

Mass Spectrometric and Molecular Analyses of Biological Agents
In Environmental Compartments

by

Erica Marie Hartmann

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2012 by the
Graduate Supervisory Committee:

Rolf Halden, Chair
Giovanna Ghirlanda
Randall Nelson

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

This thesis discusses the use of mass spectrometry and polymerase chain reaction (PCR), among other methods, to detect biomarkers of microorganisms in the environment. These methods can be used to detect bacteria involved in the degradation of environmental pollutants (bioremediation) or various single-celled pathogens, including those posing potential threats as bioterrorism agents.

The first chapter introduces the hurdles in detecting in diverse environmental compartments in which they could be found, a select list of single-celled pathogens representing known or potential bioterrorism agents. These hurdles take the form of substances that interfere either directly or indirectly with the detection method. In the case of mass spectrometry-based detection, many of these substances (interferences) can be removed via effective sample pretreatment.

Chapters 2 through 4 highlight specific methods developed to detect bioremediation or bioterrorism agents in environmental matrices. These methods are qualitative mass spectrometry, quantitative PCR, and quantitative mass spectrometry, respectively. The targeted organisms in these methods include several bioremediation agents, e.g. *Pseudomonas putida* F1 and *Sphingomonas wittichii* RW1, and bioterrorism agents, e.g. norovirus and *Cryptosporidium parvum*. In Chapter 2, I identify using qualitative mass spectrometry, biomarkers for three bacterial species involved in bioremediation. In Chapter 3, I report on a new quantitative PCR method suitable for monitoring of a key gene in yet another bioremediation agent, *Sphingomonas wittichii* RW1; furthermore, I apply this

method to track the efficacy of bioremediation in bioaugmented environmental microcosms. In Chapter 4, I report on the development of new quantitative mass spectrometry methods for two organisms, *S. wittichii* RW1 and *Cryptosporidium parvum*, and evaluate two previously published methods for their applicability to the analysis of complex environmental samples.

In Chapter 5, I review state-of-the-art methods for the detection of emerging biological contaminants, specifically viruses, in environmental samples. While this summary deals exclusively with viral pathogens, the advantages and remaining challenges identified are also applicable to all single-celled organisms in environmental settings. The suggestions I make at the end of this chapter are expected to be valid not only for future needs for emerging viruses but also for bacteria, eukaryotic pathogens, and prions. In general, it is advisable to continue the trend towards quantification and to standardize methods to facilitate comparison of results between studies.

ACKNOWLEDGMENTS

This work would not have been possible without the support and guidance of so many people. This text pales in comparison to the full extent of the work it represents, and my acknowledgements similarly can never truly express the gratitude I feel towards the innumerable people who helped me reach this point. I cannot name every individual and describe their contribution, but I have tried to highlight a few whose influence stands out in my memory.

I would like to thank the people at Arizona State University who have mentored me at various points throughout my studies, especially Drs. Giovanna Ghirlanda, Randall Nelson, and Cody Youngbull. I have to thank Jonathan Badalamenti for teaching me all of the genetic techniques I should have already known. I also received invaluable help from Drs. Benny Pycke and Nicole Hansmeier regarding both my research and my career.

I owe a tremendous debt of gratitude to Dr. Angela Jansen for introducing me to collaborators, giving me access to any instrumentation I wanted, and most of all, bringing a little Baltimore charm to the desert. I am grateful for the assistance I received from my program administrator, Maria Hanlin, and my lab manager, Diane Hagner. Thanks also to all of my colleagues in the Swette Center for Environmental Biotechnology and my classmates in the Biological Design Graduate Program.

I would also like to acknowledge some wonderful people outside ASU who contributed to my studies. Dr. David Colquhoun was my inspiration at the Johns Hopkins Bloomberg School of Public Health. I can never repay Dr.

Toshiyuki Sakaki and his group at Toyama Prefectural University for their indescribable generosity during my internship in their lab.

I would be remiss to not thank my family, especially my parents for pushing me to value my education so highly, and my cousins, Dennis and Cindy, and their children, Kayla and Matty, for graciously letting me into their home.

Of course, none of this would have ever happened without my advisor, Dr. Rolf Halden, who took me on as an undergraduate intern and convinced me to come to Arizona to pursue my graduate degree. I am forever indebted to him for introducing me to research science and giving me free reign to develop my projects independently.

Throughout my doctoral studies I received financial support from the Science Foundation of Arizona, the National Science Foundation, the Japan Society for the Promotion of Science, and the Arizona State University Security and Defense Systems Initiative.

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	ix
LIST OF FIGURES	x
PREFACE	xii
CHAPTER	
1. Challenges of Detecting Bioterrorism Agents in Complex Matrices	1
1.1 Introduction.....	1
1.1.1 The target signal: non-signal problem	1
1.2 Potential Matrices	3
1.3 Potential Solutions	3
1.3.1 Instrumentation	4
1.3.2 Target selection	6
1.3.3 Sample preparation	9
1.4 Overview of Non-Bioterrorism Agent Studies	10
1.4.1 Detection of toluene dioxygenase from <i>Pseudomonas</i> <i>putida</i> F1	10
1.4.2 Characterization of mastitic <i>Escherichia coli</i> in bovine milk	13
1.4.3 Metaproteomics of bacteria in the Sargasso Sea.....	15
1.5 Case Study: Norovirus Detection in Stool	17
1.6 Conclusions.....	21
1.7 Acknowledgements.....	22

CHAPTER	Page
1.8 Transition 1	23
2. Identification of Putative Biomarkers for Toluene-degrading <i>Burkholderia</i> and Pseudomonads by MALDI-TOF MS and PMF	25
2.1 Transition 2	35
3. Quantitative PCR for Tracking Megaplasmid-borne Biodegradation Potential of a Model Sphingomonad.....	37
3.1 Transition 3	49
4. Targeted, Quantitative MALDI-TOF MS Analysis of Biological Agents from Multiple Kingdoms in Environmental Samples.....	50
4.1 Introduction.....	50
4.2 Experimental	54
4.2.1 Peptides	54
4.2.2 Samples	54
4.2.3 Environmental samples	54
4.2.4 Protein extraction	56
4.2.5 SDS PAGE	56
4.2.6 MALDI-TOF/TOF MS	57
4.2.7 Data analysis	58
4.3 Results and Discussion	58
4.4 Conclusions.....	67
4.5 Transition 4	68

CHAPTER	Page
5. Analytical Methods for the Detection of Viruses in Food by Example of CCL-3	
Bioagents.....	70
5.1 Introduction.....	70
5.2 Sample Acquisition.....	74
5.3 Sample Processing	75
5.3.1 Polyethylene glycol (PEG) precipitation	78
5.3.2 Antibody-based concentration	79
5.3.3 Filtration.....	81
5.3.4 Enzymatic digestion.....	82
5.4 Detection	84
5.4.1 Electron microscopy (EM).....	85
5.4.2 Polymerase chain reaction (PCR)	85
5.4.3 Nucleic acid sequence-based amplification (NASBA).....	89
5.4.4 Polymerase chain reaction-electrospray ionization-mass spectrometry (PCR-ESI-MS)	91
5.4.5 Capsid protein detection	92
5.5 Interpretation.....	94
5.5.1 Sample process controls.....	95
5.5.2 Controls for inhibition.....	97
5.5.3 Infectivity.....	98
5.6 Conclusions.....	98
6. Summary and Recommendations	101

CHAPTER	Page
WORKS CITED	105
APPENDIX	
A. Proteomic Profiling of the Dioxin-Degrading Bacterium <i>Sphingomonas wittichii</i> RW1	123
B. PERMISSIONS	144

LIST OF TABLES

Table	Page
1.1 Comparison of case study methods.....	12
2.1 Identification of Bands Excised from SDS PAGE Gels of <i>Pseudomonas putida</i> F1, Showing Ranges of Mascot Scores, Percent Coverage and Number of Matched Peptides.....	31
3.1 Primer Locations and Sequences	40
4.1 Peptides used for quantitation.....	61
5.1 Viruses on the third Contaminant Candidate List.....	72
5.2 Virus concentration methods.	77
5.3 Virus detection methods	86

LIST OF FIGURES

Figure	Page
1.1 The metabolic processes by which complex organic substrates are converted to methane during anaerobic digestion	2
2.1 Schematic of the system boundary (bordered in the dashed blue line) used for aerobic SRT calculations. The conventional system is represented without the AD recycle stream (represented in red).....	30
3.1 Influent and effluent TCOD concentrations in mg COD/L for all trains and phases, excluding phases with three or fewer data points.....	42
3.2 Percent overall system TCOD removal in the liquid and solid phases by phase beginning with phase 9 (excluding methane production).....	45
4.1 Spectra of the diagnostic prion peptide, VVEQMCITQYQR ($m/z = 1497.8$), and its heavy-labeled standard, VV*EQMCITQYQR ($m/z = 1503.8$) (A); the norovirus peptide, TLDPIEVPLEDVR ($m/z = 1495.8$), and its heavy-labeled standard, TLDPIEVPLEDV*R ($m/z = 1501.8$) (B); and the dioxin dioxygenase, GVSEGYIAR ($m/z = 951.5$), and its heavy-labeled standard, GV*SEGYIAR ($m/z = 957.5$) (C).....	59
4.2 To test the dynamic range of the method, serial 10-fold dilutions of target analytes with parallel dilutions of standard peptides in a variety of environmental backgrounds.....	64

Figure	Page
4.3 To test the tolerance of the method to differences in the relative amounts of target and standard, serial 10-fold dilutions of standards were mixed with a constant amount, approximately 50 fmol for the prion protein (A) and norovirus capsid protein (B) and 500 fmol for the dioxin dioxygenase (C), of target.	66
5.1 Examples of the raw data output from various methods of virus detection....	87

PREFACE

It is impossible to culture all microorganisms. The vast diversity of microenvironments and the intricate interactions between microorganisms and their environments cannot all be recreated in the lab. There are pathogens such as norovirus that have long resisted cell culture despite years of concerted efforts (Duizer et al., 2004). There are microbial consortia that perform important ecological functions but whose members cannot perform their roles individually (Ziv-El et al., 2011). Many bacteria that degrade environmental pollutants like dioxins have been detected by PCR but not cultured (Kimura & Kamagata, 2009). Even microorganisms that can be cultured may behave very differently in artificial laboratory conditions compared to their natural environments (Madsen, 1991).

Challenges notwithstanding, it is still important to study these microorganisms, whether for their ecological benefits or potential risks to human health. The impracticality of cell culture and differences between *in situ* and *ex situ* phenomena make it necessary to use mass spectrometric and molecular methods to study biological agents in environmental settings.

The Good, the Bad, and the Ugly

This work addresses variously the detection of biological agents for the remediation of environmental contamination and for the prevention of bioterrorism and protection of public health. While these ends are diametrically opposed, the means to achieve them are the same. In both instances, the goal is to

identify and quantify a single-celled microorganism, preferably revealing some phenotypic information, and in both cases, the challenges of identification and quantification are magnified by the presence of interferences from the matrix in which the target is located. Both applications require the selection of appropriate biomarkers and the development of methods suitable to a wide range of environmental samples. Finally, the constraints of time and cost are equally applicable.

Genes vs. Proteins

Cells are composed of four major classes of macromolecules: proteins, nucleic acids, carbohydrates, and lipids. While any of these macromolecules could theoretically serve as molecular targets for identification and quantification, the latter two are severely limited.

Nucleic acids are by far the most commonly assayed. The polymerase chain reaction (PCR) is the undeniable advantage of nucleic acid-based methods, allowing even very low copy number targets to be amplified almost without limit. Nucleic acid-based methods can reveal the genetic identity (genotype) of the target. Targeting functional genes can also show the metabolic capability. The disadvantage of these assays is that genes are not necessarily expressed, and even the observation of mRNA transcripts does not guarantee a functional enzyme. If the end goal of an assay is ultimately to draw a conclusion about a phenotype, whether that be activity or infectivity, nucleic acid-based techniques leave a gap.

Proteins are more closely linked to phenotype and activity, but the techniques for their detection and quantification are still being refined. Also, the utility of these methods is highly dependent on their limit of detection or quantification because protein targets cannot be amplified in the same way as nucleic acids. Mass spectrometry-based methods offer rapid identification and quantification of whole cells, proteins, or peptides. Improvements in sample processing and instrumentation, notably the advent of tandem mass spectrometry, have lowered limits of detection and quantification and continue to do so.

Quantification

Quantification (or quantitation) is especially important. Whether looking at genes or proteins, the ability to absolutely quantify observations is crucial for interpretation of results and comparison across studies (Chao, Hansmeier, & Halden, 2010). Both PCR- and mass spectrometry-based methods can be rendered quantitative with minor modifications, and examples of both are highlighted later in this dissertation.

Structure of the Dissertation

This dissertation is divided into six chapters. The first discusses the difficulties presented by analyzing microorganisms and their characteristic biomarkers in environmental samples. While it deals specifically with mass spectrometry-based detection, the same principal difficulties hold true for gene-based detection as well. Chapters 2 through 4 cover the development and

application of methods for the observation of various microorganisms in environmental settings. In Chapter 2, I describe the use of a simple separation technique coupled with mass spectrometry to identify biomarkers for biodegradation by bacteria of varying degrees of characterization. In Chapter 3, I describe the development of a quantitative PCR method to track a biodegradative gene located on a megaplasmid in *Sphingomonas wittichii* RW1 and use this method to evaluate the prevalence of the megaplasmid in environmental microcosms. In Chapter 4, I describe the development of quantitative mass spectrometry-based methods to track *S. wittichii* RW1 and *Cryptosporidium parvum* and show the applicability of these and other methods in environmental samples. In Chapter 5, I compare the merits of molecular and mass spectrometry-based techniques in detecting viruses in food samples. By analogy, this discussion also applies to other target organisms. In the final chapter, I summarize the work contained in this dissertation and suggest future avenues of research.

Publication Status of Papers Contained in This Thesis

Chapter 1: Challenges of Detecting Bioterrorism Agents in Complex Matrices (Published in *NATO-Science for Peace and Security-Chemistry and Biology Series: Detection of Biological Agents for the Prevention of Bioterrorism*. ISBN: 9048198143, 2010)

Chapter 2: Identification of Putative Biomarkers for Toluene-Degrading *Burkholderia* and Pseudomonads by Matrix-Assisted Laser Desorption/Ionization

Time-of-Flight Mass Spectrometry and Peptide Mass Fingerprinting (Published in: *Bioscience Biotechnology and Biochemistry*, 2010, 74(7), 1470-1472.)

Chapter 3: Quantitative PCR for Tracking Megaplasmid-borne Biodegradation Potential of a Model Sphingomonad (Published: *Applied and Environmental Microbiology*, 2012, 78(12), 4493-4496)

Chapter 4: Targeted, Quantitative MALDI-TOF MS Analysis of Biological Agents from Multiple Kingdoms in Environmental Samples (Submitted: *Rapid Communications in Mass Spectrometry*, 2012)

Chapter 5: Analytical Methods for the Detection of Viruses in Food by Example of CCL-3 Bioagents (Published: *Analytical and Bioanalytical Chemistry*). Invited paper, 2012, special issue on emerging contaminants in biota.

Appendix A: Proteomic Profiling of the Dioxin-Degrading Bacterium *Sphingomonas wittichii* RW1 (Submitted: *Journal of Biomedicine and Biotechnology*, 2012). Invited paper, focus issue on proteomics.

1. Challenges of Detecting Bioterrorism Agents in Complex Matrices

The target signal: non-target signal problem. Historically, we have used mass spectrometry (MS) to analyze bioterrorism agents as pure samples, but we are now in the position to examine them in more realistic settings, such as environmental or clinical samples. One major obstacle that has traditionally stood in the way of such analysis is what we refer to here metaphorically as the signal-to-non-target signal (S:N) problem. Probing for peptide biomarkers in environmental samples equates to searching for the proverbial needle in a haystack, even when starting with relatively concentrated samples. These biomarkers must be reproducible and unique to the bioterrorism agent in question, and screening for biomarkers can involve the investigation of hundreds of candidate proteins and peptides before finding success. Matrix effects, i.e., the presence of legitimate, albeit undesirable signals from non-target substances in the sample, make it vastly more difficult to find these biomarkers in more complex samples. As the number of non-target peaks increases, signal from the target is more difficult to recognize and observe (see Fig. 1.1).

To avoid ambiguity, it is appropriate to pause here for a few definitions. In this chapter, we use the term “non-target signal” to encompass both true background “noise” and undesirable, interfering signals from sources other than the biomarker(s) of the bioterrorism agent of interest. Detector response from the inquired-after biomarker(s) is referred to as the “target signal.” The “sample matrix” is comprised of everything present in the sample aside from the targets

themselves.

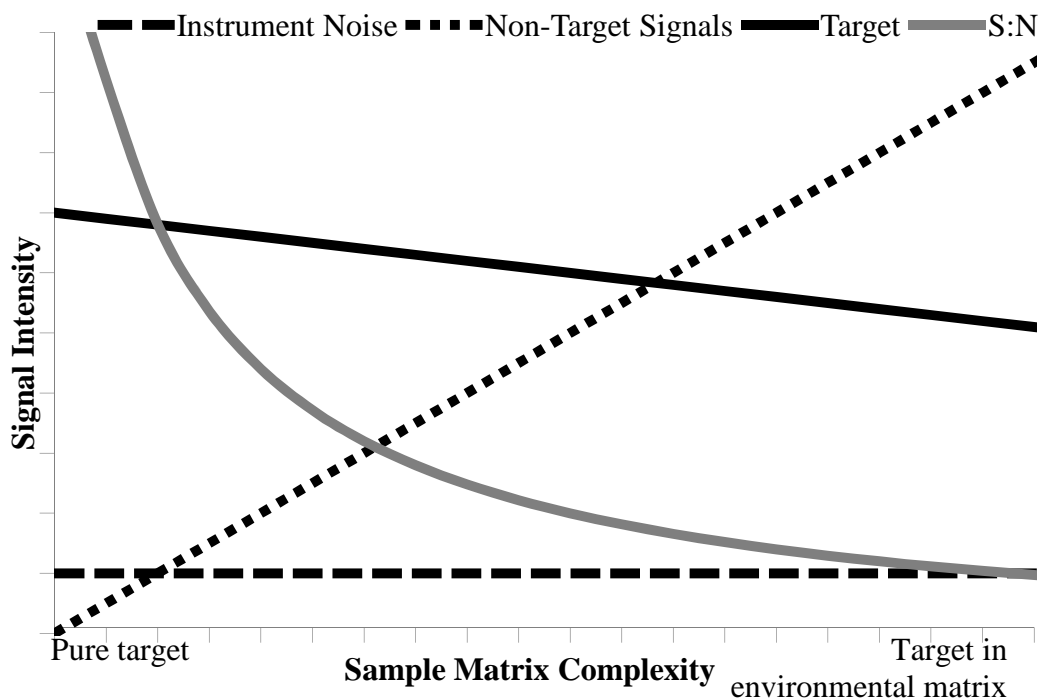


Figure 1.1. A graphical representation of analytical challenges arising from increasing levels of sample matrix complexity. As the sample becomes more complex, phenomena such as ion suppression may reduce the intensity of the signal from a given target. At the same time, signals from non-target masses increase in number and intensity. The ratio of target signal to other signal sources (background chemical noise plus all non-target, signal producing sample constituents) decreases correspondingly, thereby making it progressively more challenging to identify and quantify the target in samples of increasing complexity.

Throughout this chapter, we will refer to the *spectrum of sample matrix complexity*. The spectrum ranges from the simplest sample possible, a purified protein, to the ultimate challenge of environmental or clinical samples, bearing in mind that not all environmental or clinical matrices are equally complex.

While there are many methods of overcoming the S:N problem, many of them involve labor- and time-intensive sample preparation. Since MS-based

techniques should be rapid and designed for high-throughput, these additional measures taken for sample purification may diminish significantly the intrinsic advantages of MS. In the quest for the optimal MS assay, it is therefore desirable to evaluate sample preparation based not only on its effectiveness but also on the time it requires.

Potential Matrices

A general list of potential matrices includes air/aerosols, soil, sediment, water, food (Pellerin, 2000; Sowell et al., 2009), as well as biological specimens of exhaled breath condensate, urine, saliva, blood and stool. Each of these matrices presents its own unique challenges, and considerable differences exist with respect to non-target sample content that may interfere with successful analysis of the target. Matrix signals result from cells and tissues of non-target organisms, proteins, peptides, as well as from other sample constituents including surfactants and/or salts (Apweiler et al., 2009).

Potential Solutions.

As with many obstacles, there are multiple avenues for addressing the S:N problem. The following discussion of potential solutions is not meant to be exhaustive. Rather, it highlights a few areas where progress has already been

made and then focuses on sample preparation techniques because these represent the most accessible and practical avenue in many situations.

Instrumentation. Since the invention of soft ionization, the field of biopolymer MS has already seen progressive technical amelioration over many generations of instrumentation. In general, mass spectrometers have become less expensive and smaller (Griffiths, Jonsson, Liu, Rai, & Wang, 2001). At the same time, resolution, mass accuracy and sensitivity have all improved (Apweiler et al., 2009; Griffiths et al., 2001).

The main soft ionization techniques currently used are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). A variation on ESI known as nanospray has also been developed. Nanospray grants favorable ionization efficiencies, lower limits of detection and enhanced signal intensity (Stutz, 2005). MALDI has seen a number of technical improvements as well. In addition to the hardware, the selection of MALDI matrix—not to be confused with sample matrix—also dramatically impacts the mass range and ionization efficiencies (Renato & Richard, 1998).

These ionization techniques can then be coupled to a suite of MS detectors. MALDI is most commonly paired with a time-of-flight (TOF) detector. While the flight tubes for these TOFs can be quite long, reflectors and, for tandem MS, orthogonal systems have helped to condense these systems. Additional mass analyzers include linear ion traps (LTQ), triple quadrupoles, quadrupole ion traps,

orbitraps, Fourier-transform ion cyclotron resonance (FTICR) systems, and others (Cravatt, Simon, & Yates, 2007; Griffiths et al., 2001).

The advent of tandem MS (MS/MS) has also greatly expanded the realm of protein investigation, making it possible to study all levels of protein structure, from primary to quaternary (Cravatt et al., 2007). It is now possible to determine many post-translational modifications and partial or full amino acid sequences by MS/MS (Apweiler et al., 2009). This latter development has been particularly beneficial for protein identification. The generation of fragment ions for analysis in the second MS can be as simple as post-source decay in MALDI. Here, a small population of ions spontaneously dissociate when induced with high energy during passage through the first flight tube. For more complete fragmentation and the formation of predictable ion series, a host of dissociation techniques, including collision induced dissociation, surface-induced dissociation, black-body infrared radioactive dissociation and electron-capture-induced dissociation (Griffiths et al., 2001), are available. For improved performance and other advantages, MS detectors can be paired in hybrid MS/MS instruments, resulting in mass accuracy below one part per million (Cravatt et al., 2007).

Whereas MS/MS historically was reserved for specialists, this technique is now increasingly accessible (Cravatt et al., 2007; Griffiths et al., 2001). Software packages that take advantage of constantly increasing computing power make it relatively easy to analyze the thousands of spectra that are generated in an MS/MS run. Especially with on-line systems, where a pre-separation step such as

liquid chromatography (LC) is incorporated into the analysis, much of the process can be automated (Apweiler et al., 2009). This degree of automation goes hand in hand with the high-throughput capabilities of MS and MS/MS analyses of proteins.

The increased capacity for proteomics studies is also spurred on by the ever expanding genomic databases and developments in bioinformatics (Apweiler et al., 2009; Griffiths et al., 2001). Without the vast genomic dataset, many of the techniques and studies discussed in this chapter would not be possible.

It is also worth noting that several methods for peptide quantitation have been crafted, including isotope-coded affinity tags (Shiio & Aebersold, 2006) and stable isotope labelling (D. R. Colquhoun, 2007; Ong et al., 2002).

Bioinformatics-based quantitation, such as *Exponentially Modified Protein Abundance Index* (emPAI) (Ishihama et al., 2005; Lippolis, Bayles, & Reinhardt, 2009) and spectral counting (Cravatt et al., 2007; Sowell et al., 2009), are also being used. The ability to not only detect but also quantify targets greatly adds to the appeal of using MS to study protein and peptide biomarkers.

Target Selection. One approach to detecting bioterrorism agents using MS is fingerprinting whole cells to obtain mass spectral “barcodes” (Von Seggern & Halden, 2009; Wahl et al., 2002). This technique can be very informative, providing identification at the sub-strain level for pure samples (Siegrist, 2007) and at the species level for mixtures (Wahl et al., 2002). For bacterial source tracking the reproducibility of MS-based techniques can surpass that of DNA

fingerprinting (Siegrist, 2007). MS of whole cells can be performed with minimal sample preparation requirements.

