to degrade phenol when aged for extended periods of time (16 and 131 d), and it is likely that this is due to a phenotypic adaptation from long storage in a phenol-free medium. The increase in the area of COOH vibration (C=O at 1754.5 - 1710.2cm^{-1}) during the degradation of phenol suggests that FT-IR spectroscopy is detecting the products of phenol degradation. Further analyses could target key enzymes involved in the degradation of phenol, such as monooxygenase phenol hydroxylase and catechol-2, 3- dioxygenase. However, this is rather difficult in complex communities. GC-MS analysis was able to detect significant changes in a number of metabolites during the degradation of phenol and shows the power of employing mass spectrometry-based metabolomics approaches (Dunn *et al.*, 2005, Goodacre, 2004, Winder *et al.*, 2008) for the analysis of activated sludge samples exposed to 5 mM phenol. In conclusion, high-throughput metabolic fingerprinting with FT-IR spectroscopy has been used to monitor phenotypic shifts within complex microbial communities, and this snapshot of biochemical changes supports the further investigation of this important bioremedial process using mass spectrometry-based metabolomics and this is an area of current study.

6 Conclusions and Future Work

Active pharmaceutical ingredients and their metabolites are ubiquitous in the environment and their occurrence in the aquatic environment is of growing concern. However, despite the fact that these may cause harmful effects in organisms found within this niche, little is currently known about the effects of APIs in the aquatic environment. Chiral pharmaceuticals are of particular concern as the enantiomers may be metabolised differently, with the potential for the production of an array of harmful compounds. There are many racemic APIs for treating human and animal conditions, and even in these target organisms the pharmacodynamic effects of the enantiomers are not always known. Within recent years the importance of the interactions of these compounds within the aquatic environment has been realised and information regarding the fate and biodegradation of such environmental pollutants is of great importance. The advent of post-genomic technologies has proved advantageous in the study of the effects of these environmental pollutants. The aim of this study was therefore, to investigate the effects of chiral APIs and other environmental pollutants on environmentally relevant microorganisms and the metabolic pathways in which these pollutants undergo degradation at the metabolome level.

The effects of the chiral APIs Atenolol and Propranolol were investigated in a number of prokaryotic and eukaryotic systems in order to provide a comprehensive study of the effects of the APIs in the aquatic environment. FT-IR spectroscopy was employed in the first instance as a metabolic fingerprinting tool for the phenotypic analysis of a range of environmentally relevant bacteria. In addition, FT-IR microspectroscopy was employed for the investigation of the phenotypic and localised effects of chiral APIs in a eukaryotic system.

Initially the selected bacteria were exposed to a range of concentrations of the pure enantiomers and the racemate for each of the APIs. FT-IR spectra were acquired for these exposures and a range of multivariate analysis methods were employed in the analysis of these data. Initial results showed that the data reduction strategy PCA followed by the clustering method CVA was sufficient to characterise the metabolic effects exerted on the bacteria from the FT-IR spectra. Propranolol was found to have a notable effect on the bacterial phenotype, whilst Atenolol did not appear to have a considerable phenotypic effect. The regression method PLSR was employed in the

quantification of the metabolic effects exerted on *P. putida* KT2440 when exposed to increasingly large concentrations of (±)- Propranolol and the algorithm was able to model and predict these effects successfully. In particular the pseudomonads employed in this study displayed differing phenotypic effects when exposed to Propranolol and were thus selected for exposure to a range of enantiomeric ratios of the API at a fixed concentration. Analysis of the FT-IR spectra revealed chiral-specific effects when the bacteria were exposed to the various enantiomeric ratios of Propranolol. Any potential degradation of the APIs during the growth period was investigated with HPLC. However, the results indicated that degradation of the APIs did not occur and it was therefore not possible to link the differential phenotypic effects observed in the bacterial cells to the degradation of the APIs. It was concluded that the observed phenotypic effects were more likely to be a secondary effect of the API and it is likely that this was due to interactions of Propranolol with the bacterial cell membranes. Further work could aim to provide more accurate quantification of the APIs both in the remaining supernatant and the within the bacterial cells.

It was important to consider the metabolic effects of the APIs in both prokaryotic and eukaryotic microorganisms and therefore, the green alga *Micrasterias hardyi* was selected for investigation of the chiral specific effects of Atenolol and Propranolol. Previous reports have demonstrated that Propranolol exerts a phenotypic effect on the green alga (Patel *et al*., 2008). However, the effects of Atenolol in the alga have not previously been reported, and the chiral specific effects of Propranolol were unknown for *M. hardyi*. FT-IR microspectroscopy was employed to generate chemical maps in order to monitor the general phenotypic effects throughout the whole cell and within localised areas of the algal cells. A wide range of multivariate analysis methods were employed in order to analyse both the general and localised effects of the APIs. No observable difference was reported for the algal cells exposed to Atenolol and this finding corresponded with that of the prokaryotic systems (Chapter 2). By contrast, noticeable differences were observed in the FT-IR spectra of the Propranolol exposed cells. PC-CVA revealed a significant phenotypic difference in cells exposed to Propranolol; however, the algorithm was not able to differentiate between the enantiomeric ratios. Functional group mapping was successfully employed to investigate the localised effects of the APIs within the algal cells. In addition, an enhanced metabolic effect was reported for the cells exposed to the (*R*)- enantiomer and the racemate, an effect which was not noted in the bacterial species exposed to the APIs. PCA mapping was able to analyse the entire FT-IR spectrum in a spatially

relevant manner, and was employed in order to explore the localised phenotypic effects within the algal cells further. However, this method was unable to differentiate any significant difference in the algal cells according to the experimental conditions. It was concluded that Atenolol did not exert a detectable phenotypic effect in the algal cells. Whilst Propranolol exerted a significant phenotypic effect on *M. hardyi* and all of the bacterial species studied, and a notable phenotypic effect was observed between the enantiomeric ratios.

As changes in the FT-IR spectra suggested that lipids were being altered then future work could concentrate on this and combine a GC-MS approach for exploring central metabolism with LC-MS for the analysis of the lipids. Particular attention could be paid to the difference observed in the cells exposed to the (*R*)- enantiomer and the racemate in comparison to the remaining (*R*) : (S) ratios.

The greatest phenotypic effect observed in all of the microorganisms studied was in the pseudomonads following exposure to Propranolol. GC-MS was therefore employed for the metabolic profiling of the two pseudomonads following growth in the presence of a range of enantiomeric ratios of Propranolol at a fixed concentration. Initial results demonstrated a clear phenotypic difference between the bacterial species. In addition chiral specific metabolic effects were observed for both pseudomonads when exposed to the enantiomeric ratios of Propranolol, corresponding with the results of the FT-IR spectroscopy. In particular *P. aeruginosa* PA14 displayed a different metabolic response to each of the enantiomeric ratios of (*R*)-:(*S*)- Propranolol. The multivariate and univariate analyses of the GC-MS data highlighted significant differences in a number of metabolites between the varying enantiomeric ratios which could not be identified.

Further work should aim to provide identification of these metabolites. In addition, complementary techniques such as LC-MS could be employed in order to provide a comprehensive analysis of the metabolic effects of these APIs in aquatic microorganisms. The enantiomers of the APIs are likely to be present in the environment as a range of concentrations and therefore a range of enantiomeric ratios. Further work may investigate the more subtle differences observed by GC-MS in the bacterial cells exposed to the 75:25 and 25:75 ratios (*R*)- : (*S*)- Propranolol in order to provide a more detailed characterisation of the metabolic effects of the Propranolol enantiomers. In addition, the methods employed here may be further developed to

analyse mixed culture samples which are more representative of the aquatic environment and samples from an aquatic environment known to contain API pollutants.

Aromatic hydrocarbons represent another important class of environmental pollutants. The phenotypic effects of the aromatic hydrocarbons phenol and toluene and the chiral APIs Atenolol and Propranolol were investigated in more complex biological communities. Activated sludge previously developed for the degradation of phenol in coking effluent was employed (Manefield *et al.*, 2005, Manefield *et al.*, 2002) and FT-IR spectroscopy was utilised to monitor the phenotypic effects of the environmental pollutants during a 48 h period. In addition HPLC was employed to quantify the pollutants remaining in the supernatant following incubation. However, HPLC did not successfully quantify the pollutants due to the complex nature of the community flocs. FT-IR spectroscopy was only able to monitor the phenotypic effects occurring in the microbial community supplemented with phenol and results from the multivariate analyses suggested that these changes were due to shifts in the levels of thiocyanates found within the activated sludge, and this was further explored with a range of ageing experiments and metabolic fingerprinting (Chapter 5). By contrast, no notable difference was observed in the communities exposed to the other environmental pollutants. In contrast to the findings presented in Chapter 2 and Chapter 3, the APIs did not appear to exert a phenotypic effect on the activated sludge community. In addition, no detectable phenotypic effects were observed in the FT-IR analysis of the activated sludge following exposure to the aromatic hydrocarbon toluene.

Whilst this complex microbial community had been previously investigated using molecular approaches for the identification of the main phenol degrader (Manefield *et al.*, 2005, Manefield *et al.*, 2002), little is known regarding the changes in the community phenotype that may occur within the activated sludge when it is actively degrading phenol. FT-IR spectroscopy was employed in a temporal manner to monitor the community phenotype during phenol degradation. Phenol consumption was confirmed using a colorimetric assay for phenol and degradation took *ca*. 36 h but only in 'young' cultures compared to communities that had been allowed to age for between 16 and 131 days. Multivariate analyses of the FT-IR spectra showed different phenotypic trajectories depending on whether the community was able to degrade phenol or whether this metabolic potential was lost.

During incubation changes in the FT-IR spectra of the control samples were dominated by changes in the levels of thiocyanates. By contrast, a new peak was only observed in the FT-IR spectra of the phenol supplemented communities and this could be attributed to a COOH group vibration that was likely due to phenol degradation products from either the ortho- or meta-cleavage pathways. Armed with this knowledge the activated sludge community was monitored during the active degradation of phenol with the use of GC-MS. Further work could target these pathways and monitor the activated sludge for the production of these intermediates. It would be particularly interesting to see if changes in the community structure occur during phenol degradation as this may provide evidence of cross-feeding, and this could be achieved using the RNA-SIP molecular methods detailed in Manefield *et al.,* (2002).

This work demonstrated the utility of FT-IR spectroscopy to investigate the phenotypic variations in complex biological communities during exposure to environmental pollutants. Further work in this area should develop the methods presented here for monitoring phenotypic changes in complex communities. In addition alternative methods for quantifying the remaining levels of the pollutants should be explored, which may involve specific solvent extractions.

In conclusion, the work presented in this thesis has shown for the first time that metabolomics allows subtle phenotypes in microorganisms to be revealed when they are exposed to chiral forms of APIs which are commonly found in the environment. Despite these APIs not being designed for any interaction with bacteria and aquatic life in general these are significant findings and may have implications as more and more APIs become detectable and concentrated in the environment due to continued use in man and indeed animals or aquaculture. I believe that metabolomics will play an important role in toxicological studies, in understanding pollutant-species interactions, and should become an important tool for environmental monitoring.

7 References

ABUHAMED, T., BAYRAKTAR, E., MEHMETOGLU, T. & MEHMETOGLU, U. (2004) Kinetics model for growth of *Pseudomonas putida* F1 during benzene, toluene and phenol biodegradation. *Process Biochemistry,* 39**,** 983-988.

AGARRY, S. E., DUROJAIYE, A. O. & O., S. B. (2008) Microbial degradation of phenols: a review. *International Journal of Environment and Pollution,* 32**,** 12-28.

AHMARUZZAMAN, M. & SHARMA, D. K. (2005) Adsorption of phenols from wastewater. *Journal of Colloid and Interface Science,* 287**,** 14-24.

ALLEN, J., DAVEY, H. M., BROADHURST, D., ROWLAND, J. J., OLIVER, S. G. & KELL, D. B. (2004) Discrimination of Modes of Action of Antifungal Substances by Use of Metabolic Footprinting. *Applied and Environmental Microbiology,* 70**,** 6157-6165.

ALLWOOD, J. W., ELLIS, D. I., HEALD, J. K., GOODACRE, R. & MUR, L. A. J. (2006) Metabolomic approaches reveal that phosphatidic and phosphatidyl glycerol phospholipids are major discriminatory non-polar metabolites in responses by *Brachypodium distachyon* to challenge by *Magnaporthe grisea*. *The Plant Journal,* 46**,** 351-368.

ALSBERG, B. K., WADE, W. G. & GOODACRE, R. (1998) Chemometric Analysis of Diffuse Reflectance-Absorbance Fourier Transform Infrared Spectra Using Rule Induction Methods: Application to the Classification of *Eubacterium* Species. *Applied Spectroscopy,* 52**,** 823-832.

ALVAREZ, P. J. & VOGEL, T. M. (1991) Substrate interactions of benzene, toluene, and para-xylene during microbial degradation by pure cultures and mixed culture aquifer slurries. *Applied and Environmental Microbiology,* 57**,** 2981-2985.

ANDRUS, P. G. L. & STRICKLAND, R. D. (1998) Cancer grading by Fourier transform infrared spectroscopy. *Biospectroscopy,* 4**,** 37-46.

ARANÌBAR, N., SINGH, B. K., STOCKTON, G. W. & OTT, K.-H. (2001) Automated Mode-of-Action Detection by Metabolic Profiling. *Biochemical and Biophysical Research Communications,* 286**,** 150-155.

ARNHEIM, N. & ERLICH, H. (1992) Polymerase Chain Reaction Strategy. *Annual Review of Biochemistry,* 61**,** 131-156.

ARNIM, T. V. & WELMAN, E. (1981) Propranolol in ischaemic-reperfused working rat heart: Dissociation of beta-adrenergic blocking and protective effect. *Journal of Molecular and Cellular Cardiology,* 13**,** 521-524.

ASHTON, D., HILTON, M. & THOMAS, K. V. (2004) Investigating the environmental transport of human pharmaceuticals to streams in the United Kingdom. *Science of The Total Environment,* 333**,** 167-184.

AYSCOUGH, N. J., FAWELL, J., FRANKLIN, G. & YOUNG, W. (2000) Review of human pharmaceuticals in the environment. Environment Agency.

BANWELL, C. N. & MCCASH, E. M. (1994) *Fundamentals of Molecular Spectroscopy,* London, McGraw-Hill.

BARRETT, A. M. & CULLUM, V. A. (1968) The biological properties of the optical isomers of propranolol and their effects on cardiac arrhythmias. *British Journal of Pharmacology,* 34**,** 43-55.

BARTLETT, J. (2001) Technology evaluation: SAGE, Genzyme Molecular Oncology. *Current Opinion in Molecular Therapeutics,* 3**,** 85-96.

BARTON, H. A. & NORTHUP, D. E. (2007) Geomicrobiology in cave environments: past, current and future perspectives. *Journal of Cave and Karst Studies,* 69**,** 163–178.

BEARDALL, J., BERMAN, T., HERAUD, P., OMO KADIRI, M., LIGHT, B. R., PATTERSON, G., ROBERTS, S., SULZBERGER, B., SAHAN, E., UEHLINGER, U. & WOOD, B. (2001) A comparison of methods for detection of phosphate limitation in microalgae. *Aquatic Sciences - Research Across Boundaries,* 63**,** 107-121.

BELTON, P. S., SAFFA, A. M. & WILSON, R. H. (1987) Use of Fourier transform infrared spectroscopy for quantitative analysis: a comparative study of different detection methods. *Analyst,* 112**,** 1117 - 1120.

BHARGAVA, R., WANG, S.-Q. & KOENIG, J. L. (1998) FT-IR Imaging of the Interface in Multicomponent Systems Using Optical Effects Induced by Differences in Refractive Index. *Applied Spectroscopy,* 52**,** 323-328.

BINO, R. J., HALL, R. D., FIEHN, O., KOPKA, J., SAITO, K., DRAPER, J., NIKOLAU, B. J., MENDES, P., ROESSNER-TUNALI, U., BEALE, M. H., TRETHEWEY, R. N., LANGE, B. M., WURTELE, E. S. & SUMNER, L. W. (2004) Potential of metabolomics as a functional genomics tool. *Trends in Plant Science,* 9**,** 418-425.

BITTON, G. (2005) *Wastewater Microbiology,* New York, John Wiley & Sons.

BOPP, S. K. & LETTIERI, T. (2007) Gene regulation in the marine diatom *Thalassiosira pseudonana* upon exposure to polycyclic aromatic hydrocarbons (PAHs). *Gene,* 396**,** 293-302.

BOROS, L. G., BRACKETT, D. J. & HARRIGAN, G. G. (2003) Metabolic Biomarker and Kinase Drug Target Discovery in Cancer Using Stable Isotope-Based Dynamic Metabolic Profiling (SIDMAP). *Current Cancer Drug Targets,* 3**,** 445-453.

BOROS, L. G., CASCANTE, M. & PAUL LEE, W.-N. (2002) Metabolic profiling of cell growth and death in cancer: applications in drug discovery. *Drug Discovery Today,* 7**,** 364-372.

BOROS, L. G., SERKOVA, N. J., CASCANTE, M. S. & LEE, W.-N. P. (2004) Use of metabolic pathway flux information in targeted cancer drug design. *Drug Discovery Today: Therapeutic Strategies,* 1**,** 435-443.

BOSAKOVA, Z., CURINOVA, E. & TESAROVA, E. (2005) Comparison of vancomycin-based stationary phases with different chiral selector coverage for enantioselective separation of selected drugs in high-performance liquid chromatography. *Journal of Chromatography A,* 1088**,** 94-103.

BOUHEDJA, W., SOCKALINGUM, G. D., PINA, P., ALLOUCH, P., BLOY, C., LABIA, R., MILLOT, J. M. & MANFAIT, M. (1997) ATR-FTIR spectroscopic investigation of *E. coli* transconjugants β-lactams-resistance phenotype. *FEBS Letters,* 412**,** 39-42.

BRADLEY, B. & THEODORAKIS, C. (2002) The post-genomic era and ecotoxicology. *Ecotoxicology,* 11**,** 7-9.

BRINDLE, J. T., ANTTI, H., HOLMES, E., TRANTER, G., NICHOLSON, J. K., BETHELL, H. W. L., CLARKE, S., SCHOFIELD, P. M., MCKILLIGIN, E., MOSEDALE, D. E. & GRAINGER, D. J. (2002) Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nature Medicine,* 8**,** 1439-1445.

BRO, C. & NIELSEN, J. (2004) Impact of 'ome' analyses on inverse metabolic engineering. *Metabolic Engineering,* 6**,** 204-211.

BROADHURST, D. & KELL, D. B. (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics,* V2**,** 171-196.

BROWN, M., DUNN, W. B., DOBSON, P., PATEL, Y., WINDER, C. L., FRANCIS-MCINTYRE, S., BEGLEY, P., CARROLL, K., BROADHURST, D., TSENG, A., SWAINSTON, N., SPASIC, I., GOODACRE, R. & KELL, D. B. (2009) Mass

spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *The Analyst,* 134**,** 1322-1332.

BROWN, M., DUNN, W. B., ELLIS, D. I., GOODACRE, R., HANDL, J., KNOWLES, J. D., O'HAGAN, S., SPASIC, I. & KELL, D. B. (2005) A metabolome pipeline: from concept to data to knowledge. *Metabolomics,* 1**,** 39-51.

BUCHHOLZ, A., HURLEBAUS, J., WANDREY, C. & TAKORS, R. (2002) Metabolomics: quantification of intracellular metabolite dynamics. *Biomolecular Engineering,* 19**,** 5-15.

BULLEN, H. A., OEHRLE, S. A., BENNETT, A. F., TAYLOR, N. M. & BARTON, H. A. (2008) The Use of Attenuated Total Reflectance Fourier Transformed Infrared (ATR-FTIR) Spectroscopy to Identify Microbial Metabolic Products on Carbonate Mineral Surfaces. *Applied and Environmental Microbiology***,** AEM.02936-07.

BUSER, H.-R., POIGER, T. & MULLER, M. D. (1999) Occurrence and Environmental Behavior of the Chiral Pharmaceutical Drug Ibuprofen in Surface Waters and in Wastewater. *Environmental Science & Technology,* 33**,** 2529-2535.

BUSSELL, J. A., GIDMAN, E. A., CAUSTON, D. R., GWYNN-JONES, D., MALHAM, S. K., JONES, M. L. M., REYNOLDS, B. & SEED, R. (2008) Changes in the immune response and metabolic fingerprint of the mussel, *Mytilus edulis* (Linnaeus) in response to lowered salinity and physical stress. *Journal of Experimental Marine Biology and Ecology,* 358**,** 78-85.

BYLER, D. M. & SUSI, H. (1986) Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers,* 25**,** 469-487.

CARLSSON, C., JOHANSSON, A.-K., ALVAN, G., BERGMAN, K. & KÜHLER, T. (2006a) Are pharmaceuticals potent environmental pollutants?: Part II: Environmental risk assessments of selected pharmaceutical excipients. *Science of The Total Environment,* 364**,** 88-95.

CARLSSON, C., JOHANSSON, A.-K., ALVAN, G., BERGMAN, K. & KUHLER, T. (2006b) Are pharmaceuticals potent environmental pollutants?: Part I: Environmental risk assessments of selected active pharmaceutical ingredients. *Science of The Total Environment,* 364**,** 67-87.

CARLSSON, G., ÖRN, S. & LARSSON, D. G. J. (2009) Effluent from bulk drug production is toxic to aquatic vertebrates. *Environmental Toxicology and Chemistry,* 28**,** 2656-2662.

CHAPMAN, R. W. (2001) EcoGenomics--a consilience for comparative immunology? *Developmental & Comparative Immunology,* 25**,** 549-551.

CHICKOS, J. S., GARIN, D. L., HITT, M. & SCHILLING, G. (1981) Some solid state properties of enantiomers and their racemates. *Tetrahedron,* 37**,** 2255-2259.

CHOO-SMITH, L. P., MAQUELIN, K., VAN VREESWIJK, T., BRUINING, H. A., PUPPELS, G. J., THI, N. A. N., KIRSCHNER, C., NAUMANN, D., AMI, D., VILLA, A. M., ORSINI, F., DOGLIA, S. M., LAMFARRAJ, H., SOCKALINGUM, G. D., MANFAIT, M., ALLOUCH, P. & ENDTZ, H. P. (2001) Investigating Microbial (Micro)colony Heterogeneity by Vibrational Spectroscopy. *Applied and Environmental Microbiology,* 67**,** 1461-1469.

CLARK, T., LEE, S., SCOTT, L. R. & WANG, S. M. (2002) Computational Analysis of Gene Identification with SAGE. *Journal of Computational Biology,* 9**,** 513-526.

CRANE, M., WATTS, C. & BOUCARD, T. (2006) Chronic aquatic environmental risks from exposure to human pharmaceuticals. *Science of the Total Environment,* 367**,** 23-41.

CSTEE (2001) Opinion on ‘‘Draft discussion paper on environmental risk assessment of medicinal products for human use’’. Brussels, Belgium, European Commission.

CURK, M. C., PELEDAN, F. & HUBERT, J. C. (1994) Fourier transform infrared (FTIR) spectroscopy for identifying *Lactobacillus* species. *FEMS Microbiology Letters,* 123**,** 241-248.