However, whole cells cannot be identified without extremely pure samples and a library of standard spectra, a bioterrorism monitoring scenario representing the exception rather than the norm. The need for purity, a minimum amount of biomass, and dependence of the method's outcome on growth condition (vegetative state) of the biomass assayed often necessitate cultivation of the sample. Because of the wide mass range required for whole cell fingerprinting, the resolution of the peaks is relatively low, and the number of reproducibly detectable peaks typically is limited (Wunschel et al., 2005). This results in poor statistical power of the identification. Ultimately, this technique is limited by the size of the fingerprint (microbial barcode) library.

MS has also been combined with bioinformatics-driven proteomics. In the top-down approach, intact proteins are introduced first into the MS, fragmented in a collision cell, and the fragment ions are analyzed in the subsequent MS. This approach yields better peak resolution than whole cells because the mass range queried is smaller. The greater resolution allows for the determination of post-translational modifications and amino acid sequence information (Zabrouskov, Senko, Du, Leduc, & Kelleher, 2005). This technique also includes as targets genetically engineered novel pathogens if the latter express a known toxin or virulence factor, as proteins can potentially be identified from transformed organisms (Shiaw-Lin, Ian, William, & Barry, 2004).

Top-down approaches are potentially powerful but require expensive high-resolution instrumentation, such as an FTICR-MS, and the analysis of fragmentation spectra can be challenging and time-consuming because fragmentation patterns are complicated and few databases exist (Zabrouskov et al., 2005). Several improvements are required for this approach, including better fragmentation and faster and more reproducible methods for introducing the sample into the MS (Cravatt et al., 2007). Furthermore, this technique can require more sample preparation because cells typically have to be lysed, and the protein content of the cell separated from the lipids and nucleic acids.

In contrast, the bottom-up proteomics approach is more common, easier to perform, and more tools and inexpensive techniques are available for performing such analyses (Apweiler et al., 2009). In this technique, proteins are digested to render multiple peptides. These peptides provide multiple reference ions, which boost the statistical power of target identifications. Identifications can be obtained using peptide mass fingerprinting (PMF), in which observed peptide masses are compared with theoretically generated mass values, or tandem mass spectrometry, in which characteristic peptides are further fragmented.

Although the bottom-up approach confers many advantages, it can also involve the longest sample preparation, primarily due to digestion. Similar to the top-down approach, this technique is also limited by the available genomic dataset.

All of these methods are valid approaches to identifying bioterrorism agents. Depending on the situation, individual methods may be more or less suited to the task at hand than others. As previously mentioned, time is a very important factor in determining which methods will be used for monitoring of bioterrorism agents in environmental media.

Sample Preparation. Regardless of the target selected, some sample preparation is necessary. Some steps may be simply required for proper functioning of the mass spectrometer. For example, desalting may be used to remove adducts that would shift the observed peaks. Other steps are more geared towards boosting the signal of the target relative to non-target sample constituents. These latter steps often are essential and thus shall be discussed further.

Many different avenues of purification and signal amplification have been investigated, and a few will be discussed further in the next section. Commonly employed purification options include size fractionation by sorting, physical screening, chemical treatment (e.g., precipitation), concentration via affinity and chromatographic separation. Additional options exist but, in the interest of brevity, are not discussed here. Culturing may also be used to increase the biomass available for analysis, thereby amplifying the target mass and improving the target to non-target ratio within the sample matrix. However, this approach may not always be practical and may be time-prohibitive.

As with target selection, choice of sample preparation methods greatly impacts the time it takes to get from sample collection to identification. Some methods may also make the protocol prohibitively costly or difficult. It is therefore crucial to consider not just the efficacy of the preparation but also whether or not it is appropriate for the situation, i.e. routine high-throughput monitoring for bioterrorism agents.

Overview of Non-Bioterrorism Agent Studies

While not concentrating explicitly on the issue of bioterrorism monitoring, the following studies can serve to highlight the challenges of the task at hand. All studies discussed in this section concentrate on the identification of proteins and bacteria in complex sample matrices using MS. The methods used in these studies are also directly applicable to the monitoring for bioterrorism agents. For a brief comparison of these methods see Table 1.1.

Detection of toluene dioxygenase from *pseudomonas putida* F1. This study aimed to identify the toluene dioxygenase (gi|148548093) expressed by *Pseudomonas putida* F1 as a marker for growth on toluene (Hartmann, Colquhoun, & Halden, 2010).

Previous work indicated that it is possible to identify, with almost no sample preparation, catabolic biomarkers for biodegradation from a pure culture of aerobic, biodegradative bacteria grown on minimal media supplemented with the substrate of interest (Halden, Colquhoun, & Wisniewski, 2005b). The

procedure involved the extraction and digestion of the whole soluble proteome of *Sphingomonas wittichii* RW1, a dioxin mineralizing bacterium that, for convenience was grown on inexpensive and readily available dibenzofuran for expression of the dioxin degradation pathway. This digest was then analyzed using MALDI-TOF MS. The dioxin dioxygenase could then be identified from these spectra using PMF (Halden et al., 2005b).

Successful detection of a defined target protein in bacterial whole cell extracts by PMF is extraordinary because bacteria express hundreds, if not thousands, of proteins at any given point in time. That a single protein can be identified from the whole soluble proteome without any sort of separation implies that this protein is very highly expressed, yields tryptic peptides in the mass range examined, and ionizes favorably using the chosen ionization technique, in this case MALDI. Indeed, when the protocol was applied to other bacteria grown on other toxic substrates for expression of distinct catabolic enzymes other than the dioxin dioxygenase, their signals, while present were drowned out by non-target proteins (Hartmann et al., 2010).

Table 1.1. Comparison of case study methods. Mass spectrometry (MS), sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE), strong cation exchange (SCX), solid phase extraction (SPE), reverse-phase high-performance (RP-HP) liquid chromatography (LC), matrix-assisted laser desorption/ionization (MALDI), time-of-flight (TOF), quadrupole (Q), electrospray (ESI), linear ion trap (LTQ), peptide mass fingerprinting (PMF), tandem mass spectrometry (MS/MS), National Center for Biotechnology Information non-redundant protein sequence database (NCBIInr), environmental protein coding sequences (eCDs)

Matrix	Bacterial Cell	Bovine Milk	Sargasso Seawater	Human Stool
Target	Toluene dioxygenase	<i>E. coli</i> proteins	SAR11 proteins	Norovirus capsid proteins
Chemical Treatment				Vertrel XF
Separation from Matrix	Centrifugation	Sucrose gradient centrifugation	Filtration	Filtration (2X)
Protein Separation	SDS-PAGE			
Peptide Separation #1		SCX	SPE or SCX	
Peptide Separation #2		RP-HPLC	LC	
MS	MALDI-TOF	Nanospray-Q-TOF	ESI-LTQ	Nanospray-Q-TOF
Analysis Software	MASCOT PMF	MASCOT MS/MS	SEQUEST	MASCOT MS/MS
Database	NCBIInr	SWISSPROT [<i>E. coli</i>]	In-house eCDs	NCBIInr
Time Estimate	9 hours	20 hours	12 hours	8 hours

The problem of competing signals from other proteins within the organism is the first increment in the spectrum of sample matrix complexity from purified protein to environmental samples. To reduce the sample complexity, the whole cell lysates were separated using SDS PAGE, and individual bands were excised for MALDI-TOF MS analysis (Hartmann et al., 2010). This simple step was enough to allow for the identification of several proteins—including the target enzyme for *P. putida* F1 (Hartmann et al., 2010). However, the additional time required to run the gel, stain it, excise the bands, destain them, and extract the sample from the gel amounted to about four hours. The gel-based step also theoretically raises the limit of detection (Link et al., 1999), although quantitative work was not done.

Although these disadvantages make the technique undesirable for the detection of bioterrorism agents, the results demonstrate that it is possible to get meaningful identifications from pure cultures with a single separation step. More rapid separation steps that result in less sample loss would be more amenable to counterterrorism work.

Characterization of mastitic *Escherichia coli* in bovine milk. The goal of this study was to determine if growth on milk, as opposed to Luria-Bertani (LB) broth, influenced the expression of proteins linked with pathogenicity in *Escherichia coli* (Lippolis et al., 2009). To that end, the researchers cultured *E. coli* on both bovine milk and LB broth. They collected two fractions of the proteome, cytosolic and membrane-associated (Lippolis et al., 2009).

To separate the targets from interfering proteins from the media, especially caseins, the researchers washed the harvested cells twice with cold Dulbecco's Phosphate Buffered Saline, centrifuged twice in a sucrose gradient, and then washed six more times. Although Western blots indicated that the caseins were removed from the *E. coli* cultures, proteins of bovine origin were still identified (Lippolis et al., 2009).

Prior to MS analysis, samples were further fractionated using strong cation exchange (SCX) followed by reverse-phase high performance (RP-HP) liquid chromatography (LC). These steps are useful to separate proteins from within the target organism, as was seen in the previous example. The researchers successfully identified 633 proteins from *E. coli*, several of which had biologically relevant functions for growth in milk and some may be involved in pathogenesis. However, they also identified 25 bovine proteins. Using emPAI, they determined that there were over 100 bacterial proteins that were more abundant than the most abundant bovine protein identified in the cytosolic fraction. However, the most abundant bovine protein in the membrane-associated fraction was the 24th most abundant protein identified in that fraction (Lippolis et al., 2009).

Returning to the spectrum of sample matrix complexity, we have added interfering signals from the matrix, i.e., bovine proteins, as well as signals from endogenous proteins. This procedure takes an estimated 20 h from sample collection to protein identification, excluding the substantial time required for

microbial cultivation. It would therefore take roughly 3 days for an analyst to perform this protocol. Much of the time involved in this protocol is devoted to the column-based SCX and RP-HPLC separations. The most interesting step, in terms of overcoming challenges presented by a complex matrix, is the additional centrifugation steps to remove milk proteins, which take approximately 160 minutes. These results demonstrate that, for relatively simple liquid sample matrices, centrifugation is sufficient to remove enough interferences to enable successful identification of the biological agent of interest.

Metaproteomics of bacteria in the Sargasso Sea. The purpose of this study was to observe proteins expressed by SAR11, a clade (phylogenetically related group) of abundant marine bacteria (Sowell et al., 2009). Samples were taken from the Sargasso Sea, where bacterial growth is often nutrient-limited. These samples contain interferences from the medium, i.e., seawater, and from other bacteria, especially of the *Synechococcus* and *Prochlorococcus* genera. The presence of multiple SAR11 proteins also confounds the analysis because the multiple targets must be unequivocally identified, and some peptides are not unique to a single protein (Sowell et al., 2009). As such, this study used a true environmental sample, one towards the far extreme of the spectrum of sample matrix complexity and target-to-non-target ratios.

To separate and concentrate the bacterial biomass, samples were passed through tandem Millipore Pellicon systems with 30 kDa regenerated cellulose filters (Sowell et al., 2009). No attempt was made to separate bacterial species

prior to MS analysis. Cells were lysed and their contents digested. Resultant peptides were processed using either solid phase extraction (SPE) or SCX and then separated using LC coupled to ESI-MS/MS (Sowell et al., 2009). This procedure takes approximately 12 h, excluding sample collection.

To identify proteins from the SAR11 clade, the investigators constructed a database of environmental protein coding sequences (eCDSs). One problem that they encountered is that not all of these eCDSs were unique to SAR11. To test the specificity, they queried the observed SAR11 peptides against similarly constructed databases for *Synechococcus* and *Prochlorococcus* as well as a database compiled from the rest of the metagenomic data from the Sargasso Sea. Of the total 2,215 peptides they used to identify SAR11 proteins, 24 overlapped with the *Synechococcus* eCDSs, 20 with the *Prochlorococcus* eCDSs, and 1,226 with the remaining Sargasso Sea eCDSs. Despite this overlap, the investigators were able to confidently identify 236 proteins of SAR11 origin (Sowell et al., 2009).

Again, we see the use of column-based techniques to separate out both peptides from within the target organism and those from without. In contrast with the previous study, additional signals from the media were removed by filtration, as opposed to centrifugation. To deal with the additional species, the investigators added a bioinformatics component to the analysis. The results of this study demonstrate how bioinformatics, when used in complement with sample preparation, can be used to compensate for sample matrix complexity.

Case Study: Norovirus Detection in Stool

This is a landmark study because it tackles both an extremely difficult target and an extremely difficult sample matrix. Notable here is the simplicity of the method, which is also very rapid and enabled detection at clinically relevant copy numbers of the bioterrorism agent. Because of the relevance of the target and the speed of the method, this study highlights the possibilities and current limitations of MS-based monitoring of bioterrorism agents in complex samples.

Norovirus refers to the *Norovirus* genus of viruses that cause acute gastroenteritis ("Noroviruses," 2006). The CDC has classified norovirus as a Class B bioterrorism agent, meaning that it is "moderately easy to disseminate" with "moderate morbidity rates" and necessitates "specific enhancements of CDC's diagnostic capacity and enhanced disease surveillance" (Onisko et al., 2007). It is estimated that as few as ten virus particles can cause an infection, although clinical virus titers in stool range from 100 to 1,000 fmol/ml ("Noroviruses," 2006).

One of the difficulties of working with norovirus is that it cannot be cultured outside of human hosts. This limitation rules out culturing of samples to increase the amount of target copies. On the upside, norovirus capsids are assembled from a single protein that occurs at a high copy number. To protect analysts from infection, this study made use of virus-like particles (VLPs) consisting of 180 identical capsid proteins which are identical to the bioterrorism

agent except for the fact that they are devoid of any viral, infectious RNA. Capsid proteins comprising virus capsids were attractive targets for protein-based detection methods because they frequently occur in multiple copies per virus particle. In the case of the human norovirus the sequence of the protein (gi|34223984) already had been determined and entered into online genomic databases. An *in silico* tryptic digest showed 15 possible fragments, equating to a potential sequence coverage of up to 58.8% in the 500 to 5,000 *m/z* range (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006).

In terms of sample matrix composition, stool is localized at the far side of extreme complexity (Oleksiewicz, Kjeldal, & Klenø, 2005). In addition to the target, stool samples may contain host (human) cells, animal and plant (food) cells, microbial cells from the gut community (Oleksiewicz et al., 2005), and non-proteinaceous organic and inorganic chemicals (Rang & Dale, 1991).

Before attempting to identify the VLPs in samples of increasing complexity, the investigators examined the pure intact protein using 1D MS (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). With this method, monomers and dimers of the capsid protein were detectable at 60 pmoles. This relatively high detection limit was due to the aggregation of the capsid proteins. Nevertheless, the resultant spectra demonstrated the purity of the synthesized VLPs (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006).

Using the PMF approach with trypsin digestion, the investigators next assayed dilutions of the VLP standard and confidently detected the capsid protein down to levels of 50 fmoles. Detection at 100 fmoles was highly reproducible (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). These detection limits are comparable to the clinically relevant range in which norovirus may occur in watery stool of acutely ill individuals.

The investigators then applied PMF to VLPs spiked into processed human stool extract (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). As could be expected, the S:N ratio was unfavorable and effectively prevented identification of the target. However, a putative peak of m/z 1,495.8 was consistently observable in norovirus-fortified samples. To capitalize on this finding, sample splits were introduced into a nano-ESI-MS/MS instrument to force fragmentation and identification of the putative 1495.8 Da target peptide. In the nanospray single MS spectrum, the peptide was observable at m/z 748.4 as the double charged ion. The investigators therefore selected this peak for collision induced dissociation using MS/MS. Analysis of the VLP-fortified stool extract with this approach produced a nano-ESI-MS/MS spectrum showing 16 fragments corresponding to eight of the 12 y ion series fragments of the capsid protein peptide of sequence TLPDTIEVPLEDVR (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006).

This MS/MS-based method gives a detection limit of $3 \cdot 10^8$ viruses per sample. This detection limit translates to a sample volume requirement of

approximately 125 μ l of stool to enable successful analysis using this approach (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006).

The entire would take about 8 h to complete when leveraging rapid digestion techniques. As performed, the sample processing scheme involved dilution in ammonium bicarbonate buffer, extraction with Vertrel XF to remove lipids, passage through a 0.22 μ m filter followed by a 100 kDa MW cutoff filter, and concentration down to 100 μ l in the final sample volume (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). For the purposes of this study, samples were digested overnight with trypsin (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). More rapid digestion procedures, such as immobilized trypsin columns, could easily be substituted for the overnight digestion. This would allow for a significant reduction in prep time, thereby facilitating execution of the analysis in a single work shift, taking no longer than 8 h.

The approach taken here for increasing S:N ratios differs from the aforementioned studies. The investigators did not rely on gels or LC-separation to concentrate their target and remove undesirable sample constituents. Instead, they reduced chemical interferences by treating with Vertrel XF and they concentrated their target by employing a physical screening approach involving size fractionation by sequential filtration. In the two-step filtration process, the first step serves to remove unwanted, large cells from the sample. The second step served to separate the VLPs but not the smaller dissolved molecules stemming

from stool and cell debris. Finally, MS/MS and bioinformatics tools were used to cope with other endogenous and exogenous proteins present.

As an extension of this method, the investigators also developed a quantitative method using stable isotope-labeled standards and single reaction monitoring. They were able to detect their target at 500 attomoles (D. R. Colquhoun, 2007). However, this method was not tested in environmental samples and thus remains to be proven as a viable method for the detection of norovirus in stool.

Conclusions

During monitoring of environmental samples for bioterrorism agents, the concentration of the target and the complexity of the sample vary widely. Best-case scenarios of samples to be analyzed involve pure or semi-pure powders of dry microbial spores or vegetative cells. In these instances, traditional methods including mass spectral “barcoding” of samples may be sufficient to enable successful analysis. More likely, however, is that the analyst is challenged with the detection of minute target quantities in relatively dilute samples of great complexity. In this situation, sophisticated MS equipment and sample preparation techniques may still allow to determine a given biological agent reproducibly and with confidence. The successful detection of norovirus particles in stool at clinically relevant concentrations served to illustrate the applicability of MS approaches even in these extremely unfavorable conditions.

However, the analysis of dilute environmental samples for bioterrorism agents is still in its infancy. More work will be required in future years to abridge existing sample preparation protocols and to introduce new ones to advance the field. Insights can be gained from reviewing medical and environmental studies that do not fall into the domain of monitoring for bioterrorism agents. Alongside with the development of streamlined sample processing techniques, advances in instrument development, bioinformatics and computing power will be critical to propel the research field of bioterrorism monitoring forward.

Acknowledgements

The authors gratefully acknowledge Drs. Tzu-Chiao Chao and Nicole Hansmeier for their help in editing this manuscript.

Transition 1

In Chapter 1, we learned about the various difficulties impeding the identification of markers indicating the presence of microorganisms (biomarkers) in environmental samples. The chief problem is that biomarkers may be present at very low levels, so the target signal can be masked by interferences from the sample. This problem can largely be overcome by separation techniques implemented prior to detection, but at the cost of time, limit of detection, and money.

In both bioterrorism preparedness and bioremediation work, it is imperative to arrive at a positive identification as quickly as possible. At the same time, the limit of detection cannot be sacrificed; a limit of detection that is higher than relevant levels of the bioagent in question is not diagnostically useful. Furthermore, the ideal assay should not be cost-prohibitive.

One of the studies cited in the Introduction showed the use of a relatively simple, time- and cost-effective separation technique, sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis. In Chapter 2, we examine this technique more in depth and see how successful this minimal sample preparation protocol can be when seeking to track bacteria of importance for the bioremediation of toluene, namely *Pseudomonas putida* F1, *P. mendocina* KR1, and *Burkholderia* sp. JS150.

One challenge of working with environmental samples is that full genome sequences are not available for all organisms. It is therefore of interest to examine

an effective analysis method not only for well characterized bioagents but also for ones that have been only partially described. In the following chapter, three bioremediation agents of differing degrees of characterization are analyzed using a combination of sodium dodecyl sulfate polyacrylimide gel electrophoresis and mass spectrometry to identify protein biomarkers of biodegradative activity.

2. Identification of Putative Biomarkers for Toluene Degrading

***Burkholderia* and Pseudomonads by MALDI-TOF MS and PMF**

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) can be used to identify certain bacterial strains that have applications in the field of bioremediation (Halden et al., 2005b). This technique is promising because it is faster than other, traditional methods (C. Fenselau & Demirev, 2001; Lay Jr & Liyanage, 2005), such as 16S rRNA analysis and the measurement of relative substrate concentrations or chromatographic analysis of proteins, and because it can implicate both the presence and the efficacy of the biological agent in a single step. Ideally, the catabolic proteins involved in pollutant degradation both identify the bacterium and implicate it in degradation.

Peptide mass fingerprinting (PMF) can easily reveal the phenotype of the bacterium, but it is yet undetermined whether PMF can be used on a variety of strains and can distinguish between different bacterial strains grown using the same substrate.

MALDI-TOF/TOF MS can be used with isobaric tags for relative and absolute quantitation (iTRAQ) to yield quantitative data (C. Fenselau, 2007), Quantitative evaluation of catabolic enzymes was used in this study to corroborate qualitative findings from MALDI-TOF MS.

This study attempted to use PMF and MALDI-TOF MS to identify *Pseudomonas putida* F1, *Burkholderia* sp. JS150, and *P. mendocina* KR1 grown using toluene as the sole carbon and energy source. All of the strains assayed had been identified as capable of toluene degradation and their degradative pathways characterized (Johnson & Olsen, 1997; Parales, Parales, Pelletier, & Ditty, 2008). Amino acid sequences corresponding to known toluene degradation genes represent the most attractive targets for mass spectrometric analysis because of their implications for bioremediation. *P. putida* F1 was selected as the study organism for MALDI-TOF/TOF MS analysis because it is the only strain to have been fully characterized (NC 009512).

P. putida F1, *Burkholderia* sp. JS150, and *P. mendocina* KR1 were kindly provided by Dr. Rebecca E. Parales, Microbiology, UC Davis; Dr. Jim Spain, School of Civil and Environmental Engineering, Georgia Institute of Technology; and Dr. Paul Hatzinger, the Shaw Group, Inc., respectively. Selective medium consisted of 50 ml of ATCC medium: 1898 (0.5 g NH₄Cl, 670.0 mg Na₂HPO₄•7H₂O, 340.0 mg KH₂PO₄, 112.0 mg MgSO₄•7H₂O, 14.0 mg CaCl₂, 5.0 mg ZnSO₄•7H₂O, 2.5 mg Na₂MoO₄•2H₂O, and 0.13 mg FeCl₃ per liter) supplemented with 40 µl of high performance liquid chromatography (LC) grade toluene. Toluene cultures were sealed with a rubber stopper to prevent substrate volatilization and human health risk from inhalation. Cultures were incubated at 30°C with shaking at 250 rpm for 23 h. All bacteria were grown in 50 ml of Luria-

Bertani (LB) liquid medium (Sambrook, Fritsh, & Maniatis, 1989) as negative controls for the inducible toluene degradation pathway.

Nine ml of each culture was harvested by centrifugation (10,000 x g, 25°C, 10 minutes). Samples were then prepared as the whole-cell extract fraction and were analyzed by MALDI-TOF MS, as described previously (Halden et al., 2005b).

Spectra were analyzed using Data Explorer software (Applied Biosystems, Foster City, CA). Peaks within 2 standard deviations of the background noise were removed using the Noise Filter/Smooth tool. The remaining peaks were deisotoped to generate monoisotopic spectra. The 100 most intense peaks were searched against all taxa in the NCBI nr database (16 November 2007; 5,633,163 entries) using the Mascot PMF software (<http://www.matrixscience.com/>). A peptide error tolerance of ± 150 ppm but no missed cleavages were allowed. The variable oxidation of methionine was accounted for.

The mass spectra obtained for whole-cell extracts of the three toluene-grown bacterial cultures were distinct from those resulting from the corresponding control biomass grown on rich medium. However, attempts to match detected peaks to predicted peptide masses yielded no positive identifications ($p < 0.05$, Mascot score > 81).

To simplify the sample, whole-cell extracts were separated by SDS PAGE (4-12% bis-tris ReadyGel, Bio-Rad, Hercules, CA) run at 150 V for 90 min, and stained using SimplyBlue SafeStain (Invitrogen, Carlsbad, CA). Distinct protein

banding patterns were observed between the toluene-grown and rich-medium cultures.

The SDS PAGE gels of *Burkholderia* sp. JS150 (two biological replicates) contained fewer than 20 bands, whereas the gels from the other cultures contained at least 30. Three bands present in the toluene cultures and absent in the rich medium cultures (approximate molecular weights, 58, 25, and 20 kDa) were identified by MALDI-TOF MS and PMF as components of a chlorobenzene dioxygenase sequenced from *Pandoraea pnomenusa*.

Several differences in banding appeared between the toluene and the rich medium *P. mendocina* KR1 cultures, but only two (approximate molecular weights, 55 and 30 kDa) might have been identified. The low number of identifications were due to the lack of sequence data. The bands that could be positively identified, however, were strain-specific components of the toluene-4-monooxygenase system.