DAUGHTON, C. G. & TERNES, T. A. (1999) Pharmaceuticals and Personal Care Products in the Environment: Agents of Subtle Change? *Environmental Health Perspectives,* 107**,** 907-938.

DAVIES, C. L. (1990) Chromatography of β-adrenergic blocking agents. *Journal of Chromatography B: Biomedical Sciences and Applications,* 531**,** 131-180.

DEGEN, I. A. (1997) *Tables of characteristic group frequencies for the interpretation of infrared and Raman spectra,* Harrow, UK, Acolyte Publication.

DEPLEDGE, M. H. (1994) Genotypic toxicity: Implications for individuals and populations. *Environmental Health Perspectives,* 102**,** 101-104.

DREWES, J. E., HEBERER, T. & REDDERSEN, K. (2002) Fate of pharmaceuticals during indirect potable reuse. *Water Science & Technology,* 46**,** 73–80.

DU PLESSIS, C. A., BARNARD, P., MUHLBAUER, R. M. & NALDRETT, K. (2001) Empirical model for the autotrophic biodegradation of thiocyanate in an activated sludge reactor. *Letters in Applied Microbiology,* 32**,** 103-107.

DUMONT, M. G. & MURRELL, J. C. (2005) Stable isotope probing - linking microbial identity to function. *Nature Reviews Microbiology,* 3**,** 499-504.

DUNN, W. B., BAILEY, N. J. C. & JOHNSON, H. E. (2005) Measuring the metabolome: current analytical technologies. *The Analyst,* 130**,** 606-625.

DUNN, W. B. & ELLIS, D. I. (2005) Metabolomics: Current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry,* 24**,** 285-294.

EGGEN, R. I. L. & SUTER, M. J. F. (2007) Analytical chemistry and ecotoxicology - Tasks, needs and trends. *Journal of Toxicology and Environmental Health - Part A: Current Issues,* 70**,** 724-726.

EMERSON, E. (1943) The condensation of aminoantipyrine. II. A new color test for phenolic compounds. *The Journal of Organic Chemistry,* 8**,** 417-428.

EMERSON, E. & KELLY, K. (1948) The condensation of aminoantipyrine. VI. A study of the effect of excess base on the reaction of amino-antipyrine with phenolic compounds in the presence of oxidizing agents. *The Journal of Organic Chemistry,* 13**,** 532-534.

ENVIRONMENT AGENCY (1995) Proposed environmental quality standards for phenol in water. Environment Agency, United Kingdom.

ERIKSSON, L., ANTTI, H., GOTTFRIES, J., HOLMES, E., JOHANSSON, E., LINDGREN, F., LONG, I., LUNDSTEDT, T., TRYGG, J. & WOLD, S. (2004) Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Analytical and Bioanalytical Chemistry,* 380**,** 419-429.

ESCHER, B. I., BRAMAZ, N., EGGEN, R. I. L. & RICHTER, M. (2005) In Vitro Assessment of Modes of Toxic Action of Pharmaceuticals in Aquatic Life. *Environmental Science & Technology,* 39**,** 3090-3100.

FABIAN, H. & MÄNTELE, W. (2002) Infrared spectroscopy of proteins. IN CHALMERS, J. M. & P.R., G. (Eds.) *Handbook of Vibrational Spectroscopy.* Chichester, John Wiley & Sons Ltd.

FAN, T., LANE, A., HIGASHI, R., FARAG, M., GAO, H., BOUSAMRA, M. & MILLER, D. (2009) Altered regulation of metabolic pathways in human lung cancer discerned by ^{13}C stable isotope-resolved metabolomics (SIRM). *Molecular Cancer,* 8**,** 41.

FEIST, A. M., HENRY, C. S., REED, J. L., KRUMMENACKER, M., JOYCE, A. R., KARP, P. D., BROADBELT, L. J., HATZIMANIKATIS, V. & PALSSON, B. O. (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology,* 3.

FENT, K., WESTON, A. A. & CAMINADA, D. (2006) Ecotoxicology of human pharmaceuticals. *Aquatic Toxicology,* 76**,** 122-159.

FIEHN, O. (2001) Combining Genomics, Metabolome Analysis, and Biochemical Modelling to Understand Metabolic Networks. *Comparative and Functional Genomics,* 2**,** 155-168.

FIEHN, O. (2002) Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biology,* 48**,** 155-171.

FIEHN, O., KOPKA, J., DÃ¶RMANN, P., ALTMANN, T., TRETHEWEY, R. N. & WILLMITZER, L. (2000) Metabolite profiling for plant functional genomics. *Nature Biotechnology,* 18**,** 1157-1161.

FONO, L. J. & SEDLAK, D. L. (2005) Use of the Chiral Pharmaceutical Propranolol to Identify Sewage Discharges into Surface Waters. *Environmental Science & Technology,* 39**,** 9244-9252.

FORBES, N. S., MEADOWS, A. L., CLARK, D. S. & BLANCH, H. W. (2006) Estradiol stimulates the biosynthetic pathways of breast cancer cells: Detection by metabolic flux analysis. *Metabolic Engineering,* 8**,** 639-652.

FORSTER, J., FAMILI, I., FU, P., PALSSON, B. Ã. & NIELSEN, J. (2003) Genome-Scale Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network. *Genome Research,* 13**,** 244-253.

GASPARIC, J., SVOBODOVAA, D. & MATYSOVA, A. (1974) Identification of organic compounds : LXXXI. A contribution to the chromatography of phenols as antipyrine dyes. *Journal of Chromatography A,* 88**,** 364-369.

GELADI, P. (1988) Notes on the history and nature of partial least squares (PLS) modelling. *Journal of Chemometrics,* 2**,** 231-246.

GELADI, P. & KOWALSKI, B. R. (1986) Partial least-squares regression: a tutorial. *Analytica Chimica Acta,* 185**,** 1-17.

GEMPERLINE, P. J., LONG, J. R. & GREGORIOU, V. G. (1991) Nonlinear multivariate calibration using principal components regression and artificial neural networks. *Analytical Chemistry,* 63**,** 2313-2323.

GEVAERT, K. & VANDEKERCKHOVE, J. (2000) Protein identification methods in proteomics. *Electrophoresis,* 21**,** 1145-1154.

GILLIVER, M. A., BENNETT, M., BEGON, M., HAZEL, S. M. & HART, C. A. (1999) Enterobacteria: Antibiotic resistance found in wild rodents. *Nature,* 401**,** 233-234.

GIORDANO, M., KANSIZ, M., HERAUD, P., BEARDALL, J., WOOD, B. & MCNAUGHTON, D. (2001) Fourier transform infrared spectroscopy as a novel

tool to investigate changes in intracellular macromolecular pools in the marine microalga *Chaetoceros muellerii* (Bacillariophyceae). *Journal of Phycology,* 37**,** 271-279.

GOODACRE, R. (2004) Metabolic profiling: pathways in discovery. *Drug Discovery Today,* 9**,** 260-261.

GOODACRE, R., BROADHURST, D., SMILDE, A., KRISTAL, B., BAKER, J., BEGER, R., BESSANT, C., CONNOR, S., CAPUANI, G., CRAIG, A., EBBELS, T., KELL, D., MANETTI, C., NEWTON, J., PATERNOSTRO, G., SOMORJAI, R., SJÖSTRÖM, M., TRYGG, J. & WULFERT, F. (2007) Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics,* 3**,** 231-241.

GOODACRE, R., TIMMINS, E. M., BURTON, R., KADERBHAI, N., WOODWARD, A. M., KELL, D. B. & ROONEY, P. J. (1998) Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks. *Microbiology,* 144**,** 1157-1170.

GOODACRE, R., VAIDYANATHAN, S., DUNN, W. B., HARRIGAN, G. G. & KELL, D. B. (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology,* 22**,** 245-252.

GOURAUD, H. (1971) Continuous Shading of Curved Surfaces. *IEEE Transactions on Computers,* C-20**,** 623-629.

GRIFFITHS, P. R. & DE HASETH, J. A. (2007) *Fourier Transform Infrared Spectrometry,* Hoboken, New Jersey, John Wiley & Sons.

HALL, R. D. (2006) Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist,* 169**,** 453-468.

HALLING-SØRENSEN, B., NORS NIELSEN, S., LANZKY, P. F., INGERSLEV, F., HOLTEN LÜTZHØFT, H. C. & JØRGENSEN, S. E. (1998) Occurrence, fate and effects of pharmaceutical substances in the environment- A review. *Chemosphere,* 36**,** 357-393.

HANDL, J., KNOWLES, J. & KELL, D. B. (2005) Computational cluster validation in post-genomic data analysis. *Bioinformatics,* 21**,** 3201-3212.

HANNA, G. M. & EVANS, F. E. (2000) Optimization of enantiomeric separation for quantitative determination of the chiral drug propranolol by 1H-NMR spectroscopy utilizing a chiral solvating agent. *Journal of Pharmaceutical and Biomedical Analysis,* 24**,** 189-196.

HARRIGAN, G. G. & GOODACRE, R. (2003) *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis,* London, Kluwer Academic Publishers.

HARRIGAN, G. G., LAPLANTE, R. H., COSMA, G. N., COCKERELL, G., GOODACRE, R., MADDOX, J. F., LUYENDYK, J. P., GANEY, P. E. & ROTH, R. A. (2004) Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity. *Toxicology Letters,* 146**,** 197-205.

HARVELL, C. D., KIM, K., BURKHOLDER, J. M., COLWELL, R. R., EPSTEIN, P. R., GRIMES, D. J., HOFMANN, E. E., LIPP, E. K., OSTERHAUS, A. D., NBSP, M, NBSP, E, OVERSTREET, R. M., PORTER, J. W., SMITH, G. W. & VASTA, G. R. (1999) Emerging Marine Diseases--Climate Links and Anthropogenic Factors. *Science,* 285**,** 1505-1510.

HARWOOD, C. S. & GIBSON, J. (1988) Anaerobic and aerobic metabolism of diverse aromatic compounds by the photosynthetic bacterium *Rhodopseudomonas palustris*. *Applied and Environmental Microbiology,* 54**,** 712-717.

HEBERER, T. (2002) Occurrence, fate, and removal of pharmaceutical residues in the aquatic environment: a review of recent research data. *Toxicology Letters,* 131**,** 5-17.

HELM, D. & NAUMANN, D. (1995) Identification of some bacterial cell components by FT-IR spectroscopy. *FEMS Microbiology Letters,* 126**,** 75-79.

HERAUD, P., CAINE, S., SANSON, G., GLEADOW, R., WOOD, B. R. & MCNAUGHTON, D. (2007) Focal plane array infrared imaging: a new way to analyse leaf tissue. *New Phytologist,* 173**,** 216-225.

HERAUD, P., WOOD, B. R., TOBIN, M. J., BEARDALL, J. & MCNAUGHTON, D. (2005) Mapping of nutrient-induced biochemical changes in living algal cells using synchrotron infrared microspectroscopy. *FEMS Microbiology Letters,* 249**,** 219-225.

HIMMELREICH, U., SOMORJAI, R. L., DOLENKO, B., LEE, O. C., DANIEL, H.-M., MURRAY, R., MOUNTFORD, C. E. & SORRELL, T. C. (2003) Rapid Identification of *Candida* Species by Using Nuclear Magnetic Resonance Spectroscopy and a Statistical Classification Strategy. *Applied and Environmental Microbiology,* 69**,** 4566-4574.

HOOD, T. E., CALABRESE, E. J. & ZUCKERMAN, B. M. (2000) Detection of an Estrogen Receptor in Two Nematode Species and Inhibition of Binding and Development by Environmental Chemicals. *Ecotoxicology and Environmental Safety,* 47**,** 74-81.