These bands were produced by all three cultures, and a statistically significant identification ($p < 0.05$, Mascot score > 81) was obtained for one of the three biological replicates. Toluene-4-monooxygenase proteins were found in the top 10 Mascot hits regardless of significance, but these results were not significant at the 95% confidence level. It is uncertain why significant scores were not obtained for all samples, but it might have been due at least in part to the number of queries used in the Mascot search, fewer than 100 masses, because no more than 100 masses were observed from these bands. In cases resulting in

significant scores, fewer than 55 queries were submitted. Conversely, cases with more than or exactly 55 queries did not yield significant scores, although the same protein identification was produced.

For *P. putida* F1, seven bands (corresponding to approximate molecular weights of 182, 90, 72, 69, 57, 29, and 22 kDa) appeared to be more intense in the toluene cultures than in the rich-medium cultures (Fig. 2.1). MALDI-TOF MS and PMF revealed two dehydrogenases, a TonB-dependent/ligand-gated channel, heat shock proteins, a growth factor, and the toluene dioxygenase β subunit (gi|148548093) (Table 2.1). Neither the α subunit nor a *meta*-cleavage enzyme were identified, possibly as a result of unfavorable ionization tendencies with MALDI. This finding mirrors results of prior work in which the β subunit of the dioxin dioxygenase was known to be present but was not detected (Halden et al., 2005b).

The heat-shock proteins found were Hsp90, DnaK, and GroEL, which are also induced by growth on benzoate in *P. putida* P8 (Cao & Loh, 2008). Other bands produced *Pseudomonas*-specific identifications of constitutive proteins, *e.g.*, DNA-directed RNA polymerase (data not shown). Still other bands appeared more highly expressed in the induced sample, but either could not be excised precisely or could not be identified. While positive identification of the strain-

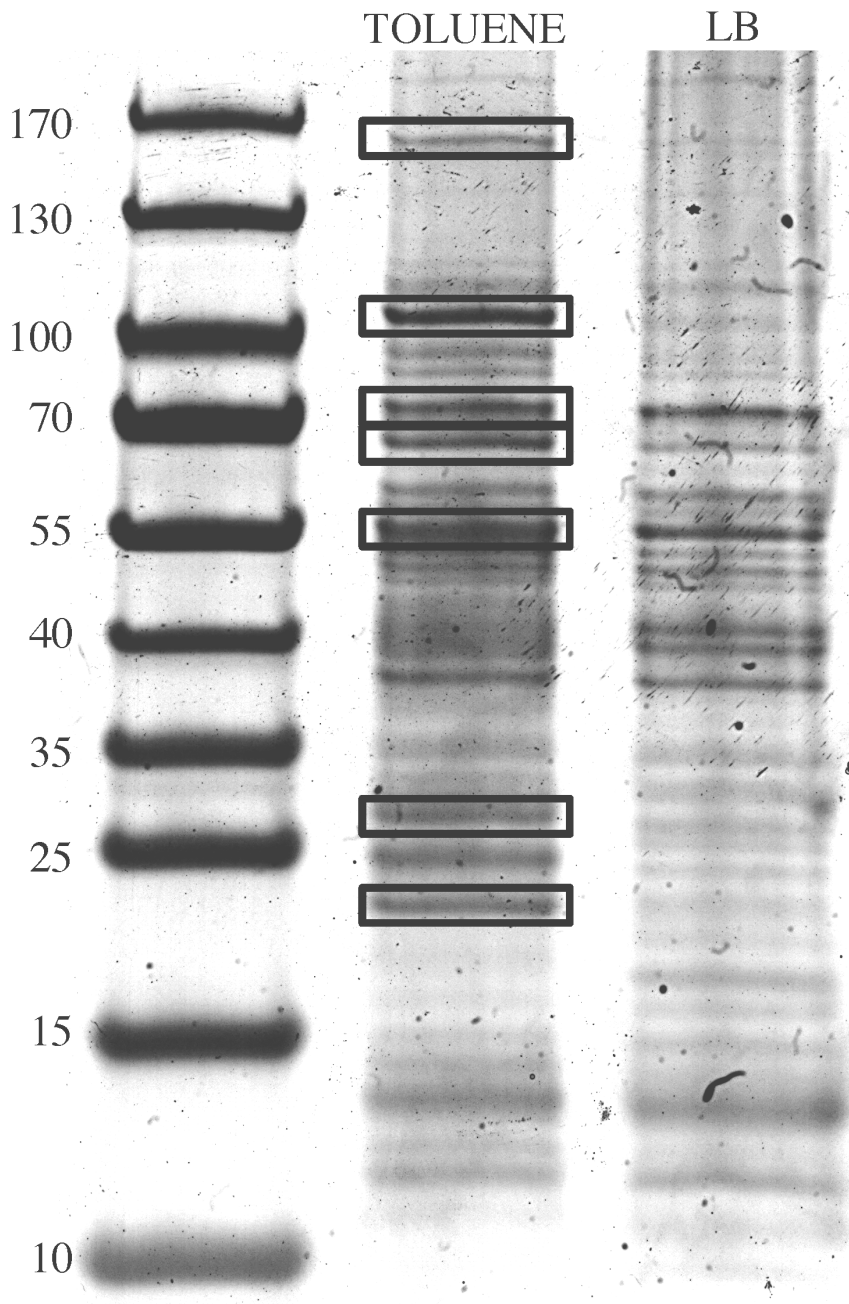


Figure 2.1. SDS PAGE Gel Comparison of the Soluble Proteome of *Pseudomonas putida* F1 Grown on Toluene and on Rich Media (LB). Standard masses are given in kDa. Highlighted bands were identified using Mascot. The results are shown in Table 2.1.

Table 2.1. Identification of Bands Excised from SDS PAGE Gels of *Pseudomonas putida* F1, Showing Ranges of Mascot Scores, Percent Coverage and Number of Matched Peptides. Approximate molecular weights reflect the nominal masses of the identified proteins, to the nearest kDa.

Approx. mol. weight (kDa)	ID	Mascot score % coverage matched peptides
182	NAD-glutamate dehydrogenase <i>Pseudomonas putida</i> F1	100 - 118
		11 - 14%
		(24 - 28)
90	TonB-dependent siderophore receptor <i>Pseudomonas putida</i> F1	84 - 255
		25 - 43%
		(13 - 31)
72	heat shock protein 90 <i>Pseudomonas putida</i> F1	96 - 156
		26 - 35%
		(16 - 22)
69	Pyrrolo-quinoline quinone <i>Pseudomonas putida</i> F1	111 - 163
		27 - 38%
		(17 - 22)
57	chaperonin GroEL <i>Pseudomonas putida</i>	96 - 130
		20 - 39%
		(19 - 10)
29	2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase <i>Pseudomonas putida</i> F1	100 - 146
		40 - 61%
		(11 - 15)
22	toluene dioxygenase <i>Pseudomonas putida</i> F1	84 - 94
		32 - 40%
		(6 - 7)

specific catabolic target, toluene dioxygenase, was produced from only one of the three biological replicates, the corresponding band was observed in all three cultures.

Two toluene-grown samples and two controls of the *P. putida* F1 cultures were prepared as described above for quantitative proteomic analysis. An iTRAQ (Applied Biosystems, Foster City, CA) labeling protocol was conducted following the manufacturer's directions. LC spotting and MALDI-TOF/TOF analysis were performed following the manufacturer's recommendations using a Tempo spotter and AB 4800 respectively (Applied Biosystems).

Spectra were analyzed using a customized off-line database for *P. putida* F1 generated from the NCBI GenBank file. Peptides were searched against this database using the Mascot PMF software with an error tolerance setting of ± 50 ppm; no missed cleavages were allowed. The variable oxidation of methionine was accounted for, as well as iTRAQ8plex modifications on N-termini, lysine, and tyrosine residues. Confidence intervals (95%) and *p*-values were obtained with ProteinPilot (Applied Biosystems) for each iTRAQ reagent ratio comparison. Only those proteins demonstrating significant ($p < 0.05$) quantitative differences in all four pairwise comparisons were considered to be modulated.

Quantitative analysis of the toluene dioxygenase revealed that expression increased between 5.7- and 2.0-fold (95% confidence). Similarly, the ring hydroxylating α subunit (gi|148548093) was also more highly expressed in the toluene-grown cultures, by a factor between 4.0 and 2.7. These results corroborate

the observations by the 1D MS technique, and support the postulate that toluene dioxygenase is a suitable biomarker for F1.

The results for *P. putida* F1 and *P. mendocina* KR1 indicate the potential use of PMF and MALDI-TOF MS for strain-specific identification and characterization of biodegradative bacteria. The results for *Burkholderia* sp. JS150 suggest that this technique may also provide valuable insight for determining homology and identifying evolutionary pathways. Although the enzyme identified in the *Burkholderia* sp. JS150 culture was not the expected metabolic biomarker, it is a metabolic enzyme with a function similar to the expected target, and might therefore be related.

It is plausible that a chlorobenzene dioxygenase might have functional homology to a toluene dioxygenase due to the structural similarities of their respective substrates. Furthermore, in this study, the 2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase from *P. putida* F1 was identified in tandem with a chlorobenzene dioxygenase (gi|61394236). In a genetic study exploring a potential mechanism for the evolution of the chlorobenzene degradation pathway in *Ralstonia* sp. JS705, 90% sequence homology was found between the *mcbAa* gene, thought to encode a terminal oxygenase, and the *todCI* gene from *P. putida* F1 (Muller, Werlen, Spain, & van der Meer, 2003), which encodes the α subunit of the oxygenase component of the toluene dioxygenase (Zylstra & Gibson, 1991).

The use of MS-based proteomics in the field of biodegradation is expanding, but it is still primarily limited to methods including two-dimensional separation and chromatography techniques (Seung, Choi, & Kahng, 2007). The present study indicates that simple MALDI-TOF MS can provide meaningful results with a minimal amount of sample preparation for multiple bacterial strains of varying degrees of characterization. This technique can also be useful in screening for the presence of degradative enzymes and in suggesting possible genes of interest.

Transition 2

The previous chapter showed the utility of a relatively simple sample process stream to identify biomarkers for the presence and efficacy of several organisms of importance to bioremediation. The combination of mass spectrometry and peptide mass fingerprinting was used to identify key degradative enzymes that could serve as biomarkers for the use of *Pseudomonas putida* F1, *P. mendocina* KR1, and *Burkholderia* sp. JS150 in the biodegradation of toluene.

This work was based on a similar approach taken to investigate another organism, *Sphingomonas wittichii* RW1 (Halden et al., 2005b), wherein similar techniques were used to identify the dioxin dioxygenase as a biomarker for dioxin biodegradation. In this study, the sodium dodecyl sulfate polyacrylimide gel electrophoresis was obviated because the target enzyme is so highly expressed that the investigators were able to identify it from the cellular background without any additional separation.

While these studies have demonstrated the advantages of such mass spectrometry-based methods for the field of environmental engineering and specifically bioremediation, their use is far from mainstream. In the next chapter, a biomarker discovered using mass spectrometry is translated to a more mature, widely accepted technique: quantitative PCR.

Chapter 3 describes the development of a quantitative PCR method for the gene encoding the large subunit of the dioxin dioxygenase, *dxnA1*. This method is

then used to examine the plasmid copy number relative to the population size of *S. wittichii* RW1 in landfill leachate microcosms.

3. Quantitative PCR for Tracking Megaplasmid-borne Biodegradation

Potential of a Model Sphingomonad

The genus *Sphingomonas* is an unusual group of α -proteobacteria known for their extraordinary ability to degrade recalcitrant environmental pollutants (White, Sutton, & Ringelberg, 1996). A survey of 18 sphingomonads showed that almost all strains carried 2 to 5 different, large (50 to 500 kb) megaplasmids (Basta, Keck, Klein, & Stolz, 2004a). Their biodegradative genes are often extrachromosomally encoded (Basta et al., 2004a), but they may not be organized in traditional operons or located on the same plasmid (Tabata et al., 2011). The frequency with which biodegradative sphingomonads are discovered, coupled with their unique genetic structure and reluctance to share plasmids amongst other genera, suggests that these megaplasmid-harboring organisms have an adaptive advantage to metabolize anthropogenic environmental pollutants (Stolz, 2009).

S. wittichii RW1 is the first identified aerobic bacterium that can degrade mono- through hexa-chlorinated dioxins (Halden, Halden, & Dwyer, 1999b; Hong, Chang, Nam, Fortnagel, & Schmidt, 2002; Nam, Kim, Schmidt, & Chang, 2006) and use dibenzo-*p*-dioxin or dibenzofuran as a sole source for carbon and energy (Sakaki & Munetsuna, 2010). Polychlorinated dibenzo-*p*-dioxins are ubiquitous environmental pollutants from various sources, including waste incineration (Hays & Aylward, 2003). *S. wittichii* RW1 has been investigated as a bioremediation agent and shows promise (Nam et al., 2005; Nam et al., 2006). It is also a model sphingomonad, thanks to its high degree of characterization and

the complete sequencing of its genome and two megaplasmids (Miller et al., 2010).

Dibenzo-*p*-dioxin degradation proceeds through the dioxin dioxygenase, which is encoded by the genes *dxnA1* and *dxnA2*, together denoted *dxnA1A2*, on the 220-kb plasmid, pSWIT02 (GI:148550845) (Miller et al., 2010). This enzyme catalyzes the initial dihydroxylation of the aromatic ring (Sakaki & Munetsuna, 2010), leading to its spontaneous cleavage into 2,2',3-trihydroxybiphenyl, which is dihydroxylated by a second enzyme into 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate and then hydrolyzed to form salicylic acid (Seah et al., 2007). These downstream metabolites can then be funneled into the citric acid cycle and used for carbon and energy. The dioxin dioxygenase also transforms chlorinated congeners of dibenzo-*p*-dioxin, resulting in the formation of chlorinated salicylates (Seah et al., 2007). This initial transformation destroys and detoxifies the planar structure of the molecule (Nojiri & Omori, 2002), but the resultant chlorinated intermediates typically do not support growth.

Due to its crucial role in dioxin degradation, the dioxin dioxygenase is therefore a promising candidate biomarker (Halden et al., 2005b). A protein-based detection method for the dioxin dioxygenase exists (Halden et al., 2005b), and standard PCR primers have been published for the study of the *dxnA1A2* cistron (Armengaud, Happe, & Timmis, 1998a; Basta et al., 2004a) and for ring-hydroxylating dioxygenases in general (Kimura & Kamagata, 2009). However, due to the amplicon length or lack of specificity, these primers are not appropriate

for quantitation, which is important in linking microbial activity with chemical transformations to track bioremediation (van der Zaan et al., 2010).

In the present study, we developed quantitative polymerase chain reaction (qPCR) primers to detect the *dxnA1* gene with sufficient sensitivity and specificity for use in environmental samples. We then used this method to determine the copy number of the pSWIT02 megaplasmid and monitor *S. wittichii* RW1 in bioaugmented landfill leachate.

S. wittichii RW1 was routinely grown in pure culture in minimal medium supplemented with dibenzofuran as previously described (Halden et al., 2005b). Plasmid DNA was extracted using a BACMAX DNA Purification Kit (EPICENTRE, Madison, WI). The *dxnA1A2* gene cluster was amplified as previously (Basta et al., 2004a). To generate a positive control, the product was treated with the QiaQuick PCR Purification Kit (Qiagen, Valencia, CA), cloned into the pCR[®]4-TOPO[®] plasmid using the TOPO TA Cloning Kit for Sequencing (Invitrogen, Carlsbad, CA), and sequenced using an Applied Biosystems 3730 capillary sequencer. A Basic Local Alignment Search Tool (BLAST; <http://www.ncbi.nlm.nih.gov/>) search of the sequence of the cloned product showed 99.95% identity to *dxnA1* (GI:4007779) with 0% gaps and complete coverage.

The qPCR primers for *dxnA1* were designed using the Primer3 software v1.1.4 (Rozen & Skaletsky, 2002) and evaluated for secondary structures using NetPrimer (PREMIER Biosoft International, <http://www.premierbiosoft.com/>).

The specificity of the primers was checked *in silico* using Primer-BLAST, which found no non-specific amplification for the selected primer pair (product size between 60 and 300 base pairs, minimum of 6 mismatches to ignore targets).

Optimized qPCR conditions were as follows: 95°C for 2 min; 40 cycles of 95°C for 10 s, 58°C for 20 s, and 68°C for 30 s, and a melting curve generated at the end of the last cycle. Primers were supplied at a concentration of 300 nM. The qPCR procedure was performed using 5 PRIME RealMasterMix for SYBR Green (Fisher Scientific, Pittsburgh, PA) on a Mastercycler ep realplex (Eppendorf, Hamburg, Germany). qPCRs contained 4 µl of template DNA in a final reaction volume of 10 µl.

The primer pair 321-451 (Table 3.1) produced a standard curve using tenfold serial dilutions of our *dxnA1* control plasmids with a slope of 3.33, indicating a near ideal amplification efficiency of 1.01, linearity over 8 orders of magnitude, and a limit of quantification of 62 copies (Figure 1). The corresponding melting curve showed only one peak, indicating a unique product.

To evaluate primer specificity, DNA extracted from *S. wittichii* RW1 was introduced at concentrations ranging from 1 ng/µl to 0.001 ng/µl (10^6 to 10^2 copies), into a background of DNA extracted from an activated sludge sample

Table 3.1. Primer Locations and Sequences

Primer	5' Base	3' Base	Sequence (5' – 3')
dxnA1fwd321	321	341	TCATG GCTGG GTGTT CAATA
dxnA1rev451	431	451	CGAAA ATCAG CCCCT TGTAG

from the aeration tank of the Mesa Northwest Wastewater Reclamation Project or agricultural soil (Walters, McClellan, & Halden, 2010) (Figure 3.1). Extractions from environmental samples were performed with the NucleoSpin Soil kit (Macherey-Nagel, Düren, Germany). The total concentration of target and background DNA added to the PCR reaction was always 1 ng/μl for a total of 4 ng DNA/reaction.

Amplification of target DNA was linear over 4 orders of magnitude, from 10^2 to 10^6 copies ($R^2 = 0.99$); fluorescence signal from the background samples was similar to that of the non-template controls. This performance is similar to that of other methods used for 16S rDNA (linearity from 10^2 to 10^7 target cells, with each cell containing between 1 and 15 copies of the 16S rDNA gene, against a background of DNA from 7 non-target bacteria) (Maeda et al., 2003) and the benzylsuccinate synthase gene *bssA* (linearity from 10^2 to 10^6 genome copies against a background of 3 non-target bacteria) (Beller, Kane, Legler, & Alvarez, 2002). Probe-based, as opposed to SYBR-based, qPCR methods can achieve detection limits an order of magnitude lower (Ritalahti et al., 2006).

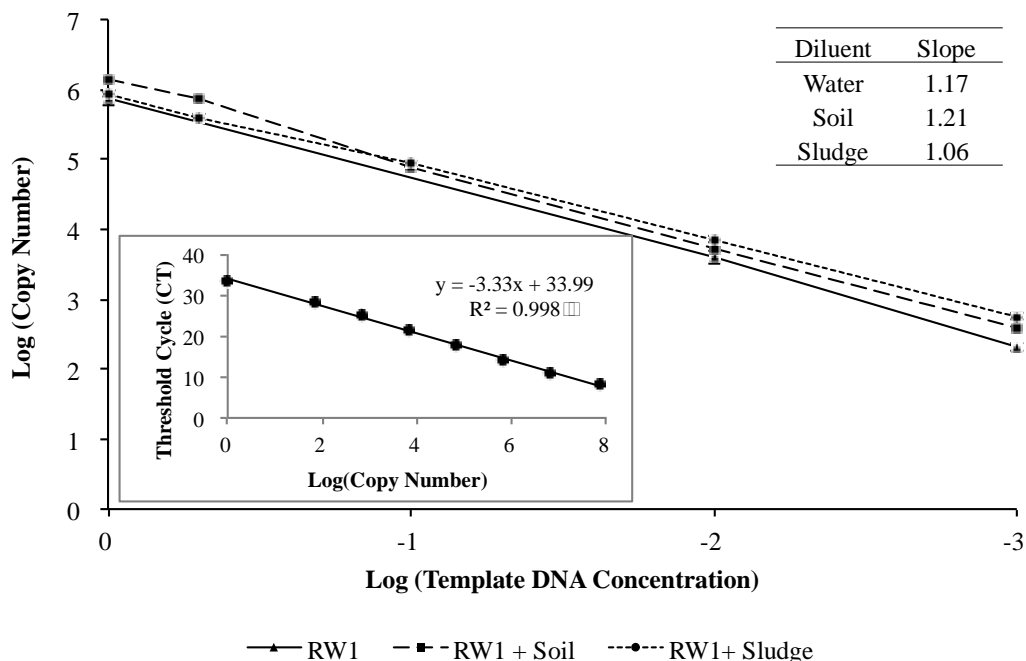


Figure 3.1. Correlation of observed copy number and concentration of template DNA from pure culture and samples spiked with genetic background. Genomic DNA from pure culture was diluted in water (RW1) or in DNA extracted from sludge (RW1 + sludge) or soil (RW1 + soil). Copy numbers calculated from the unspiked sludge and soil samples were less than the limit of quantitation. Data are shown on a log-log plot for ease of visualization. Each point is an average of three replicates; error bars representing the standard error may be smaller than the marker. Insets show the slopes of the respective dilutions and the qPCR standard curve for the designed primer set (321fwd-451rev). Standard curves were generated using a recombinant plasmid containing the gene encoding for the dioxin dioxygenase. Results shown here are typical of several qPCR runs.

To determine the copy number of the pSWIT02 megaplasmid, qPCR was also performed targeting the 16S rDNA (Maeda et al., 2003). There are two copies of this gene on the chromosome and none on either megaplasmid (Miller et al., 2010). A comparison of *dxnA1* to 16S rDNA copy number in *S. wittichii* RW1 grown on dibenzofuran showed a ratio of 1.0 ± 0.1 (average \pm standard error, $n=19$), implying a 2.0 ± 0.2 ratio of pSWIT02 to chromosome.

Microcosms were created to test the survival of *S. wittichii* RW1 in landfill leachate. The leachate was known to contain acetone, benzene, 2-butanone, 1,4-dichlorobenzene, dichloroethane, dichloroethene, dichloropropane, ethylbenzene, methylene chloride, 4-methyl-2-pentanone, tetrachloroethene, toluene, trichloroethene, vinyl chloride, and xylenes in the $\mu\text{g/L}$ range. Dibenzofuran from a methanol stock solution was added to sterile test tubes to a final concentration of 0.5 mg/ml. Landfill leachate was pasteurized to inactivate endogenous microorganisms by heating to 65°C for one hour in a water bath in sealed containers to prevent evaporative compound losses, and added to the test tubes after cooling to room temperature. Microcosms were inoculated with 50 μl of a pure culture of *S. wittichii* RW1, resulting in an approximate density of 10^7 CFU/ml, which is comparable to previous studies (Halden, Halden, et al., 1999b). This high inoculum is necessary because the population declines as the contaminant is transformed (Halden, Halden, et al., 1999b; Nam et al., 2005), likely because the chlorinated metabolites, especially chlorocatechols, can inhibit enzymes in this and other catabolic pathways (Halden, Halden, et al., 1999b; Seah

et al., 2007). The tubes were incubated horizontally at room temperature without agitation and sampled daily with triplicates taken 3, 6, 9, 12, and 14 days post inoculation. Colony counts were performed by spreading dilutions on Luria broth (LB) agar plates. DNA was extracted with the NucleoSpin Soil kit and diluted 1:10 in water for qPCR. Fluorescence signal from uninoculated leachate was not significantly different from the non-template control.

Total DNA, *dxnA1* copy numbers, and colony counts all showed a similar trend (Fig. 3.2), indicating that the population of *S. wittichii* RW1 increased initially and then was stable. Results from the qPCR assay showed less variation than colony counts because of the greater specificity of the non-culture based assay that was insensitive to a background of leachate-endogenous microorganisms that survived pasteurization and persisted at levels at one or more orders of magnitude less than *S. wittichii* RW1.

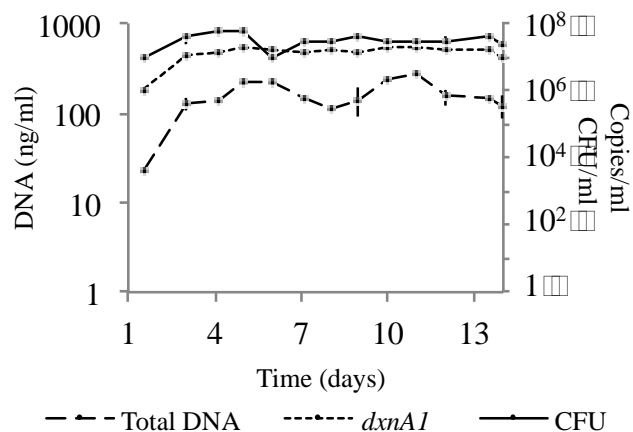


Figure 3.2. Measurement of total DNA, *dxnAI* copy number, and colony counts from *Spingomonas wittichii* RW1-bioaugmented landfill leachate microcosms. Days 3, 6, 9, 12, and 14 are the average of biological triplicates; error bars show the standard error. All other measurements are single data points. Measurements are normalized per ml of culture.

In the leachate microcosms, there were 0.5 ± 0.1 (average \pm standard error; $n=13$) *dxnAI* copies per CFU, which is lower than the pure culture copy number, suggesting that a part of the RW1 community is making use of C-sources in leachate other than dibenzofuran. The correlation between colony count and *dxnAI* was poor ($R^2 = 0.26$). Other studies have also noted discrepancies between colony counts and qPCR assays (Ludwig & Schleifer, 2000). That *dxnAI* is on a megaplasmid may contribute to the lack of correlation, as not every cell contains the plasmid. While the percentage of the population carrying the plasmid varies, it is unlikely to be transferred to other organisms, even in the presence of a selective advantage (Basta et al., 2004a).

The *dxnAI* qPCR results were consistently lower than the colony counts, which is expected if plasmid copy number is less than one. This observation

suggests that biodegradative genes can be lost, so culture-based assays or methods targeting chromosomal genes, *e.g.* 16S rDNA, may overestimate biodegradative capability when the necessary genes are plasmid-based, as is often the case with sphingomonads (Basta et al., 2004a). Similar observations have been made in the anaerobic world concerning vinyl chloride reductase genes in *Dehalococcoides* (van der Zaan et al., 2010), which are flanked by insertion sequences (Krajmalnik-Brown, Saunders, Ritalahti, & Löffler, 2007) and therefore also subject to loss in the absence of positive pressure (Futagami, Goto, & Furukawa, 2008).

Previous studies looking at *S. wittichii* RW1 activity in contaminated environments have been hampered by a lack of molecular tools to adequately evaluate the bacterial population. Instead, they rely on dioxin removal rates as a proxy for bacterial activity and survival (Halden, Halden, et al., 1999b; Nam et al., 2005). The actual population levels are therefore unknown. We believe this qPCR method will help answer questions about population levels in the face of inhibitory metabolites and the retention of biodegradative capability despite the availability of other carbon sources. We have already provided evidence that biodegradative capability can be lost in a complex medium containing multiple carbon sources by comparing *dxnA1* copy number to colony counts in landfill leachate microcosms. For *in situ* studies where it is impractical to perform colony counts due to the presence of other microorganisms, quantification of *dxnA1*

could be combined with general or *Sphingomonas*-specific 16S rDNA assays to compare population levels and biodegradative capability.

We created a qPCR method for *dxnA1*, the gene encoding for dioxin dioxygenase, an important enzyme in the dioxin degradation pathway of *S. wittichii* RW1. This qPCR method is accurate over 8 orders of magnitude and has a limit of quantitation of 62 copies per reaction. We have demonstrated the specificity of the primers over 4 orders of magnitude, down to 10^2 copies/reaction against environmental DNA backgrounds. Finally, we have used this method to quantify *dxnA1* in environmental samples and, for the first time, showed growth of *S. wittichii* RW1 in landfill leachate, suggesting that the organism could be used to remediate dioxins in contaminated leachates. Our findings indicate that loss of megaplastids occurs in *S. wittichii* RW1 populations under the conditions tested. This underscores the value of the present method for tracking the biodegradation potential of *S. wittichii* RW1 cells released into contaminated environments for remediation purposes. The method illustrated here may be adapted for targeting biodegradative, plasmid-based genes extant in other sphingomonads to accurately gauge their bioremediation potential.

The authors would like to acknowledge Drs. Nicole Hansmeier, Benny Pycke, and He-Ping Zhao for their assistance. This project was supported in part by Award Number R01ES015445 from the National Institute of Environmental Health Sciences (NIEHS). The content is solely the responsibility of the authors

and does not necessarily represent the official views of the NIEHS or the National Institutes of Health.

Transition 3

In Chapter 3, a quantitative PCR method was created for a biomarker that was originally identified using mass spectrometry. That method was then vetted in landfill leachate microcosms, showing its utility in environmental settings. While this gene-based method is immediately useful, it is based on more conventional technology and subject to limitations regarding the inference of phenotype.

In Chapter 4, I return to mass spectrometry and, with the help of my co-workers, develop a method using an **absolute quantitation** technique known as AQUA for the same biomarker discussed in Chapter 3, the dioxin dioxygenase, and for several others, including the norovirus capsid protein and the pathogenic prion protein. The use of AQUA to detect an enzyme of importance for bioremediation is groundbreaking. No such methods have previously been developed for use in the detection of bioremediation agents.

These AQUA methods are tested against a background of potential environmental interferences for all of the assayed targets. The AQUA peptide biomarkers are also tested using two forms of mass spectrometry: electrospray ionization triple quadrupole tandem mass spectrometry and the less conventional matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. This latter instrument is uncommonly used for AQUA, but its robustness is desirable when dealing with potentially dirty environmental samples.

4. Targeted, Quantitative MALDI-TOF MS Analysis of Biological Agents from Multiple Kingdoms in Environmental Samples

With the growing availability of sequence data, the feasibility and advantages of mass spectrometric (MS) protein-based detection are becoming ever more apparent. Such methods can be directed at a wide variety of targets, spanning all kingdoms, and are of use in several scientific disciplines. Pathogenic bacteria (Holland et al., 2000), viruses (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006), and prion proteins (Silva et al., 2011) can all be detected via protein-based methods in the biomedical and public health domains; beneficial bacteria (Halden et al., 2005b) are likewise targeted in environmental engineering applications, such as bioremediation (i.e., the use of bioagents to detoxify chemical pollutants). Such MS-based methods may be preferred over their gene-based counterparts because they impart phenotypic information and can give more insight into traits like infectivity or biodegradative activity.

Protein-based qualitative detection methods can be taken to the next logical and practical level by transforming them into quantitative tools, thereby enabling, for example, risk assessment or an evaluation of bioremediation efficacy. Absolute quantitation represents the ultimate goal, as it allows for a comparison of results across samples and across studies for maximal utility of the findings (Chao et al., 2010). This quantitation can be achieved through the inclusion of stable heavy-isotope labeled standards, sometimes referred to as AQUA peptides.

Electrospray ionization-quadrupole mass spectrometers or ESI-MS instrumentation has been historically favored for use in absolute quantitation of biomolecules. However, this technique is highly sophisticated and sensitive, making it less than ideal for widespread analysis of environmental samples that may contain components adversely affecting instrument performance. Matrix-assisted laser desorption/ionization (MALDI) instruments, often followed by time-of-flight (TOF) or tandem TOF mass spectrometers, provide an alternative to electrospray-based techniques and are much more robust and therefore potentially better suited for detecting targets in complex environmental matrixes. For some types of samples, the analysis can even be performed without the elimination of environmental background. For example, MALDI-TOF MS-based semi-quantitative analysis of tissues has been successfully used to measure peptide hormones (Lippolis et al., 2009). In a comparison test using isobaric tags (iTRAQ) to measure *Escherichia coli* tryptic digests, MALDI and ESI instruments performed comparably well, and samples could be archived on the MALDI target plates.(Kuzyk et al., 2009) For these practical reasons, MALDI-based quantitation should receive more attention.

Electrospray systems are usually coupled to a liquid chromatography (LC) separation, which adds another dimension that can be used to identify a target (retention time) (Anderson et al., 2012) but often represent the main bottleneck in rapid sample throughput due to well established problems with clogging and leakage. Alternative methodological approaches include liquid chromatography-

MALDI systems and MALDI-based detection following immunocapture in lieu of LC (Anderson et al., 2012; R. W. Nelson & Borges, 2011). The latter approach is excellent for concentrating target peptides from complex samples, but relies on antibodies whose availability is limited and requires further method optimization for antibody binding, thereby precluding some applications for practical or economical reasons. Immunocapture-based assays coupled with MALDI-TOF MS quantitation have, however, been successfully employed for a variety of synthetic peptides of human origin (Anderson et al., 2012), 11 diabetes-related proteins (Borges et al., 2011), and allelic variants specific to the sheep prion protein (Morel, Andreoletti, Grassi, & Clement, 2007).

In situations where antibodies are not available, relatively simple sample preparation, such as filtration (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006) or one-dimensional gel electrophoresis (Hartmann et al., 2010), may suffice. These minimal separation techniques do not require highly skilled operators or expensive and toxic solvents, nor do they require the time- and cost-intensive development of antibodies. Similar to LC, these separation methods, including immunoseparation, can provide some information to corroborate the identity of the target. In tandem mass spectrometry (e.g., TOF/TOF) systems, the identity of the target and standard peptides can also be confirmed in the second MS. Fragmentation patterns, along with particle size, molecular weight, or antibody-binding affinity, can compensate for information lost by excluding liquid chromatography from the workflow (e.g., retention time).

While the principles of MALDI-based quantitation using heavy-isotope labeled standards have been demonstrated, applications outside of the biomedical field are rare. In the present study, we examined the use of heavy-isotope labeled standards for the MALDI-TOF MS-based quantitation of proteins from several model systems in challenging environmental sample matrixes representing real-world complexities. From eukaryotic systems, we chose the prion protein, an agent widely associated with scrapie in sheep and bovine spongiform encephalitis. The target peptide is conserved for ovine, bovine, and cervine versions of the pathogenic, misfolded protein (e.g., GI:34733623, GI:37576785, GI:4235624). As a representative virus, we chose norovirus, a member of the calicivirus family and the causative agent of the majority of waterborne illness in the United States. Norovirus is also a class B bioterrorism agent. The target peptide used in this study comes from the capsid protein (GI:34223984), which protects the viral genetic material from environmental degradation and plays a role in the infection process. From the bacterial kingdom, we chose *Sphingomonas wittichii* RW1, an α -proteobacterium valued for its ability to transform polychlorinated aromatic compounds (dioxins). The target peptides come from the dioxin dioxygenase (GI:3426122), the initial enzyme in the dioxin degradation pathway. This workflow was also attempted for oocysts of *Cryptosporidium parvum*, a single-celled eukaryotic parasite that was responsible for the largest waterborne disease outbreak in recorded U.S. history. The goal of this study was to determine the

range of targets and different environmental matrices amenable to analysis by MALDI-TOF MS for absolute quantitation.

Experimental

Peptides. Peptide standards (Table 4.1) were synthesized by AnaSpec (Fremont, CA) to be > 95% purity and include one $^{13}\text{C}_6$, ^{15}N -leucine or $^{13}\text{C}_5$, ^{15}N -valine residue. Lyophilized peptides were resuspended in 0.1% trifluoroacetic acid (TFA).

Samples. The prion peptide was synthesized by AnaSpec to be >95% purity and handled as the heavy-labeled peptide standards. Norovirus-like particles were kindly provided by the laboratory of Dr. Mary Estes. Particle quantity was determined by bicichoninic acid (BCA) assay (Pierce, Rockford, IL), quality by transmission electron microscopy with negative staining, and purity by sodium dodecylsulfate polyacrylimide gel electrophoresis (SDS PAGE). The dioxin dioxygenase was produced by culturing *S. wittichii* RW1 on minimal medium supplemented with dibenzofuran (Halden et al., 2005b) and partially purified by SDS PAGE. Heat-inactivated *Cryptosporidium* oocysts (Sterling Parasitology Laboratory, Tucson, AZ) were washed twice and resuspended in 50 mM ammonium bicarbonate.

Environmental samples. Groundwater from a perchlorate-contaminated well in Mesa, AZ was collected using a bailer, transported in a cooler, and stored at -20°C until analysis. Aliquots of 5 ml were thawed, filtered and concentrated to

<500 μ l using a filter with a cutoff value of 5,000 nominal molecular weight (Agilent, Santa Clara, CA).

Fresh produce samples, consisting of strawberries and green leaf lettuce, were obtained from a local retail outlet. Samples were processed based on methods described by Tahk et al. (Tahk et al., 2012). Strawberries and lettuce were cut to produce 5 g aliquots, which were washed with approximately 50 ml 0.25 M threonine-0.3 M NaCl on a rotary shaker for 5 h. The elution was then concentrated to <500 μ l using Vivaspin20 filters featuring a 100,000 nominal molecular weight cutoff value (Sartorius Stedim North America, Bohemia, NY).

Non-homogenized, organic whole milk was obtained from a local retail outlet and stored at -80°C. Samples were processed as in Lippolis, et al. (Lippolis et al., 2009): 5 ml aliquots of milk were thawed and centrifuged at 10,000 \times g for 20 min at 4°C. The resulting pellets were washed twice and resuspended in 25 ml chilled Dulbecco's Phosphate Buffered Saline (DPBS) without calcium or magnesium (0.2 g/l KCl, 0.2 g/l KH₂PO₄, 0.8 g/l NaCl, 1.15 g/l NaH₂PO₄; pH = 7.4). Resuspensions were twice overlaid with 15 ml 45% sucrose and centrifuged at 10,000 \times g for 40 min at 4°C. Finally, pellets were washed 6 times with 45 ml DPBS and resuspended in 1 ml of 50-mM ammonium bicarbonate.

Agricultural soil samples from Baltimore, MD (Walters et al., 2010) were processed using the direct soil protein extraction method described by Chourey et al. (Chourey et al., 2010). Briefly, 5 g aliquots of frozen soil were mixed with 10 ml Alkaline-SDS buffer and heated in a boiling water bath for 10 min. The

mixture was then cooled and centrifuged at $2095 \times g$ for 10 min. Proteins were concentrated from the supernatant via a TCA-acetone precipitation and resuspended in 1 ml of 50-mM ammonium bicarbonate.

Protein extraction. Cells in environmental elutions or concentrations were lysed by sonication on ice with a SONIFIER (Branson, Danbury, CT) using 5 bursts of 10 s each with 10 s rests in between at a power setting of 30% max. Lysates were centrifuged at $21,000 \times g$ for 15 min to separate soluble proteins from cell debris. Protein concentrations were determined using the BCA assay; produce samples were analyzed directly with SDS PAGE because the presence of threonine in the sample buffer was incompatible with the BCA assay.

SDS PAGE. Sample aliquots were mixed 1:1 (v:v) with Laemmli sample buffer (Bio-Rad, Hercules, CA) containing 5% β -mercaptoethanol and heated at 95°C for 5 min. After cooling to room temperature, samples were loaded onto 4-20% TGX gels (Bio-Rad, Hercules, CA) and run at 200 V for 35 min. Gels were stained for qualitative analysis using SimplyBlue SafeStain coomassie reagent (Bio-Rad, Hercules, CA) or for mass spectrometry using the Flamingo fluorescent stain (Bio-Rad, Hercules, CA). Fluorescent stained gels were visualized using a UV light box. Bands at 57 kD, 49 kD, and 27 – 30 kD, corresponding to the norovirus capsid protein, the dioxin dioxygenase, and the diagnostic core of the prion protein, respectively, were excised using a razor blade and dried using a SpeedVac for in-gel digestion.

MALDI-TOF/TOF MS. Gel bands were digested overnight at 37°C in a solution containing 10 ng sequencing-grade modified trypsin (Promega, Madison, WI) in 20 µl of 50-mM ammonium bicarbonate. Norovirus capsid protein standards were digested in a solution containing 1:100 trypsin:norovirus and ≥ 35 µl of 50-mM ammonium bicarbonate. Approximately 10^6 washed *C. parvum* oocysts were digested with 500 ng trypsin. After digestion, large particles were removed by centrifugation at $5,000 \times g$ for 5 min. Digests were cooled to room temperature and adjusted to 0.1% TFA. For standard curves, digests of target samples were diluted in 0.1% TFA; to evaluate environmental influences, digests of target samples were diluted in digests of environmental samples, except for the dioxin dioxygenase, which was mixed with the environmental samples prior to SDS PAGE. Standard peptides were added to the acidified digests, and the mixtures were concentrated and desalted using Omix C₁₈ ZipTips (Varian, Palo Alto, CA) and directly eluted with α -cyano-4-hydroxycinnamic acid (LaserBio Labs, Sophia-Antipolis, France) onto a 384-well stainless steel target plate (AB/Sciex, Framingham, MA).

Samples were analyzed using a 4800 MALDI-TOF/TOF MS (AB/Sciex, Framingham, MA). For quantitation, spectra were acquired in positive reflector mode with a fixed laser intensity between 2800 and 3500 arbitrary units. To confirm the identity of the target peaks, MS/MS spectra were acquired in positive mode using post-source decay for fragmentation.

Data analysis. Spectra were acquired using 4000 Explorer software and exported to Data Explorer (AB/Sciex, Framingham, MA). Mass calibration was performed using internal trypsin autolysis peaks or external calibrants (Sigma, St. Louis, MO). Peaks with a signal-to-noise ratio greater than 3 or isotope clusters with a combined signal-to-noise ratio greater than 10 in the relevant mass ranges were exported to Excel (Microsoft) for quantitative analysis.

Results and Discussion

We evaluated the use of MALDI-TOF/TOF MS for the targeted, quantitative analysis of a variety of analytes as model systems of relevance for public health and environmental engineering. Targets representing multiple kingdoms (Fig. 4.1A-D) were chosen based on tryptic peptides from pre-determined biomarkers, where available. For the prion protein, the target peptide is a fragment from the proteinase K-resistant diagnostic core of the misfolded, pathogenic protein (Fig. 4.1A). For norovirus (Fig. 4.1B), the target peptide comes from the capsid protein, without which the viral RNA is no longer infectious and subject to environmental degradation. For *S. wittichii* RW1 (Fig. 4.1C), four potential peptides from the dioxin dioxygenase, the first enzyme in the dioxin degradation pathway, were evaluated. The initial work needed to determine these biomarkers should not be underestimated. For example, this workflow was extended to the sporulated, infective stage of the eukaryote *C. parvum* (Fig. 4.1D) but no representative peptides were identified, presumably due to the oocyst's

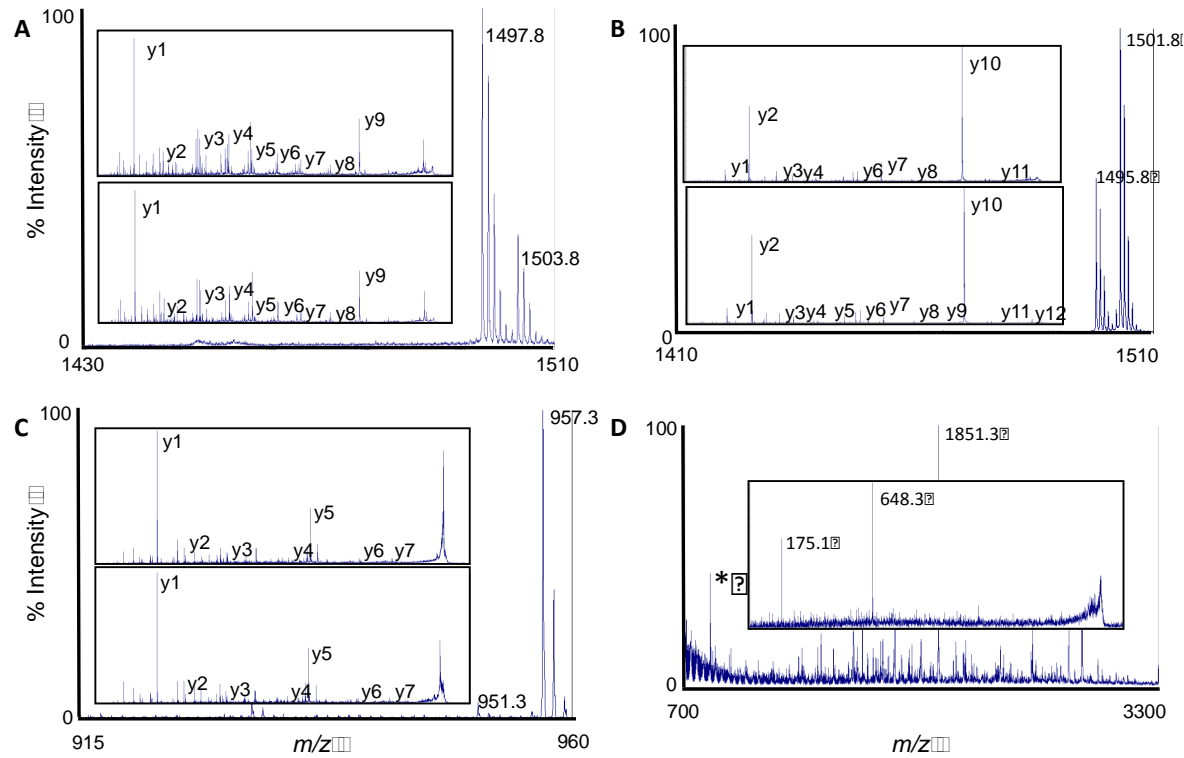


Figure 4.1. Spectra of the diagnostic prion peptide, VVEQMCITQYQR ($m/z = 1497.8$), and its heavy-labeled standard, VV*EQMCITQYQR ($m/z = 1503.8$) (A); the norovirus peptide, TLDPIEVPLEDVR ($m/z = 1495.8$), and its heavy-labeled standard, TLDPIEVPLEDV*R ($m/z = 1501.8$) (B); and the dioxin dioxygenase, GVSEGYIAR ($m/z = 951.5$), and its heavy-labeled standard, GV*SEGYIAR ($m/z = 957.5$) (C). Insets (A-C) show MS/MS of the target peptide (top) and the standard (bottom). The y-series ions are labeled. A spectrum of digested *C. parvum* oocysts and MS/MS of the $m/z = 1851.3$ peak are also shown (D); * indicates a trypsin autolysis peak ($m/z = 842.5$).

recalcitrant, heavily glycosylated cell wall. Whereas an unidentified ion reproducibly yielded a characteristic peak at m/z 1,851.3 for *C. parvum* oocysts, its diagnostic value could not be examined further due to a lack of a stable isotope-labeled analog for use as a quantitative standard.

For all other bioagents, each analyte was quantified by normalizing the observed peak height in the first MS to that of a heavy-isotope labeled standard (Table 4.1). Quantitation was evaluated as a function of the limit of detection, robustness with respect to sample interference from different environmental compartments, and tolerance to differences in the relative amounts of analytes and standards.

Mass spectrometry has already been used to detect the prion protein in tissue samples (Morel et al., 2007; Onisko et al., 2007) and biodiesel (Douma, Kerr, Brown, Keller, & Oleschuk, 2008). In the present study, we evaluated the use of a biomarker previously determined to perform well in nanoLC-MS/MS systems (Silva et al., 2011). In our MALDI system, the limit of detection for this prion peptide was 1 fmol for the pure standard. The detection of this peptide is complicated by the presence of a methionine residue, which can be oxidized during processing. A peak shift of 16 Da, indicative of the incorporation of an oxygen molecule, was observed for both the labeled and unlabeled peptides. The limit of detection for this peptide using LC-MS/MS was reported to be on the order of 100 amol (Silva et al., 2011). The order-of-magnitude discrepancy between the previously reported limit of detection and the one found in this study (1 fmol) may be attributable to differences in the ionization behavior of the

Table 4.1. Peptides used for quantitation.

Peptide	Monoisotopic Mass (Da)	Labeled Monoisotopic Mass (Da)
Prion Protein		
VV*EQMCITQYQR	1497.0	1503.0
Norovirus Capsid Protein		
TLDPIEVPLEDV*R	1495.8	1501.8
<i>Sphingomonas wittichii</i> RW1 Dioxin Dioxygenase		
GV*SEGYIAR	951.5	957.5
GL*IFGNWR	962.5	969.5
L*GHASSGFFK	1050.5	1057.5
SWL*FLGHESQIPK	1541.8	1548.8

* = heavy-isotope labeled residue

peptide during the differing ionization techniques. One of the pitfalls of mass spectrometry-based detection is that methods optimized for one type of instrumentation (e.g., ESI) may not translate to another (e.g., MALDI).

Furthermore, the limit of detection observed for one particular peptide may not be reproducible using a different instrument even when employing the very same ionization technique and operational settings. For example, in the current study the norovirus peptide could be detected at levels as low as 100 amol for the pure standard. This limit of detection is 3 orders of magnitude lower than previously reported (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). The differences can likely be attributed to the eschewal of

probability-based peptide mass fingerprinting, which relies on the detection of multiple peptides, but it can also be linked to improvements in instrumentation.

A comparable limit of detection (500 amol) for other characteristic norovirus capsid protein peptides was observed in prior work using LC-MS/MS (D. R. Colquhoun, 2007). The detection of 100 – 500 amol of the target peptide implies the presence of 10^5 – 10^6 virus particles. While these limits are still several orders of magnitude above those of gene-based detection methods, they show a dramatic improvement over other protein-based detection methods (Hartmann & Halden, 2012).