HORI, T. S. F., AVILEZ, I. M., INOUE, L. K. & MORAES, G. (2006) Metabolical changes induced by chronic phenol exposure in matrinxã *Brycon cephalus*

(teleostei: characidae) juveniles. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology,* 143**,** 67-72.

HÖSKULDSSON, A. (1995) A combined theory for PCA and PLS. *Journal of Chemometrics,* 9**,** 91-123.

HUANG, W., GOODACRE, R., ELLIOTT, G., BECKMANN, M., WORGAN, H., BAILEY, M., WILLIAMS, P., SCULLION, J. & DRAPER, J. (2005) The use of chemical profiling for monitoring metabolic changes in artificial soil slurries caused by horizontal gene transfer. *Metabolomics,* 1**,** 305-315.

HUANG, W. E., HOPPER, D., GOODACRE, R., BECKMANN, M., SINGER, A. & DRAPER, J. (2006) Rapid characterization of microbial biodegradation pathways by FT-IR spectroscopy. *Journal of Microbiological Methods,* 67**,** 273-280.

HUBER, M. M., CANONICA, S., PARK, G.-Y. & VON GUNTEN, U. (2003) Oxidation of Pharmaceuticals during Ozonation and Advanced Oxidation Processes. *Environmental Science & Technology,* 37**,** 1016-1024.

HUGGETT, D. B., BROOKS, B. W., PETERSON, B., FORAN, C. M. & SCHLENK, D. (2002) Toxicity of Select Beta Adrenergic Receptor-Blocking Pharmaceuticals (B-Blockers) on Aquatic Organisms. *Archives of Environmental Contamination and Toxicology,* 43**,** 229-235.

HUNG, C.-H. & PAVLOSTATHIS, S. G. (1997) Aerobic biodegradation of thiocyanate. *Water Research,* 31**,** 2761-2770.

HUTCHINSON, T. H., FIELD, M. D. R. & MANNING, M. J. (1999) Evaluation of immune function in juvenile turbot *Scophthalmus maximus* (L.) exposed to sediments contaminated with polychlorinated biphenyls. *Fish & Shellfish Immunology,* 9**,** 457-472.

HUTCHINSON, T. H., FIELD, M. D. R. & MANNING, M. J. (2003) Evaluation of non-specific immune functions in dab, *Limanda limanda* L., following short-term exposure to sediments contaminated with polyaromatic hydrocarbons and/or polychlorinated biphenyls. *Marine Environmental Research,* 55**,** 193-202.

IANNACCONE, P. M. (2001) Toxicogenomics: "The Call of the Wild Chip". *Environmental Health Perspectives,* 109**,** A8–A11.

JARVIS, R. M. & GOODACRE, R. (2004a) Rapid Discrimination of Bacteria Using Surface-Enhanced Raman Spectroscopy. *Analytical Chemistry,* 76**,** 40-47.

JARVIS, R. M. & GOODACRE, R. (2004b) Ultra-violet resonance Raman spectroscopy for the rapid discrimination of urinary tract infection bacteria. *FEMS Microbiology Letters,* 232**,** 127-132.

JEONG, Y.-S. & CHUNG, J. S. (2006) Biodegradation of thiocyanate in biofilm reactor using fluidized-carriers. *Process Biochemistry,* 41**,** 701-707.

JJEMBA, P. K. (2006) Excretion and ecotoxicity of pharmaceutical and personal care products in the environment. *Ecotoxicology and Environmental Safety,* 63**,** 113-130.

JOHN, D. M. & WHITE, G. F. (1998) Mechanism for Biotransformation of Nonylphenol Polyethoxylates to Xenoestrogens in *Pseudomonas putida*. *Journal of Bacteriology,* 180**,** 4332-4338.

JOHNSON, H. E., BROADHURST, D., GOODACRE, R. & SMITH, A. R. (2003) Metabolic fingerprinting of salt-stressed tomatoes. *Phytochemistry,* 62**,** 919-928.

JOHNSON, H. E., BROADHURST, D., KELL, D. B., THEODOROU, M. K., MERRY, R. J. & GRIFFITH, G. W. (2004) High-throughput metabolic fingerprinting of legume silage fermentations via Fourier transform infrared spectroscopy and chemometrics. *Applied and Environmental Microbiology,* 70**,** 1583-1592.

JOLLIFFE, I. T. (1982) A Note on the Use of Principal Components in Regression. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 31**,** 300-303.

JOLLIFFE, I. T. (1986) *Principal Component Analysis,* New York, Springer-Velarg.

JONES, O. A. H., VOULVOULIS, N. & LESTER, J. N. (2002) Aquatic environmental assessment of the top 25 English prescription pharmaceuticals. *Water Research,* 36**,** 5013-5022.

KADERBHAI, N. N., BROADHURST, D. I., ELLIS, D. I., GOODACRE, R. & KELL, D. B. (2003) Functional Genomics Via Metabolic Footprinting: Monitoring Metabolite Secretion by Escherichia Coli Tryptophan Metabolism Mutants Using FT–IR and Direct Injection Electrospray Mass Spectrometry. *Comparative and Functional Genomics,* 4**,** 376-391.

KAHNG, H.-Y., MALINVERNI, J. C., MAJKO, M. M. & KUKOR, J. J. (2001) Genetic and Functional Analysis of the tbc Operons for Catabolism of Alkyl- and Chloroaromatic Compounds in *Burkholderia* sp. Strain JS150. *Applied and Environmental Microbiology,* 67**,** 4805-4816.

KELL, D. B. (2004) Metabolomics and systems biology: making sense of the soup. *Current Opinion in Microbiology,* 7**,** 296-307.

KELL, D. B. (2006) Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discovery Today,* 11**,** 1085-1092.

KEUN, H. C., EBBELS, T. M. D., BOLLARD, M. E., BECKONERT, O., ANTTI, H., HOLMES, E., LINDON, J. C. & NICHOLSON, J. K. (2004) Geometric Trajectory

Analysis of Metabolic Responses To Toxicity Can Define Treatment Specific Profiles. *Chemical Research in Toxicology,* 17**,** 579-587.

KHACHATOURIANS, G. G. (1998) Agricultural use of antibiotics and the evolution and transfer of antibiotic-resistant bacteria. *Canadian Medical Association Journal,* 159**,** 1129-1136.

KIDDER, L. H., KALASINSKY, V. F., LUKE, J. L., LEVIN, I. W. & LEWIS, E. N. (1997) Visualization of silicone gel in human breast tissue using new infrared imaging spectroscopy. *Nature Medicine,* 3**,** 235-237.

KIME, D. E. & NASH, J. P. (1999) Gamete viability as an indicator of reproductive endocrine disruption in fish. *The Science of The Total Environment,* 233**,** 123-129.

KING, R. J., SHORT, K. A. & SEIDLER, R. J. (1991) Assay for detection and enumeration of genetically engineered microorganisms which is based on the activity of a deregulated 2,4-dichlorophenoxyacetate monooxygenase. *Applied and Environmental Microbiology,* 57**,** 1790-1792.

KLOEPFER, A., GNIRSS, R., JEKEL, M. & REEMTSMA, T. (2004) Occurrence of benzothiazoles in municipal wastewater and their fate in biological treatment. *Water Science & Technology,* 50**,** 203–208.

KOLPIN, D. W., FURLONG, E. T., MEYER, M. T., THURMAN, E. M., ZAUGG, S. D., BARBER, L. B. & BUXTON, H. T. (2002) Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999-2000: A National Reconnaissance. *Environmental Science & Technology,* 36**,** 1202-1211.

KOPKA, J., FERNIE, A., WECKWERTH, W., GIBON, Y. & STITT, M. (2004) Metabolite profiling in plant biology: platforms and destinations. *Genome Biology,* 5**,** 109.

KOPKA, J., SCHAUER, N., KRUEGER, S., BIRKEMEYER, C., USADEL, B., BERGMULLER, E., DORMANN, P., WECKWERTH, W., GIBON, Y., STITT, M., WILLMITZER, L., FERNIE, A. R. & STEINHAUSER, D. (2005) GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics,* 21**,** 1635-1638.

KOWALEWSKA, G. (1999) Phytoplankton - The main factor responsible for transport of polynuclear aromatic hydrocarbons from water to sediments in the Southern Baltic ecosystem. *ICES Journal of Marine Science,* 56**,** 219-222.

KRAFFT, C., SOBOTTKA, S., GEIGER, K., SCHACKERT, G. & SALZER, R. (2007) Classification of malignant gliomas by infrared spectroscopic imaging and linear discriminant analysis. *Analytical and Bioanalytical Chemistry,* 387**,** 1669-1677.

KRUSKAL, W. H. & WALLIS, W. A. (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association,* 47**,** 583-621.

KRZANOWSKI, W. J. (1988) *Principles of multivariate analysis: a user's perspective,* New York, USA, Oxford University Press.

KÜHN, R., PATTARD, M., PERNAK, K.-D. & WINTER, A. (1989) Results of the harmful effects of selected water pollutants (anilines, phenols, aliphatic compounds) to Daphnia magna. *Water Research,* 23**,** 495-499.

KÜMMERER, K. (2009) The presence of pharmaceuticals in the environment due to human use - present knowledge and future challenges. *Journal of Environmental Management,* 90**,** 2354-2366.

KÜMMERER, K., AL-AHMAD, A. & MERSCH-SUNDERMANN, V. (2000) Biodegradability of some antibiotics, elimination of the genotoxicity and affection of wastewater bacteria in a simple test. *Chemosphere,* 40**,** 701-710.

KUKOR, J. J. & OLSEN, R. H. (1990) Diversity of toluene degradation following long term exposure to BTEX *in situ.* IN KAMELY, D., CHAKRABARTY, A. & OMENN, G. S. (Eds.) *Biotechnology and biodegradation.* Houston, Gulf Publishing Co.

KURT, S., GABRIELE, E., GERALD, Z., WERNER, K. & WOLFGANG, L. (1993) Stereoselective features of (R)- and (S)-atenolol: Clinical pharmacological, pharmacokinetic, and radioligand binding studies. *Chirality,* 5**,** 15-19.

LACERDA, C. M. R., CHOE, L. H. & REARDON, K. F. (2007) Metaproteomic Analysis of a Bacterial Community Response to Cadmium Exposure. *Journal of Proteome Research,* 6**,** 1145-1152.

LACERDA, C. M. R. & REARDON, K. F. (2009) Environmental proteomics: applications of proteome profiling in environmental microbiology and biotechnology. *Briefings in Functional Genomics and Proteomics,* 8**,** 75-87.

LAHNER, B., GONG, J., MAHMOUDIAN, M., SMITH, E. L., ABID, K. B., ROGERS, E. E., GUERINOT, M. L., HARPER, J. F., WARD, J. M., MCINTYRE, L., SCHROEDER, J. I. & SALT, D. E. (2003) Genomic scale profiling of nutrient and trace elements in *Arabidopsis thaliana*. *Nature Biotechnology,* 21**,** 1215-1221.

LANE, A. N., FAN, T. W. M., HIGASHI, R. M., TAN, J., BOUSAMRA, M. & MILLER, D. M. (2009) Prospects for clinical cancer metabolomics using stable isotope tracers. *Experimental and Molecular Pathology,* 86**,** 165-173.

LANGSLET, A. (1970) Membrane stabilization and cardiac effects of d,1-propranolol, d-propranolol and chlorpromazine. *European Journal of Pharmacology,* 13**,** 6-14.

LARSSON, D. G. J., ADOLFSSON-ERICI, M., PARKKONEN, J., PETTERSSON, M., BERG, A. H., OLSSON, P. E. & FÖRLIN, L. (1999) Ethinyloestradiol -- an undesired fish contraceptive? *Aquatic Toxicology,* 45**,** 91-97.

LARSSON, D. G. J., DE PEDRO, C. & PAXEUS, N. (2007) Effluent from drug manufactures contains extremely high levels of pharmaceuticals. *Journal of Hazardous Materials,* 148**,** 751-755.

LASCH, P., BOESE, M., PACIFICO, A. & DIEM, M. (2002) FT-IR spectroscopic investigations of single cells on the subcellular level. *Vibrational Spectroscopy,* 28**,** 147-157.

LASKOWSKI, R. A., HUTCHINSON, E. G., MICHIE, A. D., WALLACE, A. C., JONES, M. L. & THORNTON, J. M. (1997) PDBsum: a web-based database of summaries and analyses of all PDB structures. *Trends in Biochemical Sciences,* 22**,** 488-490.

LEES, P., TAYLOR, P. M., LANDONI, F. M., ARIFAH, A. K. & WATERS, C. (2003) Ketoprofen in the Cat: Pharmacodynamics and Chiral Pharmacokinetics. *The Veterinary Journal,* 165**,** 21-35.

LEI, A. P., WONG, Y. S. & TAM, N. F. Y. (2002) Removal of pyrene by different microalgal species. *Water Science and Technology,* 46**,** 195-201.

LEMMER, B., WINKLER, H., OHM, T. & FINK, M. (1985) Chronopharmacokinetics of beta-receptor blocking drugs of different lipophilicity (propranolol, metoprolol, sotalol, atenolol) in plasma and tissues after single and multiple dosing in the rat. *Naunyn-Schmiedeberg's Archives of Pharmacology,* 330**,** 42-49.

LEONARD, D. & LINDLEY, N. D. (1998) Carbon and energy flux constraints in continuous cultures of *Alcaligenes eutrophus* grown on phenol. *Microbiology,* 144**,** 241-248.

LEWIS, E. N., GORBACH, A. M., MARCOTT, C. & LEVIN, I. W. (1996) High-Fidelity Fourier Transform Infrared Spectroscopic Imaging of Primate Brain Tissue. *Applied Spectroscopy,* 50**,** 263-269.

LEWIS, E. N., TREADO, P. J., REEDER, R. C., STORY, G. M., DOWREY, A. E., MARCOTT, C. & LEVIN, I. W. (1995) Fourier Transform Spectroscopic Imaging Using an Infrared Focal-Plane Array Detector. *Analytical Chemistry,* 67**,** 3377-3381.

LILLIEFORS, H. W. (1967) On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association,* 62**,** 399-402.

LINDON, J. C., HOLMES, E. & NICHOLSON, J. K. (2001a) Pattern recognition methods and applications in biomedical magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy,* 39**,** 1-40.

LINDON, J. C., HOLMES, E. & NICHOLSON, J. K. (2001b) Pattern recognition methods and applications in biomedical magnetic resonance. *Progress in Nuclear Magnetic Resonance Spectroscopy,* 39**,** 1-40.

LINDON, J. C., NICHOLSON, J. K., HOLMES, E., ANTTI, H., BOLLARD, M. E., KEUN, H., BECKONERT, O., EBBELS, T. M., REILY, M. D., ROBERTSON, D., STEVENS, G. J., LUKE, P., BREAU, A. P., CANTOR, G. H., BIBLE, R. H., NIEDERHAUSER, U., SENN, H., SCHLOTTERBECK, G., SIDELMANN, U. G., LAURSEN, S. M., TYMIAK, A., CAR, B. D., LEHMAN-MCKEEMAN, L., COLET, J.-M., LOUKACI, A. & THOMAS, C. (2003) Contemporary issues in toxicology the role of metabonomics in toxicology and its evaluation by the COMET project. *Toxicology and Applied Pharmacology,* 187**,** 137-146.

LOPEZ-DIEZ, E. C., WINDER, C. L., ASHTON, L., CURRIE, F. & GOODACRE, R. (2005) Monitoring the Mode of Action of Antibiotics Using Raman Spectroscopy: Investigating Subinhibitory Effects of Amikacin on *Pseudomonas aeruginosa*. *Analytical Chemistry,* 77**,** 2901-2906.

LORIN-LATXAGUE, C. & MELIN, A.-M. (2005) Radical induced damage of *Micrococcus luteus* bacteria monitored using FT-IR spectroscopy. *Spectroscopy,* 19**,** 17-26.

LOVETT, R. A. (2000) TOXICOGENOMICS: Toxicologists Brace for Genomics Revolution. *Science,* 289**,** 536-537.

LU, Y., YAN, L., WANG, Y., ZHOU, S., FU, J. & ZHANG, J. (2008) Biodegradation of phenolic compounds from coking wastewater by immobilized white rot fungus *Phanerochaete chrysosporium*. *Journal of Hazardous Materials,* 165**,** 1091-1097.

MACFIE, H. J. H., GUTTERIDGE, C. S. & NORRIS, J. R. (1978) Use of Canonical Variates Analysis in Differentiation of Bacteria by Pyrolysis Gas-Liquid Chromatography. *Journal of General Microbiology,* 104**,** 67-74.

MACKENZIE, D. A., DEFERNEZ, M., DUNN, W. B., BROWN, M., FULLER, L. J., SECO DE HERRERA, S. R. M., GÜNTHER, A., JAMES, S. A., EAGLES, J., PHILO, M., GOODACRE, R. & ROBERTS, I. N. (2008) Relatedness of

medically important strains of *Saccharomyces cerevisiae* as revealed by phylogenetics and metabolomics. *Yeast,* 25**,** 501-512.

MADSEN, E. L. (2006) The use of stable isotope probing techniques in bioreactor and field studies on bioremediation. *Current Opinion in Biotechnology,* 17**,** 92-97.

MAGEE, J. T. (1993) Whole-organism fingerprinting. IN GOODFELLOW, M. & O'DONNELL, A. G. (Eds.) *Handbook of New Bacterial Systematics.* London, Academic Press.

MANEFIELD, M., GRIFFITHS, R. I., LEIGH, M. B., FISHER, R. & WHITELEY, A. S. (2005) Functional and compositional comparison of two activated sludge communities remediating coking effluent. *Environmental Microbiology,* 7**,** 715-722.

MANEFIELD, M., WHITELEY, A. S., GRIFFITHS, R. I. & BAILEY, M. J. (2002) RNA Stable Isotope Probing, a Novel Means of Linking Microbial Community Function to Phylogeny. *Applied and Environmental Microbiology,* 68**,** 5367-5373.

MANLY, B. F. J. (1994) *Multivariate Statistical Methods: A Primer,* London, Chapman and Hall.

MANN, M., HENDRICKSON, R. C. & PANDEY, A. (2001) Analysis Of Proteins And Proteomes By Mass Spectrometry. *Annual Review of Biochemistry,* 70**,** 437-473.

MAQUELIN, K., KIRSCHNER, C., CHOO-SMITH, L. P., VAN DEN BRAAK, N., ENDTZ, H. P., NAUMANN, D. & PUPPELS, G. J. (2002) Identification of medically relevant microorganisms by vibrational spectroscopy. *Journal of Microbiological Methods,* 51**,** 255-271.

MAROTI, G., TONG, Y., YOOSEPH, S., BADEN-TILLSON, H., SMITH, H. O., KOVACS, K. L., FRAZIER, M., VENTER, J. C. & XU, Q. (2009) Discovery of [NiFe] Hydrogenase Genes in Metagenomic DNA: Cloning and Heterologous Expression in *Thiocapsa roseopersicina*. *Applied and Environmental Microbiology,* 75**,** 5821-5830.

MARTENS, H. & NAES, T. (1989) *Multivariate calibration,* Chichester, John Wiley & Sons.

MARTENS, H., NIELSEN, J. P. & ENGELSEN, S. B. (2003) Light Scattering and Light Absorbance Separated by Extended Multiplicative Signal Correction. Application to Near-Infrared Transmission Analysis of Powder Mixtures. *Analytical Chemistry,* 75**,** 394-404.

MARTIN, R. W. (1949) Rapid Colorimetric Estimation of Phenol. *Analytical Chemistry,* 21**,** 1419-1420.

MARTÍNEZ-BUENO, M. A., TOBES, R., REY, M. & RAMOS, J.-L. (2002) Detection of multiple extracytoplasmic function (ECF) sigma factors in the genome of *Pseudomonas putida* KT2440 and their counterparts in *Pseudomonas aeruginosa* PA01. *Environmental Microbiology,* 4**,** 842-855.

MATSUMOTO, I. & KUHARA, T. (1996) A new chemical diagnostic method for inborn errors of metabolism by mass spectrometry - rapid, practical, and simultaneous urinary metabolites analysis. *Mass Spectrometry Reviews,* 15**,** 43-57.

MCGOVERN, A., C., BROADHURST, D., TAYLOR, J., KADERBHAI, N., WINSON, M., K., SMALL, D., A., ROWLAND, J., J., KELL, D., B. & GOODACRE, R. (2002) Monitoring of complex industrial bioprocesses for metabolite concentrations using modern spectroscopies and machine learning: Application to gibberellic acid production. *Biotechnology and Bioengineering,* 78**,** 527-538.

MCGOVERN, A. C., ERNILL, R., KARA, B. V., KELL, D. B. & GOODACRE, R. (1999) Rapid analysis of the expression of heterologous proteins in *Escherichia coli* using pyrolysis mass spectrometry and Fourier transform infrared spectroscopy with chemometrics: application to [alpha]2-interferon production. *Journal of Biotechnology,* 72**,** 157-168.

MCINTOSH, L., M., MANSFIELD, J., R., CROWSON, N. A., MANTSCH, H., H. & JACKSON, M. (1999) Analysis and interpretation of infrared microscopic maps: Visualization and classification of skin components by digital staining and multivariate analysis. *Biospectroscopy,* 5**,** 265-275.

MEHVAR, R. & BROCKS, D. R. (2001) Stereospecific pharmacokinetics and pharmacodynamics of beta-adrenergic blockers in humans. *Journal of Pharmacy and Pharmaceutical Sciences,* 4**,** 185-200.

MEINDL, U. (1993) *Micrasterias* Cells as a Model System for Research on Morphogenesis. *Microbiological Reviews,* 57**,** 415-433.

MEINDL, U., ZHANG, D. & HEPLER, P. K. (1994) Actin microfilaments are associated with the migrating nucleus and the cell cortex in the green alga Micrasterias. Studies on living cells. *Journal of Cell Science,* 107**,** 1929-1934.

MINHALMA, M. & DE PINHO, M. N. (2002) Development of nanofiltration/steam stripping sequence for coke plant wastewater treatment. *Desalination,* 149**,** 95-100.

MINHALMA, M. & DE PINHO, M. N. (2004) Integration of nanofiltration/steam stripping for the treatment of coke plant ammoniacal wastewaters. *Journal of Membrane Science,* 242**,** 87-95.

MONSINJON, T. & KNIGGE, T. (2007) Proteomic applications in ecotoxicology. *Proteomics,* 7**,** 2997-3009.

MOORE, M. N. (2002) Biocomplexity: the post-genome challenge in ecotoxicology. *Aquatic Toxicology,* 59**,** 1-15.

MORPER, M. R. & WILDMOSER, A. (1990) Improvement of Existing Wastewater Treatment Plants' Efficiencies Without Enlargement of Tankage By Application of the Linpor-Process--Case Studies. *Water Science & Technology,* 22**,** 207-215.

MOUILLE, G., ROBIN, S., LECOMTE, M., PAGANT, S. & HÖFTE, H. (2003) Classification and identification of *Arabidopsis* cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy. *The Plant Journal,* 35**,** 393-404.

MURRAY, L. A., RAAB, A., MARR, I. L. & FELDMANN, J. (2003) Biotransformation of arsenate to arsenosugars by *Chlorella vulgaris*. *Applied Organometallic Chemistry,* 17**,** 669-674.