Of the 4 peptides analyzed from the *S. wittichii* RW1 dioxin dioxygenase, the most reliable were GVSEGYIAR and GLIFGNWR. Detection of the remaining two peptides was unreliable. In mixtures where the standards were all present at equimolar concentrations, the GVSEGYIAR and GLIFGNWR peaks were consistently higher than the LGHASSGFFK or SWLFLGHESQIPK peaks. This observation may reflect differences in ionization, a property that is hard to predict. Because the dioxin dioxygenase was purified using SDS PAGE, the limits of detection for these peptides were constrained by those of the protein stain. These limits of detection corresponded to approximately 100 and 10 fmol for GVSEGYIAR and GLIFGNWR, respectively.

Quantitation using mass spectrometry is most successful when the amounts of target analyte and standard are within the dynamic range of the instrument and also reasonably close to each other. To test the dynamic range of the instrument, serial 10-fold dilutions of the target analyte were mixed with

parallel dilutions of standard peptides and analyzed (Fig. 4.2). Dilutions were performed in 0.1% TFA or in a background of environmental extractions. In the ideal scenario, the height ratio of the target peak and the standard peak should remain constant regardless of the amount of target and standard added. Deviations from this constant ratio indicate the limit of detection at the lower end and detector saturation at the upper end.

Transmission of prion-mediated diseases is primarily a concern in edible products and animal feed, although other byproducts of the meat industry have also been investigated (Douma et al., 2008). We tested the prion protein biomarkers against a background of foods (lettuce, strawberry, and milk) as well as soil, which could serve as an environmental reservoir. Norovirus, the leading cause of food- and waterborne illness, can often be spread via uncooked foods or through contact with contaminated media (e.g., soil or water) (Hartmann & Halden, 2012). We therefore tested the norovirus biomarker in the context of lettuce, strawberry, milk, soil, and groundwater. *S. wittichii* RW1 was isolated from a river (Wittich, Wilkes, Sinnwell, Francke, & Fortnagel, 1992b) and has been used to remediate dioxin-contaminated soil (Halden, Halden, et al., 1999b), among other media. We therefore tested the dioxin dioxygenase biomarker in soil and contaminated groundwater.

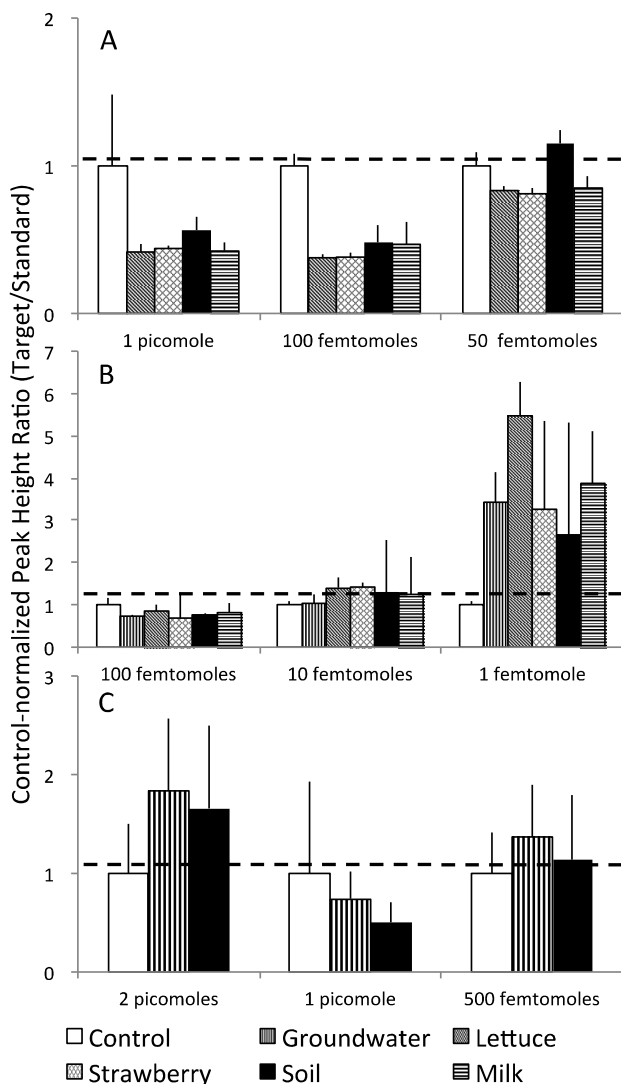


Figure 4.2. To test the dynamic range of the method, serial 10-fold dilutions of target analytes with parallel dilutions of standard peptides in a variety of environmental backgrounds. Controls were diluted in 0.1% trifluoroacetic acid for the prion protein (A) and norovirus capsid protein (B) or 50 mM ammonium bicarbonate for the dioxin dioxygenase (C). Values are presented as averaged ratios, normalized to the average measurement for the control ($n = 3$ to 12). Error bars represent the standard deviation; for clarity, only positive error bars are shown. The dashed line shows the ideal scenario where all calculated ratios are equal, regardless of the presence of environmental background.

For the prion protein, the inclusion of an environmental background, regardless of its source, increased the limit of detection to 50 fmol (Fig. 4.2A). While the possible oxidation of methionine may influence the limit of detection, it does not appear to affect the accuracy, indicating that the target and isotopically labeled peptides are oxidized at the same rate. For the norovirus capsid protein the limit of detection against an environmental background was 1 fmol (Fig. 4.2B). The limit of detection for the dioxin dioxygenase, being largely determined by the dynamic range of the upstream preparation, was not affected by the inclusion of an environmental background (Fig. 4.2C). Significant differences were observed between samples diluted in the various environmental backgrounds, but regardless of the inclusion of a background, the results are consistent within an order of magnitude and, in most cases, a factor of 3.

In the ideal scenario for quantitation based on heavy-isotope labeled standards, the standard and the analyte would be present at relatively similar (equimolar) amounts. The ratios of peaks from analytes with the same sequence have been shown to deviate from the ideal predicted value as the ratio of the amount of analyte deviates from 1 in both ESI- and MALDI-based systems (Kuzyk et al., 2009). However, sample availability or other constraints may preclude the optimization necessary to determine the proper ratio of analyte to standard. To test the tolerance of this method to differences in target and standard amounts, we analyzed 10-fold serial dilutions of the standards mixed with a constant amount, approximately 50 fmol for the prion protein and norovirus capsid protein and 500 fmol for the dioxin dioxygenase (Fig. 5.3). Ideally, the

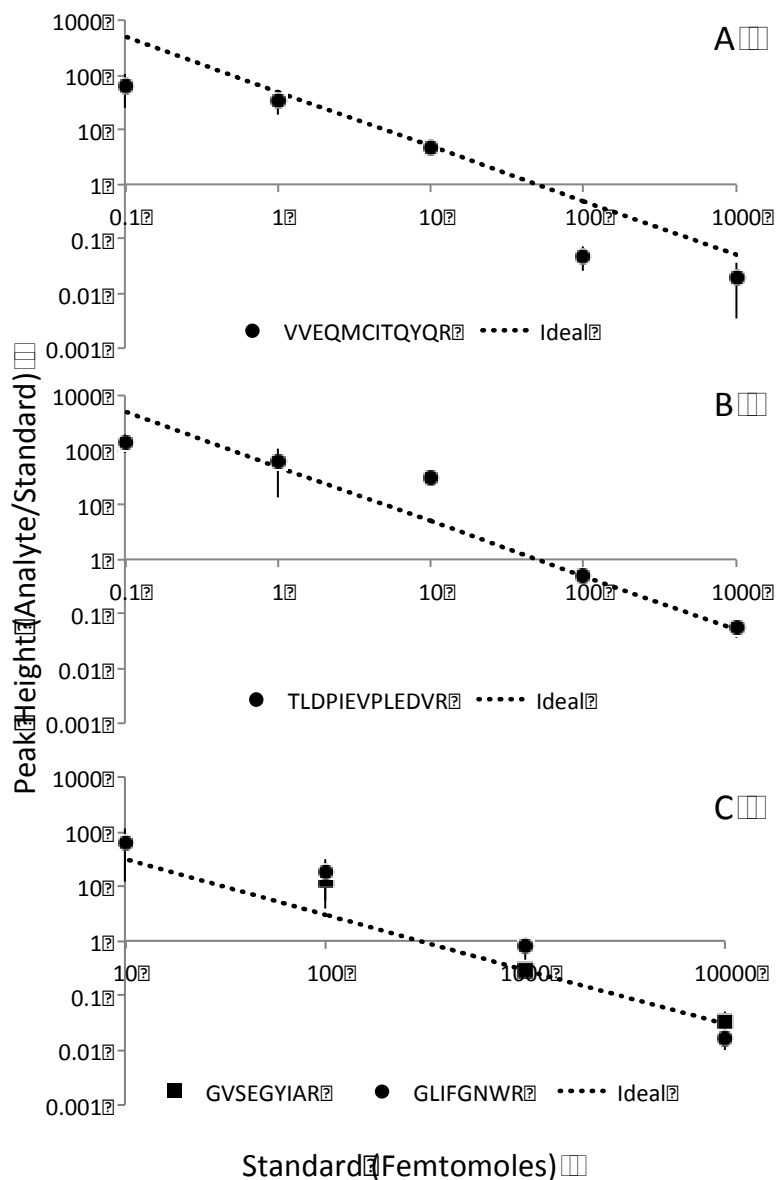


Figure 4.3. To test the tolerance of the method to differences in the relative amounts of target and standard, serial 10-fold dilutions of standards were mixed with a constant amount, approximately 50 fmol for the prion protein (A) and norovirus capsid protein (B) and 500 fmol for the dioxin dioxygenase (C), of target. Values are presented as averaged ratios ($n = 3$ to 5). Error bars represent the standard deviation. The dotted line represents the ideal ratio predicted from the best estimate of the actual target mass.

ratio of the height of the target peak to the height of the standard peak should vary linearly with the amount of standard added. Deviations from linearity indicate peak suppression, which is an inherent limitation of MALDI-TOF MS, or competition for binding sites in the extraction process directly preceding the MS analysis. While none of the curves follow the ideal exactly, this analysis shows that it is possible to roughly quantify an unknown agent by targeting a specific peptide to within an order of magnitude even with a 2-order of magnitude difference between the amounts of the analyte and the standard present.

Conclusions

The use of heavy-isotope labeled standards combined with MALDI-TOF MS is rich with untapped potential for absolute quantitation of targets ranging across multiple biological kingdoms. Using this technique, we have demonstrated detection limits as low as 1 fmol against an environmental background, and improvements in upstream sample cleanup could lower the limit of detection to 100 amol, the limit achieved for pure standards. This method has lenient requirements for the analyte-to-standard ratio, an important characteristic for applications where sample or standard availability are limited. Due to the rapid processing time and applicability to a wide range of targets (prions, viruses, and bacteria) and environmental sample types (food, soil, and groundwater), this method of quantitation has a strong potential for use in public health and environmental sciences.

Transition 4

Chapters 2 through 4 covered the development of various detection methods for bacteria, viruses, and single-celled eukaryotes. Chapter 2 describes a qualitative mass spectrometry-based method to quickly and easily identify protein biomarkers. Chapter 3 shows the design and implementation of a quantitative PCR-based method to track a functional, plasmid-encoded genetic biomarker in bioaugmented landfill leachate. Chapter 4 delves into the use of quantitative mass spectrometry to detect protein biomarkers in environmental samples, including soil, food, and groundwater.

In this dissertation, I have thus covered a variety of possible techniques for detecting biomarkers of single-celled organisms in the environment. However, the full range of possible methods is much wider. The concluding chapter is intended as an objective comparison of these and other methods in the context of viruses in food. Although this chapter focuses exclusively on virus detection methods, these methods share many qualities with methods for detecting bacteria. The desired outcome is always to detect a target within a complex background of non-target signals and ideally to extract some sort of phenotypic information from the results, whether its biodegradative activity or infectivity. As can be seen from Chapter 4, the same methods can be tailored to detect bacteria, viruses, and prion proteins. The lessons learned from virus detection methods are therefore also applicable to the field of bioremediation and any other discipline that may involve environmental detection of biomarkers.

In this fifth chapter, I review gene- and protein-based detection methods for four types of viruses found on the third Contaminant Candidate List issued by the U.S. Environmental Protection Agency. Based on the advantages and disadvantages of the available technologies, I make suggestions for future research and development in the area of the detection of biological agents in environmental compartments. These recommendations, which are also valid for the detection of bacteria in bioremediation, are that we should continue the trend towards quantification, standardize methods to facilitate inter-study comparisons, and pursue detection methods whose outcomes can be directly linked to infectivity or biodegradation.

5. Analytical Methods for the Detection of Viruses in Food by Example of CCL-3 Bioagents

The U.S. Environmental Protection Agency released its third Contaminant Candidate List (CCL-3) in September 2009 (Richardson, 2010). The CCL-3 features over 100 chemicals, but it also includes 12 microbiological contaminants. Among these emerging biocontaminants are three groups of viruses: adenoviruses, caliciviruses, and enteroviruses. It also includes hepatitis A virus, which is a picornavirus. Owing to their potential impact on food and water safety, these viruses can also be considered class B bioterrorism agents, according to the Centers for Disease Control and Prevention (CDC) ("Bioterrorism Agents/Diseases,"). In daily life, the majority of these viruses are spread through the fecal-oral route and cause gastrointestinal or respiratory illness, with the exception of hepatitis A, which causes liver disease and jaundice. At the national level, hepatitis A is the most commonly detected CCL-3 virus (European Centre for Disease Prevention and Control, 2011; Scallan et al., 2011). However, norovirus is estimated to be responsible for the majority of foodborne illnesses (Scallan et al., 2011). When an illness affects two or more people, the CDC defines that event as an outbreak. However, outbreaks are not reported to the CDC unless they affect people in multiple states; outbreaks confined to a single state are the domain of local and state health authorities ("Surveillance for Norovirus Outbreaks,"). Outbreaks of viral gastroenteritis, such as are caused by the CCL-3 viruses, are rarely reported to the federal agency. CCL-3 virus families and other pertinent information are given in Table 5.1.

For adenoviruses and caliciviruses, no standardized methods for culturing or detection exist (United States Environmental Protection Agency, 2008). There are standardized and validated procedures for culturable coxsackieviruses and echoviruses, both members of the enterovirus family (United States Environmental Protection Agency, 2008). However, these methods are not capable of distinguishing between virus types and are not applicable for all enteroviruses. Detection is challenging for several reasons. Viruses have a relatively high mutation rate, so a variety of genera, species, genotypes, and serotypes are contained within these four virus groups. Not all of these viruses are culturable, and those that are culturable do not necessarily propagate in the same cell type; and even culturable viruses may be difficult to cultivate or unreliable at producing cytopathic effects (Duizer et al., 2004; M. Kim, Lim, & Ko, 2010; Konduru & Kaplan, 2006; Straub et al., 2011; T. M. Straub et al., 2007; Timothy M. Straub et al., 2007; Zoll et al., 1992). Finally, while the infectious dose is not known for all CCL-3 viruses, it is generally thought to be between one and 10 virions (Bosch et al., 2011; Koopmans & Duizer, 2004).

Table 5.1. Viruses on the third Contaminant Candidate List.

Virus*	Symptoms	Chromosome	Culture Method	Reference
Adenovirus type 40 type 41	gastroenteritis, respiratory disease	dsDNA	293, A549, PLC/PRF5, Caco-2	(M. Kim et al., 2010)
Calicivirus norovirus sapovirus	gastroenteritis	ssRNA	3-dimensional Caco-2	(Duizer et al., 2004; Straub et al., 2011; T. M. Straub et al., 2007; Timothy M. Straub et al., 2007)
Enterovirus coxsackievirus echovirus poliovirus	various	ssRNA	BGM, RD, HEL	(Zoll et al., 1992)
Hepatitis A virus	hepatitis	ssRNA	HeLa, FRhK-4, GL37, Vero, CHO, MMH- D3, Huh7	(Konduru & Kaplan, 2006)

293, human embryonic kidney; A549, human lung epithelium; PLC/PRF5, human hepatoma; Caco-2, human epithelial colorectal adenocarcinoma; BGM, buffalo green monkey kidney; RD, rhabdomyosarcoma; HEL, human embryonal lung fibroblasts; HeLa, human cervix epithelium; FRhK-4, fetal rhesus monkey kidney; GL37, Vero, African green monkey kidney; CHO, Chinese hamster ovary; MMH-D3, murine hepatocytes; Huh7, Human hepatoma.

* Important subtypes that are common causes of food borne illness are also listed.

Naturally, much research has concentrated on the detection and quantification of these CCL-3 viruses in water (Bae & Schwab, 2008; Blaise-Boisseau, Hennechart-Collette, Guillier, & Perelle, 2010; Bosch et al., 2011; Gibson & Schwab, 2011a; Hamza, Jurzik, Überla, & Wilhelm, 2011; Schwab, DeLeon, & Sobsey, 1996; Seitz et al., 2011). However, transmission can also occur *via* food, which can become contaminated by unclean irrigation water (Bosch et al., 2011) or directly from infected food handlers (Koopmans & Duizer, 2004). Subsequent treatment may be insufficient to decontaminate many ready-to-eat products, and these viruses are remarkably stable. Norovirus, a member of the calicivirus family, can be detected in water after three years of storage and can cause infection in humans after 61 days in water (Seitz et al., 2011). Products of chief concern are leafy vegetables (Diez-Valcarce, Cook, Hernandez, & Rodriguez-Lazaro, 2011; Safaa Lamhoujeb, Fliss, Ngazoa, & Jean, 2008; Kirsten Mattison & Bidawid, 2009), soft fruits (Diez-Valcarce et al., 2011; Kirsten Mattison & Bidawid, 2009; Park, Cho, Jee, & Ko, 2008), and shellfish (Diez-Valcarce et al., 2011; Greening & Hewitt, 2008; Kou, Wu, Zhang, & Fan, 2006; Kirsten Mattison & Bidawid, 2009; Nappier, Graczyk, Tamang, & Schwab, 2010; Schwab, Neill, Estes, Metcalf, & Atmar, 1998), but the full range is much larger and includes, for example, deli meats (Safaa Lamhoujeb et al., 2008; Kirsten Mattison & Bidawid, 2009; Schwab et al., 2000). While the exact rates of contamination are unknown, norovirus infections are commonly acquired from uncooked vegetables, fruits, and shellfish ("Surveillance for Norovirus Outbreaks,"). The majority of all known foodborne outbreaks, including those of

bacterial origin, in Europe were attributed to eggs, buffet meals, and juices (European Food Safety Authority, 2012).

In the present review, we summarize the state of sample acquisition, processing, and detection methods for these viruses in food with emphasis on recently developed strategies or technologies. Concentration methods discussed include precipitation, heat release, antibody-based concentration, and filtration; detection methods span microscopy, polymerase chain reaction (PCR), nucleic acid sequence-based amplification, and mass spectrometry. Finally, suggestions are made as to where methods could be combined or improved to advance the field of virus detection in food.

For clarity, virus names written without a species qualifier refer to the strain affecting humans.

Sample Acquisition

The detection of emerging contaminant viruses in food raises interesting questions regarding sampling frequency, methods, and volumes. As previously mentioned, foods that may need to be screened for viruses include all manner of leafy vegetables, soft fruits, shellfish, and processed ready-to-eat foods. These foods can become contaminated at several points. For example, fecally impacted water can contaminate vegetables and fruits if it is used for irrigation (Bosch et al., 2011). Shellfish can become exposed if the water in their habitat is contaminated. And any food can become contaminated at any point if it is handled

by a person who is shedding virus, whether that person shows symptoms of viral infection (Koopmans & Duizer, 2004).

Briefly, viruses are collected from leafy vegetables and soft fruits by elution, and digestive organs are excised for virus collection from shellfish. Because of the enormous potential for contamination and the wide variety of food types, it is difficult to stipulate exactly how and when a particular food needs to be sampled and how much of it constitutes a representative fraction. There are no definitive answers to these questions and regulations are lacking. The details of this area of concern for quality control and monitoring purposes fall outside the scope of this review, but they have been adequately discussed elsewhere (Bosch et al., 2011; Koopmans & Duizer, 2004).

Sample Processing

Once the sample has been acquired, it is in the form of some sort of eluent, homogenate, or tissue. From this point on, two things must be achieved: the viruses, which are likely to be present at very low levels, must be concentrated, and any interfering substances need to be removed or diluted. Many strategies have been developed to fulfill these needs. We limit our discussion to polyethylene glycol (PEG) precipitation, antibody-based concentration, and filtration. We have compared these strategies based on recovery efficiency and time (Table 6.2), although other factors—such as sample and detection compatibility, availability of materials (especially antibodies), and cost—may come into play.

Any sample processing will result in sample loss. In contrast to clinical samples, wherein viruses are shed at remarkably high titers, viruses may be present at levels near their respective limit of detection for the various methods, and low recovery rates may result in false negative results. Recovery efficiency can be calculated using known quantities of virus or a sample process control (see Interpretation), although the exact efficiency will depend on the analyte as well as the type of food from which it is being extracted. For this reason, we present a range of published sample recovery efficiencies.

For ease of implementation, all methods should ideally be rapid, automatable, and high-throughput. Engineering innovations can likely help with the latter two qualities, but for some processing methods, the time required may be unavoidable and prohibitive. Sample processing times have been approximated for all of the methods discussed here, and actual times may vary.

To help reduce false positive signals resulting from damaged, nonfunctional virions, samples may be subjected to enzymatic digestions with proteases, ribonucleases (RNases), or a combination of both. We present some information on the effectiveness of this practice and its potential impacts on processing time and sample recovery.

Table 5.2. Virus concentration methods.

Concentration method	Recovery efficiency	Time	Compatibility	References
PEG	8.5 - 85%	6 - 21 h	NA & protein	(Park et al., 2008; Tahk et al., 2012)
Immunoconcentration	15 - 30%	5 - 35 h	NA & protein	(Casas & Sunen, 2002; Park et al., 2008; Schwab et al., 1996; Tahk et al., 2012)
Filtration	15 - 84%	1 h	NA & protein	(D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006; Gibson & Schwab, 2011a; Tahk et al., 2012)
Proteinase K digestion	37 - 66%	2 h	NA	(Greening & Hewitt, 2008)

Finally, nucleic acids or proteins must be extracted from the processed sample. A plethora of commercial kits are available for the extraction of nucleic acids, generally based on the same principles. Protein extraction methods tend to be custom tailored to specific targets. Because individual research groups use extraction methods based on their own preferences, and recovery efficiencies are usually reported for the method as a whole and not for each individual step, it is difficult to distinguish between analyte loss due to processing and efficiency losses associated with the extraction of nucleic acids or proteins. For that reason, extraction efficiencies are not discussed separately here, although the choice of reagents can surely affect the recovery efficiency of the entire procedure.

Polyethylene glycol (PEG) precipitation. The addition of PEG to a solution causes the viruses therein to flocculate, allowing them to be collected by centrifugation. This standard method for virus isolation is widely used because it works with any virus type and is inexpensive. However, PEG precipitation is extremely time-consuming, as samples must be incubated for extended periods of time. Furthermore, sample recovery is not uniform for all food matrices. For example, it has been found that PEG precipitation is ineffective for some leafy vegetables (Tahk et al., 2012). Recovery also varies with the elution buffer used to wash the sample. One study using strawberries had an average recovery efficiency of 85% when the berries were washed with 3% Bacto beef extract, whereas performance dropped to 8.5% when fruits were washed with 100 mM Tris-HCl or 50 mM glycine-50 mM MgCl₂ (Park et al., 2008). While this method

is appropriate for experiments focusing on downstream detection, it is highly impractical for widespread use in food quality control.

Antibody-based concentration. Concentration methods involving antibody-capture of intact virus particles have been developed for enteroviruses, norovirus, and hepatitis A. These methods have been tested for a variety of foodstuffs, including fresh vegetables, soft fruits, and shellfish (Kirsten Mattison & Bidawid, 2009; Park et al., 2008; Tahk et al., 2012). Proponents of immunological concentration postulate that these methods will help eliminate false-positive detections due to noninfectious, uncontained nucleic acids because only intact viruses with correctly expressed antigens will be collected (Casas & Sunen, 2002; Schwab et al., 1996). Indeed, heat inactivation of viruses resulted in a decreased recovery using immunocapture when compared to heat release (Schwab et al., 1996).

However, direct correlations between samples prepared using immunological concentration and culture-based infectivity assays have been inconsistent (Casas & Sunen, 2002; Hwang, Leong, Chen, & Yates, 2007; Rodriguez, Pepper, & Gerba, 2009; Schwab et al., 1996). A novel detection method, that was based on antibody binding of the target, was tested on bacteriophages and yielded detection values that were three orders of magnitude lower than those obtained on identical samples when using the infectivity plaque assay (Shirale et al., 2010). This finding underscores that antigenicity is not conclusive proof of infectivity.