NAHAR, N., ALAUDDIN, M. & QUILTY, B. (2000) Toxic effects of toluene on the growth of activated sludge bacteria. *World Journal of Microbiology and Biotechnology,* 16**,** 307-311.

NAUMANN, D., HELM, D. & LABISCHINSKI, H. (1991a) Microbiological characterizations by FT-IR spectroscopy. *Nature,* 351**,** 81-82.

NAUMANN, D., HELM, D., LABISCHINSKI, H. & GIESBRECHT, P. (1991b) The Characterization of Microorganisms by Fourier-Transform Infrared Spectroscopy (FT-IR). IN NELSON, W. H. (Ed.) *Modern Techniques for Rapid Microbiological Analysis.* New York, VCH Publishers.

NAUMANN, D., HELM, D. & SCHULTZ, C. (1994) Characterisation and identification of micro-organisms by FT-IR spectroscopy and FT-IR microscopy. IN PRIEST, F. G., RAMOS-CORMENZANA, A. & TINDALL, B. J. (Eds.) *Bacterial diversity and systematics.* New York, Plenum Press.

NAUMANN, D., SHULTZ, C. P. & HELM, D. (1996) What Can Infrared Spectroscopy Tell Us About the Structure and Composition of Intact Bacterial Cells? IN MANTSCH, H. H. & CHAPMAN, D. (Eds.) *Infrared Spectroscopy of Biomolecules.* 1 ed. London, Wiley-Liss.

NAYLER, W. G., FERRARI, R. & WILLIAMS, A. (1980) Protective effect of pretreatment with verapamil, nifedipine and propranolol on mitochondrial

function in the ischemic and reperfused myocardium. *The American Journal of Cardiology,* 46**,** 242-248.

NELSON, K. E., WEINEL, C., PAULSEN, I. T., DODSON, R. J., HILBERT, H., MARTINS DOS SANTOS, V. A. P., FOUTS, D. E., GILL, S. R., POP, M., HOLMES, M., BRINKAC, L., BEANAN, M., DEBOY, R. T., DAUGHERTY, S., KOLONAY, J., MADUPU, R., NELSON, W., WHITE, O., PETERSON, J., KHOURI, H., HANCE, I., CHRIS LEE, P., HOLTZAPPLE, E., SCANLAN, D., TRAN, K., MOAZZEZ, A., UTTERBACK, T., RIZZO, M., LEE, K., KOSACK, D., MOESTL, D., WEDLER, H., LAUBER, J., STJEPANDIC, D., HOHEISEL, J., STRAETZ, M., HEIM, S., KIEWITZ, C., EISEN, J., TIMMIS, K. N., DÜSTERHÖFT, A., TÜMMLER, B. & FRASER, C. M. (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environmental Microbiology,* 4**,** 799-808.

NELSON, M. J., MONTGOMERY, S. O., MAHAFFEY, W. R. & PRITCHARD, P. H. (1987) Biodegradation of trichloroethylene and involvement of an aromatic biodegradative pathway. *Applied and Environmental Microbiology,* 53**,** 949-954.

NEUFELD, R. D. & VALIKNAC, T. (1979) Inhibition of Phenol Biodegradation by Thiocyanate. *Research Journal of the Water Pollution Control Federation,* 51**,** 2283-2291.

NICHOLLS, A. W., HOLMES, E., LINDON, J. C., SHOCKCOR, J. P., FARRANT, R. D., HASELDEN, J. N., DAMMENT, S. J. P., WATERFIELD, C. J. & NICHOLSON, J. K. (2001) Metabonomic Investigations into Hydrazine Toxicity in the Rat. *Chemical Research in Toxicology,* 14**,** 975-987.

NICHOLSON, J. K., LINDON, J. C. & HOLMES, E. (1999) 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica,* 29**,** 1181-1189.

NIKOLAI, L. N., MCCLURE, E. L., MACLEOD, S. L. & WONG, C. S. (2006) Stereoisomer quantification of the β-blocker drugs atenolol, metoprolol, and propranolol in wastewaters by chiral high-performance liquid chromatography-tandem mass spectrometry. *Journal of Chromatography A,* 1131**,** 103-109.

NORDLUND, I., POWLOWSKI, J. & SHINGLER, V. (1990) Complete nucleotide sequence and polypeptide analysis of multicomponent phenol hydroxylase from Pseudomonas sp. strain CF600. *Journal of Bacteriology,* 172**,** 6826-6833.

NUWAYSIR, E. F., BITTNER, M., TRENT, J., BARRETT, J. C. & AFSHARI, C. A. (1999) Microarrays and toxicology: The advent of toxicogenomics. *Molecular Carcinogenesis,* 24**,** 153-159.

O'HAGAN, S., DUNN, W. B., BROWN, M., KNOWLES, J. D. & KELL, D. B. (2004) Closed-Loop, Multiobjective Optimization of Analytical Instrumentation: Gas Chromatography/Time-of-Flight Mass Spectrometry of the Metabolomes of Human Serum and of Yeast Fermentations. *Analytical Chemistry,* 77**,** 290-303.

OH, S. J. & KOENIG, J. L. (1998) Phase and Curing Behavior of Polybutadiene/Diallyl Phthalate Blends Monitored by FT-IR Imaging Using Focal-Plane Array Detection. *Analytical Chemistry,* 70**,** 1768-1772.

OKAY, O. S., DONKIN, P., PETERS, L. D. & LIVINGSTONE, D. R. (2000) The role of algae (*Isochrysis galbana*) enrichment on the bioaccumulation of benzo[a]pyrene and its effects on the blue mussel *Mytilus edulis*. *Environmental Pollution,* 110**,** 103-113.

OLIVER, S. G., WINSON, M. K., KELL, D. B. & BAGANZ, F. (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology,* 16**,** 373-378.

PALLER, G., HOMMEL, R., K. & KLEBER, H.-P. (1995) Phenol degradation by *Acinetobacter calcoaceticus* NCIB 8250. *Journal of Basic Microbiology,* 35**,** 325-335.

PARUCHURI, Y. L., SHIVARAMAN, N. & KUMARAN, P. (1990) Microbial transformation of thiocyanate. *Environmental Pollution,* 68**,** 15-28.

PASCOE, D., KARNTANUT, W. & MÜLLER, C. T. (2003) Do pharmaceuticals affect freshwater invertebrates? A study with the cnidarian Hydra vulgaris. *Chemosphere,* 51**,** 521-528.

PATANJALI, S. R., PARIMOO, S. & WEISSMAN, S. M. (1991) Construction of a uniform-abundance (normalized) cDNA library. *Proceedings of the National Academy of Sciences of the United States of America,* 88**,** 1943-1947.

PATEL, S. A., CURRIE, F., THAKKER, N. & GOODACRE, R. (2008) Spatial metabolic fingerprinting using FT-IR spectroscopy: investigating abiotic stresses on *Micrasterias hardyi*. *The Analyst,* 133**,** 1707-1713.

PATTERSON, J. W. (1985) *Industrial wastewater treatment technology,* Stoneham, Butterworth Publishers.

PEARSON, A. A., GAFFNEY, T. E., WALLE, T. & PRIVITERA, P. J. (1989) A stereoselective central hypotensive action of atenolol. *Journal of Pharmacology and Experimental Therapeutics,* 250**,** 759-763.

PENNIE, W. D., TUGWOOD, J. D., OLIVER, G. J. A. & KIMBER, I. (2000) The Principles and Practice of Toxicogenomics: Applications and Opportunities. *Toxicological Sciences,* 54**,** 277-283.

PEOPLES, M. E., SMITH, M. J. & PALMER, R. A. (1987) FT-IR Photothermal Beam Deflection Spectroscopy (PBDS) Studies of Polymer-Modified Graphite Electrodes. *Applied Spectroscopy,* 41**,** 1257-1259.

PETRICH, W., DOLENKO, B., FRÜH, J., GANZ, M., GREGER, H., JACOB, S., KELLER, F., NIKULIN, A. E., OTTO, M., QUARDER, O., SOMORJAI, R. L., STAIB, A., WERNER, G. & WIELINGER, H. (2000) Disease Pattern Recognition in Infrared Spectra of Human Sera with Diabetes Mellitus as an Example. *Applied Optics,* 39**,** 3372-3379.

PHILP, J. C., BALMAND, S., HAJTO, E., BAILEY, M. J., WILES, S., WHITELEY, A. S., LILLEY, A. K., HAJTO, J. & DUNBAR, S. A. (2003) Whole cell immobilised biosensors for toxicity assessment of a wastewater treatment plant treating phenolics-containing waste. *Analytica Chimica Acta,* 487**,** 61-74.

PICARD, R. R. & BERK, K. N. (1990) Data Splitting. *The American Statistician,* 44**,** 140-147.

PICARD, R. R. & COOK, R. D. (1984) Cross-Validation of Regression Models. *Journal of the American Statistical Association,* 79**,** 575-583.

POTTER, L. T. & SWEETLAND, C. B. (1967) Uptake of Propranolol by Isolated Guinea-pig Atria. *Journal of Pharmacology and Experimental Therapeutics,* 155**,** 91-100.

POULIQUEN, H., LE BRIS, H. & PINAULT, L. (1992) Experimental study of the therapeutic application of oxytetracycline, its attenuation in sediment and sea water, and implications for farm culture of benthic organisms. *Marine Ecology Progress Series,* 89**,** 93-98.

POYNTON, H. C. & VULPE, C. D. (2009) Ecotoxicogenomics: Emerging technologies for emerging contaminants. *Journal of the American Water Resources Association,* 45**,** 83-96.

POYNTON, H. C., WINTZ, H. & VULPE, C. D. (2008) Progress in ecotoxicogenomics for environmental monitoring, mode of action, and toxicant identification. *Advances in Experimental Biology.*

PRATER, B. E. & FISHER, R. (1991) Recent developments in the treatment of coke oven wastewaters. *ENCOSTEEL World Conference: Environmental Control in the Steel Industry.* Verlag Stahleisen.

PURDOM, C. E., HARDIMAN, P. A., BYE, V. V. J., ENO, N. C., TYLER, C. R. & SUMPTER, J. P. (1994) Estrogenic Effects of Effluents from Sewage Treatment Works. *Chemistry and Ecology,* 8**,** 275 - 285.

QIAN, Y., WEN, Y. & ZHANG, H. (1994) Efficacy of pre-treatment methods in the activated sludge removal of refractory compounds in coke-plant wastewater. *Water Research,* 28**,** 701-707.

QUINTANA, J. B., WEISS, S. & REEMTSMA, T. (2005) Pathways and metabolites of microbial degradation of selected acidic pharmaceutical and their occurrence in municipal wastewater treated by a membrane bioreactor. *Water Research,* 39**,** 2654-2664.

RAAMSDONK, L. M. C. (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature biotechnology,* 19**,** 45-50.

RAM, R. J., VERBERKMOES, N. C., THELEN, M. P., TYSON, G. W., BAKER, B. J., BLAKE, R. C., II, SHAH, M., HETTICH, R. L. & BANFIELD, J. F. (2005) Community Proteomics of a Natural Microbial Biofilm. *Science,* 308**,** 1915-1920.

RASHED, M. S. (2001) Clinical applications of tandem mass spectrometry: ten years of diagnosis and screening for inherited metabolic diseases. *Journal of Chromatography B: Biomedical Sciences and Applications,* 758**,** 27-48.

REARDON, K., F., MOSTELLER, D., C. & BULL ROGERS, J., D. (2000) Biodegradation kinetics of benzene, toluene, and phenol as single and mixed substrates for *Pseudomonas putida* F1. *Biotechnology and Bioengineering,* 69**,** 385-400.

REASONER, D. J. & GELDREICH, E. E. (1985) A new medium for the enumeration and subculture of bacteria from potable water. *Applied and Environmental Microbiology,* 49**,** 1-7.

RICE, K., C. & BAYLES, K., W. (2003) Death's toolbox: examining the molecular components of bacterial programmed cell death. *Molecular Microbiology,* 50**,** 729-738.

RICKARD, A. H., LEACH, S. A., BUSWELL, C. M., HIGH, N. J. & HANDLEY, P. S. (2000) Coaggregation between Aquatic Bacteria Is Mediated by Specific-Growth-Phase-Dependent Lectin-Saccharide Interactions. *Applied and Environmental Microbiology,* 66**,** 431-434.

ROBERTS, P. H. & THOMAS, K. V. (2006) The occurrence of selected pharmaceuticals in wastewater effluent and surface waters of the lower Tyne catchment. *Science of The Total Environment,* 356**,** 143-153.

ROCKETT, J. C. & DIX, D. J. (1999) Application of DNA arrays to toxicology. *Environmental Health Perspectives,* 107**,** 681-685.

ROCKETT, J. C. & DIX, D. J. (2000) DNA arrays: technology, options and toxicological applications. *Xenobiotica,* 30**,** 155-177.

ROESSNER, U., WAGNER, C., KOPKA, J., TRETHEWEY, R. N. & WILLMITZER, L. (2000) Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *The Plant Journal,* 23**,** 131-142.

ROGERS, J. B. & REARDON, K., F. (2000) Modeling substrate interactions during the biodegradation of mixtures of toluene and phenol by *Burkholderia* species JS150. *Biotechnology and Bioengineering,* 70**,** 428-435.

ROTH, S. & SEEMAN, P. (1971) All lipid-soluble anaesthetics protect red cells. *Nature: New biology,* 231**,** 284-285.

SAHA, N. C., BHUNIA, F. & KAVIRAJ, A. (1999) Toxicity of Phenol to Fish and Aquatic Ecosystems. *Bulletin of Environmental Contamination and Toxicology,* 63**,** 195-202.

SAUER, U. (2005) Metabolic flux analysis: A key methodology for systems biology of metabolism. *Systems Biology.*

SCHENA, M., SHALON, D., DAVIS, R. W. & BROWN, P. O. (1995) Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science,* 270**,** 467-470.

SCHLEGEL, H. G. (1993) *General microbiology,* Cambridge, Cambridge University Press.

SCHULTZ, C. P., LIU, K.-Z., SALAMON, E. A., RIESE, K. T. & MANTSCH, H. H. (1999) Application of FT-IR microspectroscopy in diagnosing thyroid neoplasms. *Journal of Molecular Structure,* 480-481**,** 369-377.

SCHULZE, W. X., GLEIXNER, G., KAISER, K., GUGGENBERGER, G., MANN, M. & SCHULZE, E.-D. (2005) A proteomic fingerprint of dissolved organic carbon and of soil particles. *Oecologia,* 142**,** 335-343.

SCHUSTER, K. C., GOODACRE, R., GAPES, J. R. & YOUNG, M. (2001) Degeneration of solventogenic *Clostridium* strains monitored by Fourier transform infrared spectroscopy of bacterial cells. *Journal of Industrial Microbiology and Biotechnology,* 27**,** 314-321.

SECOR, R. M. (1963) Resolution of Optical Isomers by Crystallization Procedures. *Chemical Reviews,* 63**,** 297-309.

SEEMAN, P. M. & CARL, C. P. A. J. R. S. (1966) Membrane Stabilization by Drugs: Tranquilizers, Steroids, and Anesthetics. *International Review of Neurobiology.* Academic Press.

SEILER, J. P. (2002) Pharmacodynamic activity of drugs and ecotoxicology - can the two be connected? *Toxicology Letters,* 131**,** 105-115.

SHELLY, K., HERAUD, P. & BEARDALL, J. (2002) Nitrogen limitation in *Dunaliella tertiolecta* (Chlorophyceae) leads to increased susceptibility to damage by ultraviolet-B radiation but also increased repair capacity. *Journal of Phycology,* 38**,** 713-720.

SHIEH, W. K. & RICHARDS, D. J. (1988) Anoxic/Oxic Activated Sludge Treatment Kinetics of Cyanides and Phenols. *Journal of Environmental Engineering,* 114**,** 639-654.

SHINGLER, V., POWLOWSKI, J. & MARKLUND, U. (1992) Nucleotide sequence and functional analysis of the complete phenol/3,4-dimethylphenol catabolic pathway of Pseudomonas sp. strain CF600. *Journal of Bacteriology,* 174**,** 711-724.

SHUGART, L. R. & THEODORAKIS, C. (1996) Genetic ecotoxicology: The genotypic diversity approach. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology,* 113**,** 273-276.

SINGLETON, I. (1994) Microbial metabolism of xenobiotics: Fundamental and applied research. *Journal of Chemical Technology & Biotechnology,* 59**,** 9-23.

SLY, L. I. & CAHILL, M. M. (1997) Transfer of *Blastobacter natatorius* (Sly 1985) to the Genus *Blastomonas* gen. nov. as *Blastomonas natatoria* comb. nov. *International Journal of Systematic and Evolutionary Microbiology,* 47**,** 566-568.

SMEDSGAARD, J. & NIELSEN, J. (2005) Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *Journal of Experimental Botany,* 56**,** 273-286.

SMITH, K. E., BESSER, J. M., HEDBERG, C. W., LEANO, F. T., BENDER, J. B., WICKLUND, J. H., JOHNSON, B. P., MOORE, K. A., OSTERHOLM, M. T. & THE INVESTIGATION, T. (1999) Quinolone-Resistant Campylobacter jejuni Infections in Minnesota, 1992-1998. *The New England Journal of Medicine,* 340**,** 1525-1532.

SNAPE, J. R., MAUND, S. J., PICKFORD, D. B. & HUTCHINSON, T. H. (2004) Ecotoxicogenomics: the challenge of integrating genomics into aquatic and terrestrial ecotoxicology. *Aquatic Toxicology,* 67**,** 143-154.

SNIVELY, C. M., KATZENBERGER, S., OSKARSDOTTIR, G. & LAUTERBACH, J. (1999) Fourier-transform infrared imaging using a rapid-scan spectrometer. *Optics Letters,* 24**,** 1841-1843.

SOGA, T., OHASHI, Y., UENO, Y., NARAOKA, H., TOMITA, M. & NISHIOKA, T. (2003) Quantitative Metabolome Analysis Using Capillary Electrophoresis Mass Spectrometry. *Journal of Proteome Research,* 2**,** 488-494.

SPAIN, J. C., ZYLSTRA, G. J., BLAKE, C. K. & GIBSON, D. T. (1989) Monohydroxylation of phenol and 2,5-dichlorophenol by toluene dioxygenase in Pseudomonas putida F1. *Applied and Environmental Microbiology,* 55**,** 2648-2652.

STAIB, A., DOLENKO, B., FINK, D. J., FRÜH, J., NIKULIN, A. E., OTTO, M., PESSIN-MINSLEY, M. S., QUARDER, O., SOMORJAI, R., THIENEL, U., WERNER, G. & PETRICH, W. (2001) Disease pattern recognition testing for rheumatoid arthritis using infrared spectra of human serum. *Clinica Chimica Acta,* 308**,** 79-89.

STAIB, C. & LANT, P. (2007) Thiocyanate degradation during activated sludge treatment of coke-ovens wastewater. *Biochemical Engineering Journal,* 34**,** 122-130.

STONE, M. & BROOKS, R. J. (1990) Continuum Regression: Cross-Validated Sequentially Constructed Prediction Embracing Ordinary Least Squares, Partial Least Squares and Principal Components Regression. *Journal of the Royal Statistical Society. Series B (Methodological),* 52**,** 237-269.

STOVER, C. K., PHAM, X. Q., ERWIN, A. L., MIZOGUCHI, S. D., WARRENER, P., HICKEY, M. J., BRINKMAN, F. S. L., HUFNAGLE, W. O., KOWALIK, D. J., LAGROU, M., GARBER, R. L., GOLTRY, L., TOLENTINO, E., WESTBROCK-WADMAN, S., YUAN, Y., BRODY, L. L., COULTER, S. N., FOLGER, K. R., KAS, A., LARBIG, K., LIM, R., SMITH, K., SPENCER, D., WONG, G. K. S., WU, Z., PAULSEN, I. T., REIZER, J., SAIER, M. H., HANCOCK, R. E. W., LORY, S. & OLSON, M. V. (2000) Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature,* 406**,** 959(6).

STUART, B. (1996) *Modern Infrared Spectroscopy*, John Wiley & Sons.

SUMNER, L. W., MENDES, P. & DIXON, R. A. (2003) Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry,* 62**,** 817-836.

SUMPTER, J. P. & JOBLING, S. (1995) Vitellogenesis as a Biomarker for Estrogenic Contamination of the Aquatic Environment. *Environmental Health Perspectives,* 103**,** 173-178.

SUTTON, P. M., HURVID, J. & HOEKSEMA, M. (1999) Biological Fluidized-Bed Treatment of Wastewater from Byproduct Coking Operations: Full-Scale Case History. *Water Environment Research,* 71**,** 5-9.

SUZUKI, T., KANAGAWA, T. & KAMAGATA, Y. (2002) Identification of a Gene Essential for Sheathed Structure Formation in *Sphaerotilus natans*, a Filamentous Sheathed Bacterium. *Applied and Environmental Microbiology,* 68**,** 365-371.

SVOBODOVÁ, D. & GASPARIČ, J. (1971) Investigation of the colour reaction of phenols with 4-aminoantipyrine. *Microchimica Acta,* 59**,** 384-390.

TAN, M.-W., MAHAJAN-MIKLOS, S. & AUSUBEL, F. M. (1999) Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America,* 96**,** 715-720.

TAYSSE, L., TROUTAUD, D., KHAN, N. A. & DESCHAUX, P. (1995) Structure-activity relationship of phenolic compounds (phenol, pyrocatechol and hydroquinone) on natural lymphocytotoxicity of carp (*Cyprinus carpio*). *Toxicology,* 98**,** 207-214.

TERNES, T. A. (1998) Occurrence of drugs in German sewage treatment plants and rivers. *Water Research,* 32**,** 3245-3260.

TERNES, T. A., STÜBER, J., HERRMANN, N., MCDOWELL, D., RIED, A., KAMPMANN, M. & TEISER, B. (2003) Ozonation: a tool for removal of pharmaceuticals, contrast media and musk fragrances from wastewater? *Water Research,* 37**,** 1976-1982.

THACKER, P. D. (2005) Pharmaceutical Data Elude Researchers. *Environmental Science & Technology,* 39**,** 193A-194A.

THEODORSSON-NORHEIM, E. (1987) Friedman and Quade tests: BASIC computer program to perform nonparametric two-way analysis of variance and multiple comparisons on ranks of several related samples. *Computers in Biology and Medicine,* 17**,** 85-99.

THOMAS, G. H. (2001) Metabolomics breaks the silence. *Trends in Microbiology,* 9**,** 158-158.

THOMKINS, J., DEVILLE, M. M., DAY, J. G. & TURNER, M. F. (1995) Culture Collection of Algae and Protozoa, Catalogue of Strains. Ambleside, Natural Environmental Research Council.

THOMPSON, J. D. (1991) Phenotypic plasticity as a component of evolutionary change. *Trends in Ecology & Evolution,* 6**,** 246-249.

TIMMINS, E. M., HOWELL, S. A., ALSBERG, B. K., NOBLE, W. C. & GOODACRE, R. (1998) Rapid Differentiation of Closely Related *Candida* Species and Strains by Pyrolysis-Mass Spectrometry and Fourier Transform-Infrared Spectroscopy. *Journal of Clinical Microbiology,* 36**,** 367-374.