Concern also exists regarding the specificity of these methods. Critics warn that not all strains of a given enteric virus will present the same antigens, so antibody-based concentration methods may exclude a subset of viruses (Hamza et al., 2011; Kirsten Mattison & Bidawid, 2009). The use of polyclonal antibodies should address this issue, at least in part; such methods have already been used to concentrate simultaneously a variety of viruses or virus genogroups (Casas & Sunen, 2002; Park et al., 2008; Schwab et al., 1996).

One recent investigation using hepatitis A-specific antibodies to concentrate viruses from vegetables found that the method had a poor recovery efficiency compared to PEG precipitation and ultrafiltration (Tahk et al., 2012). Similarly, a study investigating the recovery of norovirus from strawberries had a recovery efficiency of 15 to 30% using immunomagnetic separation with polyclonal antibodies against norovirus capsid protein, compared to 85% using PEG precipitation (Park et al., 2008). In both studies, viruses were detected only by a PCR-based method, and infectivity was not considered, so the observed lower recovery efficiency may be due to the exclusion of noninfectious genetic material. Such results would be expected if immunoconcentration does in fact decrease the rate of false-positive PCR results; investigators therefore should keep this in mind when evaluating antibody-based methods.

Immunoconcentration methods take between approximately 5 and 35 hours from start to finish; this broad range is largely due to the length of the incubation period (Casas & Sunen, 2002; Park et al., 2008; Schwab et al., 1996).

Recovery efficiency and specificity are affected by changes in incubation time, among other factors (Park et al., 2008).

Filtration. Physical separation methods, such as filtration, can be used to concentrate viruses and bacteria simultaneously. For example, tangential flow ultrafiltration, wherein the sample flows parallel to the membrane, was used to collect bacteriophages (MS2 and PRD1), murine norovirus, poliovirus, and several bacterial species from surface and drinking waters (Gibson & Schwab, 2011a). Recovery efficiencies ranged from 15.7 to 83.7% depending on the species in question and the original seeded concentration.

Ultrafiltration was also used to recover hepatitis A from various leafy vegetables (Tahk et al., 2012). This method takes approximately 6 hours, but the bulk of that time is devoted to eluting the virus from the food; filtration takes only 10 to 40 minutes. In said study, ultrafiltration outperformed immunomagnetic separation and PEG precipitation, but was more effective for Chinese cabbage than sesame leaf or lettuce (Tahk et al., 2012).

Filtration methods do not necessarily need specialized equipment. One method using only a degreasing step followed by passage through a 0.22- μ m syringe filter had a recovery efficiency of 72% for bacteriophage MS2 in stool samples and could be completed in under three hours (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006).

Short processing times make filtration an attractive method for routine assessments, but extreme variability of the recovery efficiencies makes it especially important to use process controls (see Interpretation). Attention should

be paid to the relative recovery efficiencies of different viruses and surrogates to compensate for that variability, if possible.

Enzymatic digestion. Treatment with proteases or nucleases is one way to reduce the rate of false-positive detections caused by noninfectious materials. It is assumed that enzymatic digestion will affect damaged capsids or unprotected genetic material. For this reason, signals observed following enzymatic digestion are said to be from "putatively" infectious viruses (S. Lamhoujeb, Fliss, Ngazoa, & Jean, 2009). Treatments may use simultaneous protease and RNase (Nuanualsuwan & Cliver, 2002), sequential protease and RNase (S. Lamhoujeb et al., 2009), or RNase only (Seitz et al., 2011).

Several studies evaluating the efficacy of disinfection protocols or virus survival have used enzymatic treatment to eliminate inactivated virions. Simultaneous proteinase K and RNase digestion eliminated positive signals from hepatitis A, poliovirus type 1, and feline calicivirus inactivated with UV, hypochlorite, and thermal treatments (Nuanualsuwan & Cliver, 2002).

Sequential proteinase K and RNase digestion was used to determine infectious norovirus survival on lettuce and deli turkey (Safaa Lamhoujeb et al., 2008). Heat inactivated norovirus was tested both with and without enzymatic digestion, and only the samples treated with enzyme produced negative results. RNase treatment was used to eliminate unencapsulated RNA from samples in a study regarding the long-term survival of norovirus in water (Seitz et al., 2011). The authors observed significantly lower virus titers following RNase treatment, suggesting that the treatment did in fact help eliminate noninfectious material.

However, enzymatic digestion may not be effective at degrading all noninfectious viruses or a particular virus inactivated by all methods (Cliver, 2009; Nuanualsuwan & Cliver, 2002). Different enzymes or digestion conditions may be more appropriate for certain types of samples. For example, in one study higher concentrations of both proteinase K and RNase were needed to eliminate false positives from heat-killed norovirus than for feline calicivirus (Safaa Lamhoujeb et al., 2008). It is possible to tailor these treatments to the sample, but that requires *a priori* knowledge of the target and the inactivation method, which of course are unavailable when screening samples of unknown content. Furthermore, additional modifications will increase the processing time. For example, sequential treatment with proteinase K and RNase adds 2.5 hours of incubation time (S. Lamhoujeb et al., 2009). Finally, prolonged exposure to elevated temperatures required for functional enzymes may inactivate some viruses (Rodriguez et al., 2009). Such effects would produce artifacts, leading to underestimation of the infectious virus titer.

Protease digestion can also be used to extract viruses from oysters (Greening & Hewitt, 2008). Because mollusks are filter feeders, they naturally concentrate viruses from the surrounding water in their digestive tissue. This obviates the need for additional virus concentration in the sample preparation. Proteinase K digestion yields a recovery efficiency of about 50% (Greening & Hewitt, 2008). It is important to note that the addition of a digestion step to any protocol will affect the recovery efficiency. Furthermore, the use of proteases may affect detection methods directed at the capsid.

Detection

Virus detection methods can be largely grouped into three categories: visual inspection (microscopy), gene-based, or protein-based. For some viruses, infectivity assays can be performed on cell lines that are then examined for cytopathic effects. However, cell culture systems do not necessarily exist for all of the viruses in question. Furthermore, methods that rely on cell culture tend to be labor- and time-intensive and therefore not suitable for routine monitoring applications.

In the absence of rapid, robust infectivity assays, several molecular methods have been developed. In the creation of most of these methods, the chief concerns have been sensitivity and specificity, the latter being especially important for epidemiological source tracking. As a consequence, some methods will detect genetic material or proteins that are not associated with intact or functional virions. Thus, they are prone to giving false-positive results, in the sense that they will react positively even if the sample is not infectious.

We review electron microscopy, PCR, nucleic acid sequence-based amplification, PCR-electrospray ionization-mass spectrometry, Western blots, and mass spectrometry-based detection of proteins. These methods are compared on the basis of sensitivity and specificity, but also on their amenability to multiplexing and their ability to distinguish infectious from non-infectious materials. Results from a critical review of the peer-reviewed literature are presented in Table 5.3.

Electron microscopy (EM). EM, first used to diagnose the early cases of norovirus, remains a useful tool in the clinical setting for a variety of viruses. However, similar to enzyme-linked immunosorbent assays (ELISAs), which are also used primarily in the clinical setting, EM has been largely dismissed for environmental applications and food, samples because of its inherent lack of sensitivity (Kirsten Mattison & Bidawid, 2009).

Polymerase chain reaction (PCR). This molecular method has become the most widely researched way to detect enteric viruses in environmental samples. For the RNA viruses, it is coupled with reverse transcription (RT-PCR). RT-PCR can be performed either by first reverse-transcribing the RNA and then doing traditional PCR to detect the target or as a one-step procedure (Scipioni, Bourgot, et al., 2008; Scipioni, Mauroy, Ziant, Saegerman, & Thiry, 2008). The one-step option has the advantage of reducing material use and risk of introducing foreign RNA during sample handling. In traditional PCR, amplified products are visualized on an agarose gel. Sensitivity can be determined by dot-blotting (Fig. 5.1A).

Table 5.3. Virus detection methods.

Method	Quantitative/ Qualitative	LOD/ LOQ (copies)	Specificity/ Multiplexing	Determine Infectivity?	References
EM	Qualitative	10^5	-	No	(Kou et al., 2006; Kirsten Mattison & Bidawid, 2009)
Real time RT-PCR	Quantitative	1-20	Multiplexing	No	(Park et al., 2008; Scipioni, Mauroy, et al., 2008; Seitz et al., 2011)
RT-PCR-Luminex	Quantitative	10	Multiplexing (AdV, NVGI, NVGII, SV)	No	(J. Liu et al., 2011)
NASBA	Semi- quantitative	1.2×10^{30}	Multiplexing (NVGI, NVGII, HAV)	No	(Gracias & McKillip, 2007; Kou et al., 2006; S. Lamhoujeb et al., 2009)
PCR-ESI-MS	Quantitative	< 100	Multiplexing (AdV)	No	(Blyn et al., 2008)
Western blot	Quantitative	1.5×10^{10} / 3×10^{10}	Unknown	No	(Q. W. Liu et al., 2011)
MS	Qualitative	6×10^{10}	NV	Yes	(D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006; Wigginton, Menin, Montoya, & Kohn, 2010)

AdV, adenovirus; NVGI, norovirus genogroup I; NVGII, norovirus genogroup II; SV, sapovirus; HAV, hepatitis A.

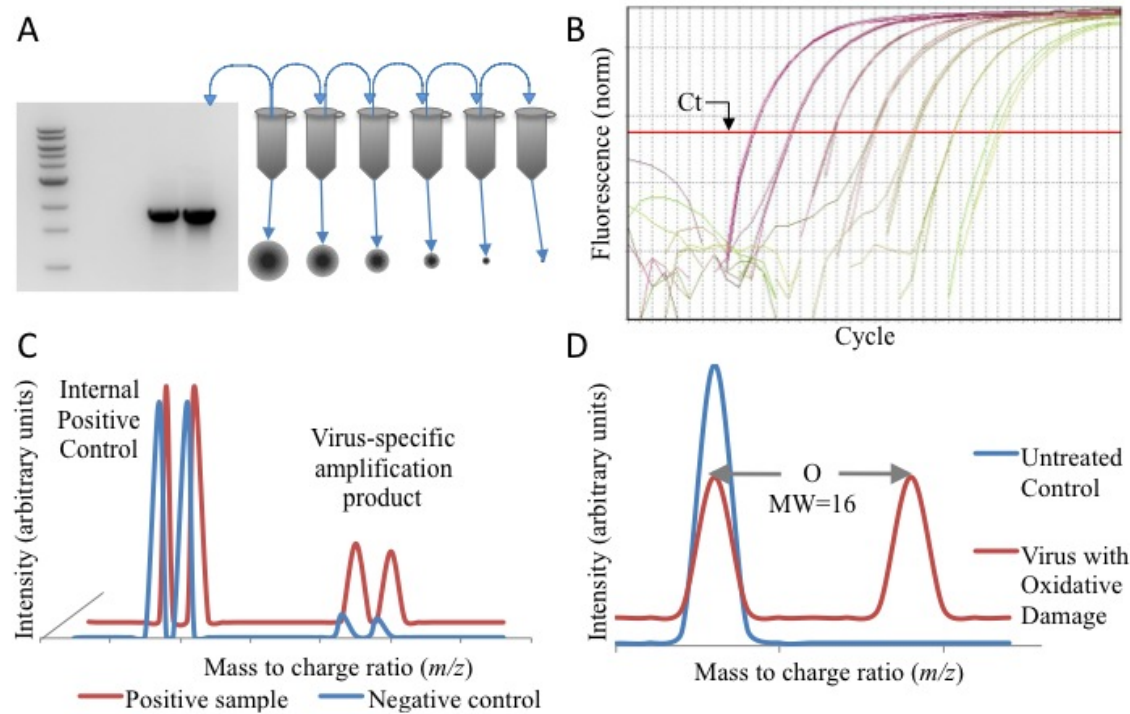


Figure 5.1. Examples of the raw data output from various methods of virus detection. A) Agarose gel electrophoresis and dot-blotting can be used with PCR, RT-PCR, and NASBA . B) Fluorescence curves for a series of standards are generated using real-time or quantitative PCR. C) Internal positive controls and sample amplicons generated using PCR are detected and quantified using mass spectrometry in PCR-ESI-MS (Blyn et al., 2008). D) The effects of oxidation on virus capsid proteins can be observed through the detection of a shift in the mass-to-charge ratio of peaks in mass spectrometry (Wigginton et al., 2010).

Methods can also be quantitative (real-time or qPCR) through the use of SYBR, TaqMan probes, or molecular beacons. Primers are directed against genes coding for capsid proteins or nonstructural proteins or noncoding regions (Rodriguez et al., 2009). Fluorescent probes bind the DNA targets. The intensity of the fluorescence increases with copy number. To quantify, the cycle number, n , at which the curve crosses the threshold, C_t , is plotted against standards of known copy number (Fig. 1B). Sensitivity and specificity depend on the exact protocol, but the most sensitive methods have a limit of detection of one to 10 genome copies (Butot et al., 2010). In general, real time methods are more sensitive than conventional methods. For example, a comparison of real time and conventional RT-PCR for norovirus detection found that the real time method had a limit of detection three to 5 times lower than the conventional method (Park et al., 2008). Regardless of the particular details, primers, probes, or reagents used, these methods all face similar challenges, namely the presence of inhibitors and the inability to distinguish between infectious and noninfectious viruses.

To address the issue of infectivity, primers can also be designed to amplify the entire viral genome (Rodriguez et al., 2009). Targeting the ends or longer fragments of the genome has been found to correlate better with infectivity than direct PCR, although these methods show decreased sensitivity and may not be indicative of loss of infectivity from all modes of inactivation (Rodriguez et al., 2009). One such long-range RT-PCR method for the detection of murine and human norovirus had a limit of detection of 10 murine norovirus plaque forming units/ml (Wolf, Rivera-Aban, & Greening, 2009). In this study, the authors used

murine norovirus to evaluate the performance of both short- and long-range real-time RT-PCR to detect infectivity following inactivation by heat and UV treatment. They observed poor correlation between both of the PCR-based methods and the infectivity assay following heat treatment; the long-range assay, but not the short-range assay, showed a similar trend to that of the infectivity assay following UV irradiation (Wolf et al., 2009).

PCR methods can also be multiplexed to detect several viruses or genogroups at the same time. For example, a combination of primers and probes was used to simultaneously detect three genogroups of norovirus (Wolf et al., 2007). The multiplexed assay performed similarly to singleplexed assays for norovirus GI and GII and had a limit of detection under 10 copies.

PCR products can also be detected using DNA microarrays. The use of DNA microarrays allows for the simultaneous detection of multiple targets, whether viral or bacterial. Several microarrays have been designed to detect norovirus GI and GII (K. Mattison et al., 2011; Seyrig et al., 2011).

However, hybridization can take up to 16 hours (K. Mattison et al., 2011).

Nucleic acid sequence-based amplification (NASBA). To overcome many of the complications of RT-PCR, a one-tube, isothermal amplification method that combines transcription and reverse transcription was created (Gracias & McKillip, 2007; Rodriguez-Lazaro, Hernandez, D'Agostino, & Cook, 2006). As this technique is exclusively for RNA amplification, it is not applicable for adenoviruses, which are DNA viruses. NASBA is reportedly more sensitive than RT-PCR while requiring less preparation because RNA fragments can be

amplified in the presence of double-stranded DNA (Rodriguez-Lazaro et al., 2006). Amplified products can then be detected much as RT-PCR products, *e.g.* on a gel or using a fluorescence detector for real time results (Fig. 1A-B). As with RT-PCR, NASBA reactions can be multiplexed (Rodriguez-Lazaro et al., 2006) and semi-quantitative (S. Lamhoujeb et al., 2009). NASBA-based techniques have been used to detect norovirus GI and GII, and hepatitis A (Kou et al., 2006; Safaa Lamhoujeb et al., 2008; S. Lamhoujeb et al., 2009; Rodriguez-Lazaro et al., 2006) in lettuce (Safaa Lamhoujeb et al., 2008), deli turkey (Safaa Lamhoujeb et al., 2008), and oysters (Kou et al., 2006).

NASBA has a theoretical limit of detection of one copy per reaction, and in practice limits of detection under 10 copies have been observed (Rodriguez-Lazaro et al., 2006). Actual reported limits of detection for virus methods are somewhat lacking. One method for norovirus detection identifies the total amount of RNA needed to give a positive signal from fecal samples and shellfish, respectively, at 5 and 100 pg (Kou et al., 2006; Rodriguez-Lazaro et al., 2006). However it is impossible to deduce the corresponding number of virions from the total RNA extracted. Another norovirus detection method, which coupled NASBA to real-time measurement, had a limit of quantitation of one NASBA particle unit (Safaa Lamhoujeb et al., 2008). While a direct unit conversion is not given, the authors report that the limit of detection of the NASBA-based method was 100-fold lower than that of the equivalent RT-PCR method.

Although less commonly used to detect viruses than RT-PCR, NASBA is a mature technology that has certain advantages with regard to analysis time and

robustness. Commercial kits are available, as are services and guides to design primers for novel applications (Gracias & McKillip, 2007).

Polymerase chain reaction-electrospray ionization-mass spectrometry (PCR-ESI-MS). This hybrid technology combines PCR with mass spectrometry to rapidly and quantitatively detect a range of genetic targets. It is capable of simultaneously detecting and identifying multiple serotypes of the same virus (Blyn et al., 2008), multiple serotypes of different viruses (Sampath et al., 2007), and an internal positive control (see Interpretation) (Blyn et al., 2008; Ecker et al., 2008; Sampath et al., 2007). Redundant primer pairs target the RNA-dependent RNA polymerase (Ecker et al., 2008) as well as other areas of the viral genome (Sampath et al., 2007). The sequence composition of the amplicons is then determined using mass spectrometry to triangulate the identity of the contaminating microorganisms. The internal positive control is also used to determine the absolute abundance of the target sequence, a method which is accurate within a factor of 4 (Ecker et al., 2008). This quantification is accomplished by measuring the ratio of the peak height of the internal positive control vs. that of the sample (Fig. 5.1C).

Quantification is based on a single internal calibrant, a common practice in mass spectrometry, so the analyte and internal calibrant must be present at relatively close abundances to allow detection of both peaks (Blyn et al., 2008). The amount of internal calibrant may need to be optimized for each sample. While mass spectrometry is extremely rapid, the entire method is still limited by the thermocycling step. Even so, an identification can be obtained in 4 to 6 hours

(Ecker et al., 2008; Sampath et al., 2007). Because the underlying principles are based on PCR, this technology is still subject to the traditional problems of inhibition (Ecker et al., 2008). This innovative technology combines the best of PCR and mass spectrometry, but it also inherits the weaknesses of both technologies, namely extended analytical time, sensitivity to inhibitors, and limited dynamic range.

Capsid protein detection. Viruses are enclosed within a capsid, which serves both to protect the genetic material and initiate infection. For the enteric viruses discussed here, the capsid is composed of one to 13 structural proteins (Cliver, 2009; Catherine Fenselau, Laine, & Swatkoski, 2011) repeated 60 (Cliver, 2009) to 180 (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006) times per virion. The capsid is altered during inactivation, the effects of which are evident through changes in antigenicity (Cliver, 2009; Schwab et al., 1996). The effects of UV and singlet oxygen inactivation on capsid proteins are also measurable using mass spectrometry (Wigginton et al., 2010).

Antibody-based detection methods include ELISA and Western blots; ELISAs are used in the clinical setting, but they are largely thought to be too insensitive for environmental or food samples. One Western blot analysis of coxsackievirus A using polyclonal antibodies achieved a limit of detection of 1 ng, or approximately 25 femtomoles, of VP0, a protein precursor that forms a component of the coxsackievirus A capsid (Q. W. Liu et al., 2011). The same method has a linear range suitable for quantitation of between 2 and 6 ng of VP0. Assuming one VP0 per capsid protein and 60 capsid protein copies per virion

(Cliver, 2009), the limit of detection corresponds to 2.5×10^{11} virions and a linear range of 5 to 15×10^{11} virions. However, the specificity of these antibodies has not been evaluated. Also the antisera reacted positively with heat-killed virus, thereby rendering inconclusive when evaluating infectivity.

Mass spectrometry can be used to identify compounds based on a characteristic mass-to-charge ratio (m/z). Analysis can be performed on whole cells, intact proteins, or peptides. In addition to identification, shifts in the mass-to-charge ratio can be used to deduce modifications, such as the incorporation of an oxygen molecule during oxidative inactivation of viruses (Wigginton et al., 2010). The effects of this treatment translate into a shift equivalent to 16 Da (the mass of a single oxygen atom) in the observed peak (Fig. 5.1D).

One mass spectrometry-based method for norovirus detection was shown to be applicable for stool samples (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006). The authors found a limit of detection of 100 femtomoles, which corresponds to approximately 3×10^8 virions. As for ELISAs and Western blots, this limit of detection is acceptable for clinical samples, but it is too high for food samples. Continued instrument and method development will likely result in a decrease in the limit of detection for mass spectrometry-based methods.

Another mass spectrometry-based method for the detection of human adenovirus type 5 has been advanced (Catherine Fenselau et al., 2011). This method uses microwave assisted acid cleavage to digest the capsid and other viral proteins from intact virions. While this method has not been tested on food

samples, it involves minimal processing for virus detection in cell culture. Indeed, the entire procedure takes less than 5 minutes. Processing of food samples will doubtlessly increase the time required. However, this method will still take less time than any PCR-based method, since most thermocycler programs require several hours, and even the isothermal amplification in NASBA takes 1.5 hours (Gracias & McKillip, 2007; Kou et al., 2006).

Although still a nascent methodology, mass spectrometric detection of proteins has several advantages. Because inactivation effects on capsid proteins are measurable (Wigginton et al., 2010), these methods may be able to differentiate between infectious and noninfectious viruses, unlike nucleic acid-based detection. Furthermore, mass spectrometry is not subject to inhibition in the same way as PCR. Mass spectrometry-based methods also have the potential to be extremely rapid (Catherine Fenselau et al., 2011; Swatkoski, Russell, Edwards, & Fenselau, 2007). At present, these mass spectrometric methods for virus detection are qualitative (D. R. Colquhoun, K. J. Schwab, R. N. Cole, and R. U. Halden, 2006) or semi-quantitative (Catherine Fenselau et al., 2011) only.

Interpretation

In addition to the standard positive and negative controls that are crucial to the interpretation of any scientific experiment, the results of virus detection are greatly enhanced by the incorporation of sample process controls and, where appropriate, controls for inhibition. We present an explanation of why these controls are so important and review possible options for these controls.

In lieu of infectivity assays, some inferences on infectivity can be drawn based on the inclusion of certain steps in the procedure. We review which steps are informative in this regard.

Sample process controls. Regardless of the concentration and detection methods used, process controls are a requirement. For quantitative measurements, process controls can be used to judge the overall recovery efficiency and thereby calculate the virus titer in the original sample material; for qualitative measurements, they rule out false negatives due to virus degradation during sample processing. Recovery efficiencies vary based on several factors, including the foodstuff being analyzed and the sample processing methods, so it is not practical to calculate theoretical recovery efficiencies. Because sample process controls must be different from the analyte in question, their recovery efficiency will by definition not be exactly the same as that of the analyte. However, they should be similar enough in structure and resistance to environmental factors to still be informative. Several candidates have been investigated to this end; we limit our discussion to feline calicivirus, murine norovirus and MS2 bacteriophage.

All three of these candidates pose no health risk to humans and have established plaque assays to measure infectivity. Feline calicivirus, although once considered to be the best surrogate for norovirus, has been largely discounted due to its sensitivity to pH (Blaise-Boisseau et al., 2010; Diez-Valcarce et al., 2011), temperature (Bae & Schwab, 2008; Safaa Lamhoujeb et al., 2008), and other inactivation methods (Nuanualsuwan & Cliver, 2002). Feline calicivirus is

demonstrably easier to inactivate than the enteric viruses of interest, and is therefore inappropriate for the purpose of a sample process control because it would lead to significant underestimation of the recovery efficiency.

In contrast, murine norovirus stability is comparable to norovirus and poliovirus (Bae & Schwab, 2008), although it is still less thermotolerant than hepatitis A (Gibson & Schwab, 2011b). Murine norovirus has been used effectively as a process control for vegetables, soft fruits, and shellfish (Diez-Valcarce et al., 2011). Because it is so closely related to norovirus, it is thought to behave similarly during sample processing. However, this similarity could also be an inconvenience for multiplexed detection if the method used is not specific enough to distinguish between the murine and human strains or if the murine norovirus interferes with the assay. For example, some norovirus primers cross-react to produce nonspecific PCR products (Bae & Schwab, 2008). Murine norovirus seems to be an appropriate process control, provided that any interference with detection can be avoided.

The bacteriophage MS2 has also been used extensively as a surrogate for norovirus and shows favorable characteristics for use as a process control (Bae & Schwab, 2008; Blaise-Boisseau et al., 2010; J. Liu et al., 2011). In addition to gene-based detection methods, MS2 would also be compatible with mass spectrometry-based protein detection methods (Swatkoski et al., 2007; Wigginton et al., 2010). However, the use of this or any other process control with immunoconcentration and antibody-based detection methods may not be practical.