TISLER, T., ZAGORC-KONCAN, J., ROS, M. & COTMAN, M. (1999) Biodegradation and toxicity of wastewater from industry producing mineral fibres for thermal insulation. *Chemosphere,* 38**,** 1347-1352.

TODD, S. J., CAIN, R. B. & SCHMIDT, S. (2002) Biotransformation of naphthalene and diaryl ethers by green microalgae. *Biodegradation,* 13**,** 229-238.

TORRES, M. A., BARROS, M. P., CAMPOS, S. C. G., PINTO, E., RAJAMANI, S., SAYRE, R. T. & COLEPICOLO, P. (2008) Biochemical biomarkers in algae and marine pollution: A review. *Ecotoxicology and Environmental Safety,* 71**,** 1-15.

TREADO, P. J., LEVIN, I. W. & LEWIS, E. N. (1994) Indium Antimonide (InSb) Focal Plane Array (FPA) Detection for Near-Infrared Imaging Microscopy. *Applied Spectroscopy,* 48**,** 607-615.

TSCHECH, A. & FUCHS, G. (1987) Anaerobic degradation of phenol by pure cultures of newly isolated denitrifying pseudomonads. *Archives of Microbiology,* 148**,** 213-217.

UDELHOVEN, T., NAUMANN, D. & SCHMITT, J. (2000) Development of a Hierarchical Classification System with Artificial Neural Networks and FT-IR Spectra for the Identification of Bacteria. *Applied Spectroscopy,* 54**,** 1471-1479.

URBANCZYK-WOCHNIAK, E. B. (2003) Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO reports,* 4**,** 989-993.

VALENZUELA, L., CHI, A., BEARD, S., ORELL, A., GUILIANI, N., SHABANOWITZ, J., HUNT, D. F. & JEREZ, C. A. (2005) Genomics, metagenomics and proteomics in biomining microorganisms. *Biotechnology Advances,* 24**,** 197-211.

VAN DEN BERG, R., HOEFSLOOT, H., WESTERHUIS, J., SMILDE, A. & VAN DER WERF, M. (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics,* 7**,** 142.

VÁZQUEZ, I., RODRÍGUEZ-IGLESIAS, J., MARAÑÓN, E., CASTRILLÓN, L. & ÁLVAREZ, M. (2007) Removal of residual phenols from coke wastewater by adsorption. *Journal of Hazardous Materials,* 147**,** 395-400.

VÁZQUEZ, I., RODRÍGUEZ, J., MARAÑÓN, E., CASTRILLÓN, L. & FERNÁNDEZ, Y. (2006) Simultaneous removal of phenol, ammonium and thiocyanate from coke wastewater by aerobic biodegradation. *Journal of Hazardous Materials,* 137**,** 1773-1780.

VELCULESCU, V. E., ZHANG, L., VOGELSTEIN, B. & KINZLER, K. W. (1995) Serial Analysis of Gene Expression. *Science,* 270**,** 484-487.

VENTER, J. C., REMINGTON, K., HEIDELBERG, J. F., HALPERN, A. L., RUSCH, D., EISEN, J. A., WU, D., PAULSEN, I., NELSON, K. E., NELSON, W., FOUTS, D. E., LEVY, S., KNAP, A. H., LOMAS, M. W., NEALSON, K., WHITE, O., PETERSON, J., HOFFMAN, J., PARSONS, R., BADEN-TILLSON, H., PFANNKOCH, C., ROGERS, Y.-H. & SMITH, H. O. (2004) Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science,* 304**,** 66-74.

VIANT, M. R. (2007) Metabolomics of aquatic organisms: the new 'omics' on the block. *Marine Ecology Progress Series,* 332**,** 301-306.

VIANT, M. R., ROSENBLUM, E. S. & TJEERDEMA, R. S. (2003) NMR-Based Metabolomics: A Powerful Approach for Characterizing the Effects of Environmental Stressors on Organism Health. *Environmental Science & Technology,* 37**,** 4982-4989.

WAGNER, C., SEFKOW, M. & KOPKA, J. (2003) Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry,* 62**,** 887-900.

WALLE, T., WALLE, U. K., WILSON, M. J., FAGAN, T. C. & GAFFNEY, T. E. (1984) Stereoselective ring oxidation of propranolol in man. *British Journal of Clinical Pharmacology,* 18**,** 741–748.

WANG, J. S., CHOU, H. N., FAN, J. J. & CHEN, C. M. (1998) Uptake and transfer of high PCB concentrations from phytoplankton to aquatic biota. *Chemosphere,* 36**,** 1201-1210.

WATANABE, H. (2007) Toxicogenomics as a tool for evaluation of chemical effects. *Yakugaku Zasshi,* 127**,** 1967-1974.

WATANABE, H. & IGUCHI, T. (2006) Using ecotoxicogenomics to evaluate the impact of chemicals on aquatic organisms. *Marine Biology,* 149**,** 107-115.

WATKINS, S. M. & GERMAN, J. B. (2002) Toward the implementation of metabolomic assessments of human health and nutrition. *Current Opinion in Biotechnology,* 13**,** 512-516.

WECKWERTH, W. (2003) Metabolomics in Systems Biology. *Annual Review of Plant Biology,* 54**,** 669-689.

WECKWERTH, W., LOUREIRO, M. E., WENZEL, K. & FIEHN, O. (2004) Differential metabolic networks unravel the effects of silent plant phenotypes. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 7809-7814.

WELMAN, E. (1979) Stabilization of lysosomes in anoxic myocardium by propranolol. *British Journal of Pharmacology,* 65**,** 479-482.

WESTERLUND, A. (1985) Central nervous system side-effects with hydrophilic and lipophilic β-Blockers. *European Journal of Clinical Pharmacology,* 28**,** 73-76.

WHITELEY, A. S. & BAILEY, M. J. (2000) Bacterial Community Structure and Physiological State within an Industrial Phenol Bioremediation System. *Applied and Environmental Microbiology,* 66**,** 2400-2407.

WHITELEY, A. S., MANEFIELD, M. & LUEDERS, T. (2006) Unlocking the 'microbial black box' using RNA-based stable isotope probing technologies. *Current Opinion in Biotechnology,* 17**,** 67-71.

WINDER, C. L., CARR, E., GOODACRE, R. & SEVIOUR, R. (2004) The rapid identification of *Acinetobacter* species using Fourier transform infrared spectroscopy. *Journal of Applied Microbiology,* 96**,** 328-339.

WINDER, C. L., CORNMELL, R., SCHULER, S., JARVIS, R. M., STEPHENS, G. & GOODACRE, R. (2007) Fourier-transform infrared spectroscopy as a metabolic fingerprinting tool for monitoring whole cell biotransformations. *In Preparation*.

WINDER, C. L., DUNN, W. B., SCHULER, S., BROADHURST, D., JARVIS, R., STEPHENS, G. M. & GOODACRE, R. (2008) Global Metabolic Profiling of *Escherichia coli* Cultures: an Evaluation of Methods for Quenching and Extraction of Intracellular Metabolites. *Analytical Chemistry,* 80**,** 2939-2948.

WINDER, C. L. & GOODACRE, R. (2004) Comparison of diffuse-reflectance absorbance and attenuated total reflectance FT-IR for the discrimination of bacteria. *The Analyst,* 129**,** 1118-1122.

WINDER, C. L., GORDON, S. V., DALE, J., HEWINSON, R. G. & GOODACRE, R. (2006) Metabolic fingerprints of *Mycobacterium bovis* cluster with molecular type: implications for genotype-phenotype links. *Microbiology,* 152**,** 2757-2765.

WINDIG, W., HAVERKAMP, J. & KISTEMAKER, P. G. (1983) Interpretation of sets of pyrolysis mass spectra by discriminant analysis and graphical rotation. *Analytical Chemistry,* 55**,** 81-88.

WINSON, M. K., GOODACRE, R., TIMMINS, E. M., JONES, A., ALSBERG, B. K., WOODWARD, A. M., ROWLAND, J. J. & KELL, D. B. (1997) Diffuse reflectance absorbance spectroscopy taking in chemometrics (DRASTIC). A hyperspectral FT-IR-based approach to rapid screening for metabolite overproduction. *Analytica Chimica Acta,* 348**,** 273-282.

WISHART, D. S., KNOX, C., GUO, A. C., EISNER, R., YOUNG, N., GAUTAM, B., HAU, D. D., PSYCHOGIOS, N., DONG, E., BOUATRA, S., MANDAL, R., SINELNIKOV, I., XIA, J., JIA, L., CRUZ, J. A., LIM, E., SOBSEY, C. A., SHRIVASTAVA, S., HUANG, P., LIU, P., FANG, L., PENG, J., FRADETTE, R., CHENG, D., TZUR, D., CLEMENTS, M., LEWIS, A., DE SOUZA, A., ZUNIGA, A., DAWE, M., XIONG, Y., CLIVE, D., GREINER, R., NAZYROVA, A., SHAYKHUTDINOV, R., LI, L., VOGEL, H. J. & FORSYTHE, I. (2009) HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research,* 37**,** D603-610.

WOLD, H. (1966) Estimation of principal components and related models by iterative least squares. IN KRISHNAIAH, P. R. (Ed.) *Multivariate Analysis.* New York, Academic Press.

WOLD, H. (1984) Partial least squares. IN JOHNSON, N. L. & KOTZ, S. (Eds.) *Encyclopedia of Statistical Science.* New York, John Wiley & Sons.

WOLD, S. (1978) Cross-Validatory Estimation of the Number of Components in Factor and Principal Components Models. *Technometrics,* 20**,** 397-405.

WOLD, S., ESBENSEN, K. & GELADI, P. (1987) Principal component analysis. *Chemometrics and Intelligent Laboratory Systems,* 2**,** 37-52.

WOLD, S., SJÖSTRÖM, M. & ERIKSSON, L. (2001a) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems,* 58**,** 109-130.

WOLD, S., TRYGG, J., BERGLUND, A. & ANTTI, H. (2001b) Some recent developments in PLS modeling. *Chemometrics and Intelligent Laboratory Systems,* 58**,** 131-150.

WOOD, A. J. J. (1984) Pharmacologic differences between beta blockers. *American Heart Journal,* 108**,** 1070-1077.

YANG, Y., SU, B., YAN, Q. & REN, Q. (2005) Separation of naproxen enantiomers by supercritical/subcritical fluid chromatography. *Journal of Pharmaceutical and Biomedical Analysis,* 39**,** 815-818.

YEN, K. M., KARL, M. R., BLATT, L. M., SIMON, M. J., WINTER, R. B., FAUSSET, P. R., LU, H. S., HARCOURT, A. A. & CHEN, K. K. (1991) Cloning and characterization of a *Pseudomonas mendocina* KR1 gene cluster encoding toluene-4-monooxygenase. *Journal of Bacteriology,* 173**,** 5315-5327.

ZELIKOFF, J. T. (1993) Metal pollution-induced immunomodulation in fish. *Annual Review of Fish Diseases,* 3**,** 305-325.

ZHAO, H., PARRY, R. L., ELLIS, D. I., GRIFFITH, G. W. & GOODACRE, R. (2006) The rapid differentiation of *Streptomyces* isolates using Fourier transform infrared spectroscopy. *Vibrational Spectroscopy,* 40**,** 213-218.

ZHAO, J.-S. & WARD, O. P. (1999) Microbial degradation of nitrobenzene and mono-nitrophenol by bacteria enriched from municipal activated sludge. *Canadian Journal of Microbiology,* 45**,** 427-432.

ZUEHLKE, S., DUENNBIER, U., LESJEAN, B., GNIRSS, R. & BUISSON, H. (2003) Long term comparison of trace organics removal performances between conventional and membrane activated sludge processes. *WEFTEC 2003.* Los Angeles.