Controls for inhibition. As the vast majority of detection methods are in some way PCR-based, a brief discussion of controls for PCR inhibition is warranted. These controls are alternatively referred to as "amplification controls" or "internal positive controls." A cautionary note: "internal" may refer to the PCR (Blyn et al., 2008; Ecker et al., 2008; Sampath et al., 2007) or the sample itself (Bosch et al., 2011). These meanings are not the same; "external control RNA" originating from outside the target organism or sample matrix may be added to the PCR and therefore serve as an "internal positive control" to the reaction.

Such controls are not necessary for detection methods that are not similarly sensitive to interference by humic acids and other inhibitory compounds that might be found in food matrices. Although not subject to the same inhibitors, NASBA-based detection methods could also benefit from the inclusion of inhibition controls (Gracias & McKillip, 2007; S. Lamhoujeb et al., 2009).

Inhibition controls include artificial constructs containing unrelated sequences (Scipioni, Bourgot, et al., 2008; Scipioni, Mauroy, et al., 2008), related sequences of altered amplicon length (Blyn et al., 2008; Ecker et al., 2008; Sampath et al., 2007; Schwab, Estes, Neill, & Atmar, 1997), or RNA targets protected by the MS2 bacteriophage coat (Armored RNA) (Gibson & Schwab, 2011a; Greening & Hewitt, 2008). Regardless of the particular RNA used, the purpose to the inhibition control is to show that the PCR has proceeded as expected. The use of both a process control and an inhibition control may seem gratuitous, but the additional information may be useful in troubleshooting should the process control fail to amplify. The choice of inhibition control is less critical

than the choice of process control, provided that the inclusion of the inhibition control does not interfere with the detection method.

Infectivity. The most direct way to assess infectivity is through the combination of molecular methods and cell culture (Yeh, Yates, Chen, & Mulchandani, 2009). However, cell culture systems do not exist for all CCL-3 viruses, and judged on current methods for other targets, tend to be cumbersome. Pending the discovery of robust, rapid cell culture-based methods, it is more practical to use indirect, molecular methods. Presence of long transcripts and terminal noncoding regions, ability to withstand pre-treatment digestion, antigenicity, and absence of damage to capsid are all lines of evidence that point toward a functional, infectious virion. However, none of these molecular methods has been demonstrated to be universally applicable to all viruses under all conditions. Further research is needed to refine these methods for use with unknown samples. Future directions could perhaps include combinations of these lines of evidence and quantitative results that can be used to perform risk assessment.

Conclusions

The detection of emerging contaminant viruses in food presents many difficult challenges. We have presented some of the most prominent or recent solutions to these challenges and highlighted their strengths and weaknesses. As further improvements and innovations are developed, it is important to remember that, while sensitivity is crucial, so is the ability to distinguish between infectious

and noninfectious material. The latter quality may seem to compromise the former, but in fact it enhances our ability to correctly assess contaminated foods that pose a threat to public health without needlessly wasting products that pose no risk.

Furthermore, some discretion is needed with regard to specificity. Many of the viruses discussed here have multiple genogroups or serotypes. While identifying the offending microorganism down to the strain level is essential for source tracking in an epidemic, it is less important for quality control purposes, where the operator is merely trying to ascertain whether the product is safe for consumption. It is certainly useful to have a very specific assay, but it is even more useful to have an assay capable of detecting multiple possible contaminants. Such assays can be based on highly conserved genes or proteins, or they can consist of multiplex reactions specific to a variety of pathogens. Assays based on conserved elements should, however, be able to differentiate virus strains that are not pathogenic to humans (Wolf et al., 2007). Due to the wide variety of pathogens of concern, it may be most expeditious to have an assay that both detects conserved markers for closely related viruses and multiplexes disparate markers for more distantly related viruses.

In summary, we do not currently have the capability to rapidly detect CCL3 viruses from the full range of pertinent foods. We may be able to make use of current technologies to attain this goal, but further work is needed to properly interpret the results for an informed risk assessment. PCR-based methods are the most mature and make use of more common equipment, so they are the most

feasible to implement. However, the inability to correlate a positive test result with a public health risk severely limits the informational value of the current DNA-based methods for the purpose of quality control and risk assessment (Baert et al., 2011). Future technologies should focus on quantitative measurement of infectivity. Standardized testing is needed to compare results across methods, targets, and samples.

The authors would like to thank Dr. Angela Jansen for her input on the manuscript.

6. Summary and Recommendations

As can be seen in the previous chapter, method development for the detection and quantification of biological agents in environmental compartments is an ongoing process with many open avenues of research. With the help of my collaborators, I have contributed to the field by identifying putative protein biomarkers for toluene degradation by *Pseudomonas putida* F1, *P. mendocina* KR1, and *Burkholderia*. sp. JS150; developing quantitative gene- and protein-based assays for dioxin transformation by *Sphingomonas wittichii* RW1; and developing a quantitative protein-based assay for the norovirus capsid protein. However, much work remains to be done, both in the realm of continued method development and in the realm of application in bioremediation and in the protection of public health.

The method advanced in Chapter 2 combines a simple separation technique, gel electrophoresis, with a powerful identification tool, mass spectrometry. In this study, three potential biomarkers for the degradation of toluene are identified. Previous work in this vein identified the dioxin dioxygenase from *S. wittichii* RW1 as a potential biomarker of dioxin degradation. The work described in Chapters 3 and 4 capitalizes on this suggestion to make quantitative gene- and protein-based assays, respectively. Future research should be devoted to doing similar biomarker discovery for other organisms of interest for bioremediation. In many instances, e.g., for *Dehalococcoides*, gene-based detection methods already exist, but methods for the corresponding enzymes have not been created. Such methods could lend

credence to claims that monitored natural attenuation is occurring or that biostimulation or bioaugmentation strategies are successful.

The results presented in Chapter 2 also raise questions regarding the evolutionary pathways that led to these enzymes of biodegradative importance. The identification algorithm used in this study (Mascot) consistently linked the target (i.e., the toluene dioxygenase) with a seemingly unrelated protein (i.e., a chlorobenzene dioxygenase). It was previously thought that no existing organisms would be able to transform many of the environmental pollutants currently targeted in bioremediation because of the anthropogenic origins and lack of naturally occurring homologs of these chemicals. Some biodegradative enzymes are therefore relatively recently evolved, and the circumstances preceding their evolution, as well as the time required for them to arise, are poorly understood. Research should be done to investigate how sequence homology in disparate enzymes can help or hinder efforts to identify biodegradative enzymes in the environment. Also, it may be possible to take advantage of this technique to identify enzymes involved in previously unobserved reactions, e.g., the biotransformation of perfluorinated compounds, the possibility of which has been speculated about but not confirmed.

Chapters 3 and 4 involve the quantification methods for the dioxin dioxygenase gene and protein, respectively. The results from Chapter 3 indicate that *S. wittichii* RW1 can lose its biodegradative genes in complex growth media, such as would be found in the environment. Studies should be conducted to examine if enzyme production, as measured using the method advanced in

Chapter 4, follows a similar trend or if biodegradative activity is independent of gene copy number. Further investigations could delve into the behavior of *S. wittichii* RW1 in different situations using both quantitative PCR and mass spectrometry. Such studies may shed light on processes controlling the expression of the dioxin degradation pathway, which is likely more complex than previously thought, and maintenance of the degradative megaplasmiids, a hallmark of sphingomonads. Similar methods could be developed and applied to investigate analogous enzymes in other sphingomonads to see if comparable regulatory phenomena can be observed.

Also in Chapter 4, a mass spectrometry-based method to detect the norovirus capsid protein is described. While comparisons have been made between the sensitivity of methods targeting the capsid protein and viral RNA, it could be useful to do a direct comparison of capsid and RNA detection on samples where deductions can be made about infectivity. It might be that the techniques complement each other and that the combination is a better indicator of infectivity than either method alone. Furthermore, with this and other quantitative methods, it should be possible to perform risk assessment on known or potential norovirus-contaminated samples. Norovirus detection methods are often tested on artificially contaminated or clinical samples. If possible, it could be more useful to test this and other methods against samples that have been implicated in a norovirus outbreak. In short, there is much room for improvement in quality control methods for norovirus, as well as for other food- and waterborne pathogens.

Although none of the techniques used in this dissertation are new inventions, their levels of maturity and acceptance vary. Quantitative PCR is by far the most widely used technique to detect microorganisms in the fields of bioremediation and environmental monitoring of pathogens. While mass spectrometry-based techniques have been explored, their use is not commonplace. Before these methods can become widespread, the research community will have to come to consensus on instrumentation and standard operating protocols. Until there are accepted ways to compare results between facilities, these techniques will not and cannot be industry standards.

One lesson to learn from the work in this dissertation is that there are many ways to measure the same information. Detection methods are often compared against each other, but it may be beneficial to look into whether different methods are complementary, contributing different lines of evidence towards the same phenomenon or providing unique information that cannot be deduced from using one method alone. Future work should take advantage of the potential synergy of orthogonal methods.

WORKS CITED

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 25(17), 3389-3402.
- Anderson, N. L., Razavi, M., Pearson, T. W., Kruppa, G., Paape, R., & Suckau, D. (2012). Precision of Heavy-Light Peptide Ratios Measured by MALDI-TOF Mass Spectrometry. *Journal of Proteome Research*, 11(3), 1868-1878. doi: 10.1021/pr201092v
- Apweiler, R., Aslanidis, C., Deufel, T., Gerstner, A., Hansen, J., Hochstrasser, D., . . . Schmitz, G. (2009). Approaching clinical proteomics: current state and future fields of application in fluid proteomics. [Review]. *Clinical Chemistry and Laboratory Medicine*, 47(6), 724-744. doi: 10.1515/cclm.2009.167
- Armengaud, J., Happe, B., & Timmis, K. N. (1998a). Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. Strain RW1: Catabolic genes dispersed on the genome. *Journal of Bacteriology*, 180(15), 3954-3966.
- Armengaud, J., Happe, B., & Timmis, K. N. (1998b). Genetic analysis of dioxin dioxygenase of *Sphingomonas* sp. Strain RW1: catabolic genes dispersed on the genome. *J Bacteriol*, 180(15), 3954-3966.
- Armengaud, J., Happe, B., & Timmis, K. N. (1998c). Genetic Analysis of Dioxin Dioxygenase of *Sphingomonas wittichii* Strain RW1: Catabolic Genes Dispersed on the Genome. *Journal of Bacteriology*, 180(15), 3954-3966.
- Armengaud, J., & Timmis, K. N. (1997). Molecular characterization of Fdx1, a putidaredoxin-type [2Fe-2S] ferredoxin able to transfer electrons to the dioxin dioxygenase of *Sphingomonas* sp. RW1. [Article]. *European Journal of Biochemistry*, 247(3), 833-842.
- Bae, J., & Schwab, K. J. (2008). Evaluation of murine norovirus, feline calicivirus, poliovirus, and MS2 as surrogates for human norovirus in a model of viral persistence in surface water and groundwater. *Applied and Environmental Microbiology*, 74(2), 477-484. doi: 10.1128/aem.02095-06
- Baert, L., Mattison, K., Loisy-Hamon, F., Harlow, J., Martyres, A., Lebeau, B., . . . Uyttendaele, M. (2011). Review: Norovirus prevalence in Belgian, Canadian and French fresh produce: A threat to human health? *International Journal of Food Microbiology*, 151(3), 261-269. doi: 10.1016/j.ijfoodmicro.2011.09.013

- Basta, T., Keck, A., Klein, J., & Stolz, A. (2004a). Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading Sphingomonas strains. *Journal of Bacteriology*, *186*(12), 3862-3872. doi: 10.1128/jb.186.12.3862-3872.2004
- Basta, T., Keck, A., Klein, J., & Stolz, A. (2004b). Detection and characterization of conjugative degradative plasmids in xenobiotic-degrading Sphingomonas strains. *J Bacteriol*, *186*(12), 3862-3872.
- Beller, H. R., Kane, S. R., Legler, T. C., & Alvarez, P. J. J. (2002). A real-time polymerase chain reaction method for monitoring anaerobic, hydrocarbon-degrading bacteria based on a catabolic gene. *Environmental Science & Technology*, *36*(18), 3977-3984. doi: 10.1021/es025556w
- Bertazzi, P. A., Consonni, D., Bachetti, S., Rubagotti, M., Baccarelli, A., Zocchetti, C., & Pesatori, A. C. (2001). Health effects of dioxin exposure: a 20-year mortality study. *Am J Epidemiol*, *153*(11), 1031-1044.
- . Bioterrorism Agents/Diseases. Retrieved January 23, 2012, from <http://www.bt.cdc.gov/agent/agentlist.asp>
- Blaise-Boisseau, S., Hennechart-Collette, C., Guillier, L., & Perelle, S. (2010). Duplex real-time qRT-PCR for the detection of hepatitis A virus in water and raspberries using the MS2 bacteriophage as a process control. *Journal of Virological Methods*, *166*(1-2), 48-53. doi: 10.1016/j.jviromet.2010.02.017
- Blyn, L. B., Hall, T. A., Libby, B., Ranken, R., Sampath, R., Rudnick, K., . . . Ecker, D. J. (2008). Rapid detection and molecular serotyping of adenovirus by use of PCR followed by electrospray ionization mass spectrometry. [Article]. *Journal of Clinical Microbiology*, *46*(2), 644-651. doi: 10.1128/jcm.00801-07
- Borges, C. R., Oran, P. E., Buddi, S., Jarvis, J. W., Schaab, M. R., Rehder, D. S., . . . Nelson, R. W. (2011). Building Multidimensional Biomarker Views of Type 2 Diabetes on the Basis of Protein Microheterogeneity. *Clinical Chemistry*, *57*(5), 719-728.
- Bosch, A., Sanchez, G., Abbaszadegan, M., Carducci, A., Guix, S., Le Guyader, F. S., . . . Sellwood, J. (2011). Analytical Methods for Virus Detection in Water and Food. *Food Analytical Methods*, *4*(1), 4-12. doi: 10.1007/s12161-010-9161-5
- Bunz, P. V., & Cook, A. M. (1993a). Dibenzofuran 4,4a-Dioxygenase from Sphingomonas sp. Strain RW1 - Angular Dioxygenation by a 3-Component Enzyme-System. [Article]. *Journal of Bacteriology*, *175*(20), 6467-6475.

- Bunz, P. V., & Cook, A. M. (1993b). Dibenzofuran 4,4a-dioxygenase from *Sphingomonas* sp. strain RW1: angular dioxygenation by a three-component enzyme system. *J Bacteriol*, *175*(20), 6467-6475.
- Bunz, P. V., & Schmidt, S. (1997). The microbial degradation of halogenated diaryl ethers. *Biotechnol Adv*, *15*(3-4), 621-632.
- Butot, S., Le Guyader, F. S., Krol, J., Putallaz, T., Amoroso, R., & Sanchez, G. (2010). Evaluation of various real-time RT-PCR assays for the detection and quantitation of human norovirus. *Journal of Virological Methods*, *167*(1), 90-94. doi: 10.1016/j.jviromet.2010.03.018
- Campanella, B. F., Bock, C., & Schroder, P. (2002). Phytoremediation to increase the degradation of PCBs and PCDD/Fs. Potential and limitations. *Environ Sci Pollut Res Int*, *9*(1), 73-85.
- Cao, B., & Loh, K.-C. (2008). Catabolic pathways and cellular responses of *Pseudomonas putida* P8 during growth on benzoate with a proteomics approach. *Biotechnology and Bioengineering*, *101*(6), 1297-1312.
- Casas, N., & Sunen, E. (2002). Detection of enteroviruses, hepatitis A virus and rotaviruses in sewage by means of an immunomagnetic capture reverse transcription-PCR assay. *Microbiological Research*, *157*(3), 169-175. doi: 10.1078/0944-5013-00152
- Chao, T.-C., Hansmeier, N., & Halden, R. U. (2010). Towards proteome standards: The use of absolute quantitation in high-throughput biomarker discovery. *Journal of Proteomics*, *73*(3), 1641-1646.
- Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K. L., Tom, L. M., . . . Hettich, R. L. (2010). Direct Cellular Lysis/Protein Extraction Protocol for Soil Metaproteomics. [Article]. *Journal of Proteome Research*, *9*(12), 6615-6622. doi: 10.1021/pr100787q
- Cliver, D. O. (2009). Capsid and Infectivity in Virus Detection. *Food and Environmental Virology*, *1*(3-4), 123-128. doi: 10.1007/s12560-009-9020-y
- Colquhoun, D. R. (2007). *Public Health Applications of Quantitative Biomarkers*. (Doctor of Philosophy), The Johns Hopkins University, Baltimore, MD.
- Colquhoun, D. R., K. J. Schwab, R. N. Cole, and R. U. Halden. (2006). Detection of Norovirus Capsid Protein in Authentic Standards and in Stool Extract by Matrix-Assisted Laser Desorption Ionization (MALDI) and Nanospray Mass Spectrometry. *Appl. Environ. Microbiol.*, *72*(5), 2442 - 24451.

- Cravatt, B. F., Simon, G. M., & Yates, J. R. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature*, *450*(7172), 991-1000. doi: 10.1038/nature06525
- Diez-Valcarce, M., Cook, N., Hernandez, M., & Rodriguez-Lazaro, D. (2011). Analytical Application of a Sample Process Control in Detection of Foodborne Viruses. *Food Analytical Methods*, *4*(4), 614-618. doi: 10.1007/s12161-011-9262-9
- Douma, M. D., Kerr, G. M., Brown, R. S., Keller, B. O., & Oleschuk, R. D. (2008). Mass spectrometric detection of proteins in non-aqueous media - The case of prion proteins in biodiesel. *Canadian Journal of Chemistry- Revue Canadienne De Chimie*, *86*(8), 774-781. doi: 10.1139/v08-083
- Duizer, E., Schwab, K. J., Neill, F. H., Atmar, R. L., Koopmans, M. P. G., & Estes, M. K. (2004). Laboratory efforts to cultivate noroviruses. *Journal of General Virology*, *85*, 79-87. doi: 10.1099/vir.0.19478-0
- Ecker, D. J., Sampath, R., Massire, C., Blyn, L. B., Hall, T. A., Eshoo, M. W., & Hofstadler, S. A. (2008). Innovation - Ibis T5000: a universal biosensor approach for microbiology. *Nature Reviews Microbiology*, *6*(7), 553-558. doi: 10.1038/nrmicro1918
- European Centre for Disease Prevention and Control. (2011). *Annual Epidemiological Report 2011. Reporting on 2009 surveillance data and 2010 epidemic intelligence data*. Stockholm: ECDC.
- European Food Safety Authority. (2012). The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks in 2010. *EFSA Journal*, *10*(3), 2597-3039. doi: 10.2903/j.efsa.2012.2597
- Fenselau, C. (2007). *A review of quantitative methods for proteomic studies*.
- Fenselau, C., & Demirev, F. A. (2001). Characterization of intact microorganisms by MALDI mass spectrometry. *Mass Spectrometry Reviews*, *20*(4), 157-171.
- Fenselau, C., Laine, O., & Swatkoski, S. (2011). Microwave assisted acid cleavage for denaturation and proteolysis of intact human adenovirus. *International Journal of Mass Spectrometry*, *301*(1-3), 7-11. doi: 10.1016/j.ijms.2010.05.026
- Futagami, T., Goto, M., & Furukawa, K. (2008). Biochemical and Genetic Bases of Dehalorespiration. *The Chemical Record*, *8*, 1-12.

- Gibson, K. E., & Schwab, K. J. (2011a). Tangential-Flow Ultrafiltration with Integrated Inhibition Detection for Recovery of Surrogates and Human Pathogens from Large-Volume Source Water and Finished Drinking Water. *Applied and Environmental Microbiology*, *77*(1), 385-391. doi: 10.1128/aem.01164-10
- Gibson, K. E., & Schwab, K. J. (2011b). Thermal Inactivation of Human Norovirus Surrogates. *Food and Environmental Virology*, *3*(2), 74-77. doi: 10.1007/s12560-011-9059-4
- Giometti, C. S. (2006). Tale of two metal reducers: comparative proteome analysis of *Geobacter sulfurreducens* PCA and *Shewanella oneidensis* MR-1. *Methods Biochem Anal*, *49*, 97-111.
- Goncalves, E. R., Hara, H., Miyazawa, D., Davies, J. E., Eltis, L. D., & Mohn, W. W. (2006). Transcriptomic assessment of isozymes in the biphenyl pathway of *Rhodococcus* sp. strain RHA1. *Appl Environ Microbiol*, *72*(9), 6183-6193.
- Gracias, K. S., & McKillip, J. L. (2007). Nucleic acid sequence-based amplification (NASBA) in molecular bacteriology: A procedural guide. *Journal of Rapid Methods and Automation in Microbiology*, *15*(3), 295-309. doi: 10.1111/j.1745-4581.2007.00099.x
- Greening, G. E., & Hewitt, J. (2008). Norovirus Detection in Shellfish Using a Rapid, Sensitive Virus Recovery and Real-Time RT-PCR Detection Protocol. *Food Analytical Methods*, *1*(2), 109-118. doi: 10.1007/s12161-008-9018-3
- Griffiths, W. J., Jonsson, A. P., Liu, S., Rai, D. K., & Wang, Y. (2001). Electrospray and tandem mass spectrometry in biochemistry. *Biochem. J.*, *355*(3), 545-561.
- Gupta, N., Tanner, S., Jaitly, N., Adkins, J. N., Lipton, M., Edwards, R., . . . Pevzner, P. A. (2007). Whole proteome analysis of post-translational modifications: applications of mass-spectrometry for proteogenomic annotation. *Genome Res*, *17*(9), 1362-1377.
- Halden, R. U., Colquhoun, D. R., & Wisniewski, E. S. (2005a). Identification and phenotypic characterization of *Sphingomonas wittichii* strain RW1 by peptide mass fingerprinting using matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Appl Environ Microbiol*, *71*(5), 2442-2451.
- Halden, R. U., Colquhoun, D. R., & Wisniewski, E. S. (2005b). Identification and phenotypic characterization of *Sphingomonas wittichii* strain RW1 by peptide mass fingerprinting using matrix-assisted laser desorption

- ionization-time of flight mass spectrometry. [Article]. *Applied and Environmental Microbiology*, 71(5), 2442-2451. doi: 10.1128/aem.71.5.2442-2451.2005
- Halden, R. U., & Dwyer, D. F. (1997). Biodegradation of Dioxin-Related Compounds: A Review. *Bioremediation Journal*, 1(1), 11-25.
- Halden, R. U., Halden, B. G., & Dwyer, D. F. (1999a). Removal of dibenzofuran, dibenzo-p-dioxin, and 2-chlorodibenzo-p-dioxin from soils inoculated with *Sphingomonas* sp. strain RW1. *Appl Environ Microbiol*, 65(5), 2246-2249.
- Halden, R. U., Halden, B. G., & Dwyer, D. F. (1999b). Removal of dibenzofuran, dibenzo-p-dioxin, and 2-chlorodibenzo-p-dioxin from soils inoculated with *Sphingomonas* sp. strain RW1. *Applied and Environmental Microbiology*, 65(5), 2246-2249.
- Halden, R. U., Tepp, S. M., Halden, B. G., & Dwyer, D. F. (1999). Degradation of 3-phenoxybenzoic acid in soil by *Pseudomonas pseudoalcaligenes* POB310(pPOB) and two modified *Pseudomonas* strains. *Appl Environ Microbiol*, 65(8), 3354-3359.
- Hamza, I. A., Jurzik, L., Überla, K., & Wilhelm, M. (2011). Methods to detect infectious human enteric viruses in environmental water samples. *International Journal of Hygiene and Environmental Health*, 214(6), 424-436.
- Hartmann, E. M., Badalamenti, J. P., Krajmalnik-Brown, R., & Halden, R. U. (2012). Quantitative PCR for Tracking Megaplasmid-borne Biodegradation Potential of a Model *Sphingomonad*. *Applied and Environment Microbiology*, 78(12).
- Hartmann, E. M., Colquhoun, D. R., & Halden, R. U. (2010). Identification of Putative Biomarkers for Toluene-Degrading Burkholderia and Pseudomonads by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry and Peptide Mass Fingerprinting. [Article]. *Bioscience Biotechnology and Biochemistry*, 74(7), 1470-1472. doi: [10.1271/bbb.100064](https://doi.org/10.1271/bbb.100064)
- Hartmann, E. M., & Halden, R. U. (2012). Analytical methods for the detection of viruses in food by example of CCL-3 bioagents. *Analytical and Bioanalytical Chemistry*(Emerging Contaminants in Biota).
- Hays, S. M., & Aylward, L. L. (2003). Dioxin risks in perspective: past, present, and future. *Regulatory Toxicology and Pharmacology*, 37(2), 202-217.

- Holland, R. D., Rafii, F., Heinze, T. M., Sutherland, J. B., Voorhees, K. J., & Lay, J. O. (2000). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometric detection of bacterial biomarker proteins isolated from contaminated water, lettuce and cotton cloth. [Article]. *Rapid Communications in Mass Spectrometry*, *14*(10), 911-917. doi: 10.1002/(sici)1097-0231(20000530)14:10<911::aid-rcm965>3.3.co;2-3
- Hong, H. B., Chang, Y. S., Nam, I. H., Fortnagel, P., & Schmidt, S. (2002). Biotransformation of 2,7-dichloro- and 1,2,3,4-tetrachlorodibenzo-p-dioxin by *Sphingomonas wittichii* RW1. *Applied and Environmental Microbiology*, *68*(5), 2584-2588. doi: 10.1128/aem.68.5.2584-2588.2002
- Hufnagel, P., & Rabus, R. (2006). Mass spectrometric identification of proteins in complex post-genomic projects. Soluble proteins of the metabolically versatile, denitrifying 'Aromatoleum' sp. strain EbN1. *J Mol Microbiol Biotechnol*, *11*(1-2), 53-81.
- Hwang, Y. C., Leong, O. M., Chen, W., & Yates, M. V. (2007). Comparison of a reporter assay and immunomagnetic separation real-time reverse transcription-PCR for the detection of enteroviruses in seeded environmental water samples. *Applied and Environmental Microbiology*, *73*(7), 2338-2340. doi: 10.1128/aem.01758-06
- Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., & Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. [Article]. *Molecular & Cellular Proteomics*, *4*(9), 1265-1272. doi: 10.1074/mcp.M500061-MCP200
- Johnson, G. R., & Olsen, R. H. (1997). Multiple pathways for toluene degradation in *Burkholderia* sp. strain JS150. *Applied and Environmental Microbiology*, *63*(10), 4047-4052.
- Kim, M., Lim, M. Y., & Ko, G. (2010). Enhancement of Enteric Adenovirus Cultivation by Viral Transactivator Proteins. [Article]. *Applied and Environmental Microbiology*, *76*(8), 2509-2516. doi: 10.1128/aem.02224-09
- Kim, S. I., Song, S. Y., Kim, K. W., Ho, E. M., & Oh, K. H. (2003). Proteomic analysis of the benzoate degradation pathway in *Acinetobacter* sp. KS-1. *Res Microbiol*, *154*(10), 697-703.
- Kimura, N., & Kamagata, Y. (2009). Impact of dibenzofuran/dibenzo-p-dioxin amendment on bacterial community from forest soil and ring-hydroxylating dioxygenase gene populations. *Applied Microbiology and Biotechnology*, *84*(2), 365-373. doi: 10.1007/s00253-009-2046-0

- Konduru, K., & Kaplan, G. G. (2006). Stable growth of wild-type hepatitis A virus in cell culture. *Journal of Virology*, *80*(3), 1352-1360. doi: 10.1128/jvi.80.3.1352-1360.2006
- Koopmans, M., & Duizer, E. (2004). Foodborne viruses: an emerging problem. [Article]. *International Journal of Food Microbiology*, *90*(1), 23-41. doi: 10.1016/s0168-1605(03)00169-7
- Kou, X. X., Wu, Q. P., Zhang, J. M., & Fan, H. Y. (2006). Rapid detection of noroviruses in fecal samples and shellfish by nucleic acid sequence-based amplification. *Journal of Microbiology*, *44*(4), 403-408.
- Krajmalnik-Brown, R., Saunders, F. M., Ritalahti, K. M., & Löffler, F. E. (2007). Environmental distribution of the trichloroethene reductive dehalogenase gene (*tceA*) suggests lateral gene transfer among *Dehalococcoides*. *FEMS Microbiology Ecology*, *59*(1), 206-214.
- Kulkarni, P. S., Crespo, J. G., & Afonso, C. A. (2007). Dioxins sources and current remediation technologies - A review. *Environ Int.*
- Kuzyk, M. A., Ohlund, L. B., Elliott, M. H., Smith, D., Qian, H., Delaney, A., . . . Borchers, C. H. (2009). A comparison of MS/MS-based stable-isotope-labeled, quantitation performance on ESI-quadrupole TOF and MALDI-TOF/TOF mass spectrometers. *Proteomics*, *9*, 3328-3340.
- Lamhoujeb, S., Fliss, I., Ngazoa, S. E., & Jean, J. (2008). Evaluation of the persistence of infectious human noroviruses on food surfaces by using real-time nucleic acid sequence-based amplification. *Applied and Environmental Microbiology*, *74*(11), 3349-3355. doi: 10.1128/aem.02878-07
- Lamhoujeb, S., Fliss, I., Ngazoa, S. E., & Jean, J. (2009). Molecular Study of the Persistence of Infectious Human Norovirus on Food-Contact Surfaces. *Food and Environmental Virology*, *1*(2), 51-56. doi: 10.1007/s12560-009-9010-0
- Lay Jr, J. O., & Liyanage, R. (2005). MALDI-TOF mass spectrometry of intact bacteria. *Chemical Analysis*, *169*, 125-152.
- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., . . . Yates, J. R. (1999). Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology*, *17*(7), 676-682.
- Lippolis, J. D., Bayles, D. O., & Reinhardt, T. A. (2009). Proteomic Changes in *Escherichia coli* When Grown in Fresh Milk versus Laboratory Media. *Journal of Proteome Research*, *8*(1), 149-158. doi: doi:10.1021/pr800458v

- Liu, J., Kibiki, G., Maro, V., Maro, A., Kumburu, H., Swai, N., . . . Houpt, E. (2011). Multiplex reverse transcription PCR Luminex assay for detection and quantitation of viral agents of gastroenteritis. *Journal of Clinical Virology*, 50(4), 308-313. doi: 10.1016/j.jcv.2010.12.009
- Liu, Q. W., Ku, Z. Q., Cai, Y. C., Sun, B., Leng, Q. B., & Huang, Z. (2011). Detection, characterization and quantitation of Coxsackievirus A16 using polyclonal antibodies against recombinant capsid subunit proteins. [Article]. *Journal of Virological Methods*, 173(1), 115-120. doi: 10.1016/j.jviromet.2011.01.016
- Ludwig, W., & Schleifer, K. H. (2000). How quantitative is quantitative PCR with respect to cell counts? *Systematic and Applied Microbiology*, 23(4), 556-562.
- Madsen, E. L. (1991). DETERMINING INSITU BIODEGRADATION - FACTS AND CHALLENGES. [Review]. *Environmental Science & Technology*, 25(10), 1663-1673.
- Maeda, H., Fujimoto, C., Haruki, Y., Maeda, T., Kokeyuchi, S., Petelin, M., . . . Takashiba, S. (2003). Quantitative real-time PCR using TaqMan and SYBR Green for *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Prevotella intermedia*, *tetQ* gene and total bacteria. *FEMS Immunology and Medical Microbiology*, 39(1), 81-86. doi: 10.1016/s0928-8244(03)00224-4
- Mattison, K., & Bidawid, S. (2009). Analytical Methods for Food and Environmental Viruses. *Food and Environmental Virology*, 1(3-4), 107-122. doi: 10.1007/s12560-009-9017-6
- Mattison, K., Corneau, N., Berg, I., Bosch, A., Duizer, E., Gutierrez-Aguirre, I., . . . Bidawid, S. (2011). Development and validation of a microarray for the confirmation and typing of norovirus RT-PCR products. *Journal of Virological Methods*, 173(2), 233-250. doi: 10.1016/j.jviromet.2011.02.011
- Mazzoli, R., Pessione, E., Giuffrida, M. G., Fattori, P., Barello, C., Giunta, C., & Lindley, N. D. (2007). Degradation of aromatic compounds by *Acinetobacter radioresistens* S13: Growth characteristics on single substrates and mixtures. *Archives of Microbiology*, 188(1), 55-68.
- Miller, T. R., Delcher, A. L., Salzberg, S. L., Saunders, E., Detter, J. C., & Halden, R. U. (2010). The Genome Sequence of the Dioxin Mineralizing Bacterium *Sphingomonas wittichii* RW1. *Journal of Bacteriology*, 192(22), JB.01030-01010. doi: [10.1128/jb.01030-10](https://doi.org/10.1128/jb.01030-10)

- Morel, N., Andreoletti, O., Grassi, J., & Clement, G. (2007). Absolute and relative quantification of sheep brain prion protein (PrP) allelic variants by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21(24), 4093-4100. doi: 10.1002/rcm.3317
- Muller, T. A., Werlen, C., Spain, J., & van der Meer, J. R. (2003). Evolution of a chlorobenzene degradative pathway among bacteria in a contaminated groundwater mediated by a genomic island in *Ralstonia*. [Article]. *Environmental Microbiology*, 5(3), 163-173.
- Nam, I. H., Hong, H. B., Kim, Y. M., Kim, B. H., Murugesan, K., & Chang, Y. S. (2005). Biological removal of polychlorinated dibenzo-p-dioxins from incinerator fly ash by *Sphingomonas wittichii* RW1. *Water Research*, 39(19), 4651-4660. doi: 10.1016/j.watres.2005.09.009
- Nam, I. H., Kim, Y. M., Schmidt, S., & Chang, Y. S. (2006). Biotransformation of 1,2,3-Tri- and 1,2,3,4,7,8-hexachlorodibenzo-p-dioxin by *Sphingomonas wittichii* strain RW1. *Applied and Environmental Microbiology*, 72(1), 112-116. doi: 10.1128/aem.72.1.112-116.2006
- Nappier, S. P., Graczyk, T. K., Tamang, L., & Schwab, K. J. (2010). Co-localized *Crassostrea virginica* and *Crassostrea ariakensis* Oysters differ in bioaccumulation, retention and depuration of microbial indicators and human enteropathogens. *Journal of Applied Microbiology*, 108(2), 736-744. doi: 10.1111/j.1365-2672.2009.04480.x
- Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., Martins dos Santos, V. A., . . . Fraser, C. M. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol*, 4(12), 799-808.
- Nelson, R. W., & Borges, C. R. (2011). Mass Spectrometric Immunoassay Revisited. *Journal of the American Society for Mass Spectrometry*, 22(6), 960-968. doi: 10.1007/s13361-011-0094-z
- Nie, L., Wu, G., Culley, D. E., Scholten, J. C., & Zhang, W. (2007). Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. *Crit Rev Biotechnol*, 27(2), 63-75.
- Nojiri, H., & Omori, T. (2002). Molecular bases of aerobic bacterial degradation of dioxins: Involvement of angular dioxygenation. [Review]. *Bioscience Biotechnology and Biochemistry*, 66(10), 2001-2016.
- . Noroviruses. (2006, August 3, 2006). *CDC Fact Sheet* Retrieved August 8, 2009, 2009, from <http://www.cdc.gov/ncidod/dvrd/revb/gastro/norovirus-factsheet.htm>

- Nuanualsuwan, S., & Cliver, D. O. (2002). Pretreatment to avoid positive RT-PCR results with inactivated viruses. *Journal of Virological Methods*, *104*(2), 217-225. doi: 10.1016/s0166-0934(02)00089-7
- Oleksiewicz, M. B., Kjeldal, H. Ø., & Klenø, T. G. (2005). Identification of stool proteins in C57BL/6J mice by two-dimensional gel electrophoresis and MALDI-TOF mass spectrometry. *Biomarkers*, *10*(1), 29-40.
- Ong, S.-E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., & Mann, M. (2002). Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol Cell Proteomics*, *1*(5), 376-386. doi: 10.1074/mcp.M200025-MCP200
- Onisko, B. C., Silva, C. J., Dynin, I., Erickson, M., Vensel, W. H., Hnasko, R., . . . Carter, J. M. (2007). Sensitive, preclinical detection of prions in brain by nanospray liquid chromatography/tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, *21*(24), 4023-4026. doi: 10.1002/rcm.3310
- Parales, R. E., Parales, J. V., Pelletier, D. A., & Ditty, J. L. (2008). *Advances in Applied Microbiology* (Vol. Volume 64). Maryland Heights, MO: Elsevier.
- Park, Y., Cho, Y. H., Jee, Y., & Ko, G. (2008). Immunomagnetic separation combined with real-time reverse transcriptase PCR assays for detection of norovirus in contaminated food. [Article]. *Applied and Environmental Microbiology*, *74*(13), 4226-4230. doi: 10.1128/aem.00013-08
- Pellerin, C. (2000). The next target of bioterrorism: Your food. *Environmental Health Perspectives*, *108*(3), A126-A129.
- Rang, H. P., & Dale, M. M. (1991). Absorption, Fate and Distribution of Drugs *Pharmacology* (2nd ed., pp. 955). Edinburgh: Churchill Livingstone.
- Renato, Z., & Richard, K. (1998). Ion formation in MALDI mass spectrometry. *Mass Spectrometry Reviews*, *17*(5), 337-366.
- Richardson, S. D. (2010). Environmental Mass Spectrometry: Emerging Contaminants and Current Issues. *Analytical Chemistry*, *82*(12), 4742-4774. doi: 10.1021/ac101102d
- Ritalahti, K. M., Amos, B. K., Sung, Y., Wu, Q. Z., Koenigsberg, S. S., & Loffler, F. E. (2006). Quantitative PCR targeting 16S rRNA and reductive dehalogenase genes simultaneously monitors multiple Dehalococcoides strains. *Applied and Environmental Microbiology*, *72*(4), 2765-2774. doi: 10.1128/aem.72.4.2765-2774.2006

- Rodriguez, R. A., Pepper, I. L., & Gerba, C. P. (2009). Application of PCR-Based Methods To Assess the Infectivity of Enteric Viruses in Environmental Samples. *Applied and Environmental Microbiology*, 75(2), 297-307. doi: 10.1128/aem.01150-08
- Rodriguez-Lazaro, D., Hernandez, M., D'Agostino, M., & Cook, N. (2006). Application of nucleic acid sequence-based amplification for the detection of viable foodborne pathogens: Progress and challenges. *Journal of Rapid Methods and Automation in Microbiology*, 14(3), 218-236. doi: 10.1111/j.1745-4581.2006.00048.x
- Rozen, S., & Skaletsky, H. J. (2002). Primer3 on the WWW for general users and for biologist programmers. In S. Krawetz & S. Misener (Eds.), *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (pp. 365-386). Totowa, NJ: Humana Press.
- Sakaki, T., & Munetsuna, E. (2010). Enzyme systems for biodegradation of polychlorinated dibenzo-p-dioxins. *Applied Microbiology and Biotechnology*, 88(1), 23-30. doi: [10.1007/s00253-010-2765-2](https://doi.org/10.1007/s00253-010-2765-2)
- Sambrook, J., Fritsh, E. F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
- Sampath, R., Hall, T. A., Massire, C., Li, F., Blyn, L. B., Eshoo, M. W., . . . Ecker, D. J. (2007). Rapid identification of emerging infectious agents using PCR and electrospray ionization mass spectrometry. In S. K. Lal (Ed.), *Biology of Emerging Viruses: Sars, Avian and Human Influenza, Metapneumovirus, Nipah, West Nile, and Ross River Virus* (Vol. 1102, pp. 109-120).
- Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., . . . Griffin, P. M. (2011). Foodborne Illness Acquired in the United States-Major Pathogens. *Emerging Infectious Diseases*, 17(1), 7-15. doi: 10.3201/eid1701.P11101
- Schiene-Fischer, C., Habazettl, J., Schmid, F. X., & Fischer, G. (2002). The hsp70 chaperone DnaK is a secondary amide peptide bond cis-trans isomerase. *Nat Struct Biol*, 9(6), 419-424.
- Schwab, K. J., DeLeon, R., & Sobsey, M. D. (1996). Immunoaffinity concentration and purification of waterborne enteric viruses for detection by reverse transcriptase PCR. *Applied and Environmental Microbiology*, 62(6), 2086-2094.
- Schwab, K. J., Estes, M. K., Neill, F. H., & Atmar, R. L. (1997). Use of heat release and an internal RNA standard control in reverse transcription-PCR

detection of Norwalk virus from stool samples. *Journal of Clinical Microbiology*, 35(2), 511-514.

- Schwab, K. J., Neill, F. H., Estes, M. K., Metcalf, T. G., & Atmar, R. L. (1998). Distribution of norwalk virus within shellfish following bioaccumulation and subsequent depuration by detection using RT-PCR. *Journal of Food Protection*, 61(12), 1674-1680.
- Schwab, K. J., Neill, F. H., Fankhauser, R. L., Daniels, N. A., Monroe, S. S., Bergmire-Sweat, D. A., . . . Atmar, R. L. (2000). Development of methods to detect "Norwalk-like viruses" (NLVs) and hepatitis A virus in delicatessen foods: Application to a food-borne NLV outbreak. *Applied and Environmental Microbiology*, 66(1), 213-218.
- Scipioni, A., Bourgot, I., Mauroy, A., Ziant, D., Saegerman, C., Daube, G., & Thiry, E. (2008). Detection and quantification of human and bovine noroviruses by a TaqMan RT-PCR assay with a control for inhibition. *Molecular and Cellular Probes*, 22(4), 215-222. doi: 10.1016/j.mcp.2008.02.003
- Scipioni, A., Mauroy, A., Ziant, D., Saegerman, C., & Thiry, E. (2008). A SYBR Green RT-PCR assay in single tube to detect human and bovine noroviruses and control for inhibition. *Virology Journal*, 5. doi: 94 10.1186/1743-422x-5-94
- Seah, S. Y. K., Ke, J. Y., Denis, G., Horsman, G. P., Fortin, P. D., Whiting, C. J., & Eltis, L. D. (2007). Characterization of a C-C bond hydrolase from *Sphingomonas wittichii* RW1 with novel specificities towards polychlorinated biphenyl metabolites. [Article]. *Journal of Bacteriology*, 189(11), 4038-4045. doi: [10.1128/jb.01950-06](https://doi.org/10.1128/jb.01950-06)
- Seaver, L. C., & Imlay, J. A. (2001). Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol*, 183(24), 7173-7181.
- Seitz, S. R., Leon, J. S., Schwab, K. J., Lyon, G. M., Dowd, M., McDaniels, M., . . . Moe, C. L. (2011). Norovirus Infectivity in Humans and Persistence in Water. *Applied and Environmental Microbiology*, 77(19), 6884-6888. doi: 10.1128/aem.05806-11
- Seung, I. K., Choi, J. S., & Kahng, H. Y. (2007). A proteomics strategy for the analysis of bacterial biodegradation pathways. *OMICS: A Journal of Integrative Biology*, 11(3), 280-294.
- Seyrig, G., Herzog, A. B., Ahmad, F., Srinivasan, S., Kronlein, M. R., Bhaduri, P., & Hashsham, S. A. (2011). Detection and Occurrence of Indicator

Organisms and Pathogens. [Article]. *Water Environment Research*, 83(10), 900-926. doi: 10.2175/106143011x13075599869056

- Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*, 68(5), 850-858.
- Shiaw-Lin, W., Ian, J., William, S. H., & Barry, L. K. (2004). A new and sensitive on-line liquid chromatography/mass spectrometric approach for top-down protein analysis: the comprehensive analysis of human growth hormone in an *E. coli* lysate using a hybrid linear ion trap/Fourier transform ion cyclotron resonance mass spectrometer. *Rapid Communications in Mass Spectrometry*, 18(19), 2201-2207.
- Shiio, Y., & Aebersold, R. (2006). Quantitative proteome analysis using isotope-coded affinity tags and mass spectrometry. *Nat. Protocols*, 1(1), 139-145.
- Shirale, D. J., Bangar, M. A., Park, M., Yates, M. V., Chen, W., Myung, N. V., & Mulchandani, A. (2010). Label-Free Chemiresistive Immunosensors for Viruses. [Article]. *Environmental Science & Technology*, 44(23), 9030-9035. doi: 10.1021/es102129d
- Siegrist, T. J., and T. R. Sandrin. (2007). Discrimination and characterization of environmental strains of *E. coli* by MALDI-TOF MS. *J. Microbiol. Methods*, 68(3), 554 - 562.
- Silva, C. J., Onisko, B. C., Dynin, I., Erickson, M. L., Requena, J. R., & Carter, J. M. (2011). Utility of Mass Spectrometry in the Diagnosis of Prion Diseases. [Article]. *Analytical Chemistry*, 83(5), 1609-1615. doi: 10.1021/ac102527w
- Singh, O. V. (2006). Proteomics and metabolomics: the molecular make-up of toxic aromatic pollutant bioremediation. *Proteomics*, 6(20), 5481-5492.
- Singh, O. V., & Nagaraj, N. S. (2006). Transcriptomics, proteomics and interactomics: unique approaches to track the insights of bioremediation. *Brief Funct Genomic Proteomic*, 4(4), 355-362.
- Sowell, S. M., Wilhelm, L. J., Norbeck, A. D., Lipton, M. S., Nicora, C. D., Barofsky, D. F., . . . Giovanonni, S. J. (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. [Article]. *Isme Journal*, 3(1), 93-105. doi: 10.1038/ismej.2008.83
- Steenland, K., & Deddens, J. (2003). Dioxin: exposure-response analyses and risk assessment. *Ind Health*, 41(3), 175-180.

- Stolz, A. (2009). Molecular characteristics of xenobiotic-degrading sphingomonads. [Review]. *Applied Microbiology and Biotechnology*, 81(5), 793-811. doi: 10.1007/s00253-008-1752-3
- Straub, T. M., Bartholomew, R. A., Valdez, C. O., Valentine, N. B., Dohnalkova, A., Ozanich, R. M., . . . Call, D. R. (2011). Human norovirus infection of Caco-2 cells grown as a three-dimensional tissue structure. [Article]. *Journal of Water and Health*, 9(2), 225-240. doi: 10.2166/wh.2010.106
- Straub, T. M., Bentrup, K. H. Z., Orosz-Coghlan, P., Dohnalkova, A., Mayer, B. K., Bartholomew, R. A., . . . Nickerson, C. A. (2007). In vitro cell culture infectivity assay for human noroviruses. [Article]. *Emerging Infectious Diseases*, 13(3), 396-403.
- Straub, T. M., Bentrup, K. H. z., Orosz-Coghlan, P., Dohnalkova, A., Mayer, B. K., Bartholomew, R. A., . . . Nickerson, C. A. (2007). Cell culture assay for human noroviruses - Response. *Emerging Infectious Diseases*, 13(7), 1117-1118.
- Stutz, H. (2005). Advances in the analysis of proteins and peptides by capillary electrophoresis with matrix-assisted laser desorption/ionization and electrospray-mass spectrometry detection. [Review]. *Electrophoresis*, 26(7-8), 1254-1290. doi: 10.1002/elps.200410130
- . Surveillance for Norovirus Outbreaks. (March 12, 2012) Retrieved March 20, 2012, from <http://www.cdc.gov/Features/dsNorovirus>
- Susin, M. F., Baldini, R. L., Gueiros-Filho, F., & Gomes, S. L. (2006). GroES/GroEL and DnaK/DnaJ have distinct roles in stress responses and during cell cycle progression in *Caulobacter crescentus*. *J Bacteriol*, 188(23), 8044-8053.
- Swatkoski, S., Russell, S., Edwards, N., & Fenselau, C. (2007). Analysis of a model virus using residue-specific chemical cleavage and MALDI-TOF mass spectrometry. [Article]. *Analytical Chemistry*, 79(2), 654-658. doi: 10.1021/ac061493e
- Tabata, M., Endo, R., Ito, M., Ohtsubo, Y., Kumar, A., Tsuda, M., & Nagata, Y. (2011). The lin Genes for gamma-Hexachlorocyclohexane Degradation in *Sphingomonas* sp MM-1 Proved to Be Dispersed across Multiple Plasmids. *Bioscience Biotechnology and Biochemistry*, 75(3), 466-472. doi: 10.1271/bbb.100652
- Tahk, H., Lee, K. B., Lee, M. H., Seo, D. J., Cheon, D. S., & Choi, C. (2012). Development of reverse transcriptase polymerase chain reaction enzyme-linked immunosorbent assay for the detection of hepatitis A virus in

vegetables. *Food Control*, 23(1), 210-214. doi:
10.1016/j.foodcont.2011.07.012

- United States Environmental Protection Agency. (2008). Microorganisms on the CCL 2 *Regulatory Determinations Support Document for CCL 2*.
- Unlu, M., Morgan, M. E., & Minden, J. S. (1997). Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis*, 18(11), 2071-2077.
- van der Zaan, B., Hannes, F., Hoekstra, N., Rijnaarts, H., de Vos, W. M., Smidt, H., & Gerritse, J. (2010). Correlation of *Dehalococcoides* 16S rRNA and Chloroethene-Reductive Dehalogenase Genes with Geochemical Conditions in Chloroethene-Contaminated Groundwater. *Applied and Environment Microbiology*, 76(3), 843-850.
- Von Seggern, C. E., & Halden, R. U. (2009). Detection of bioterrorism agents and related public health threats utilising matrix-assisted laser desorption/ionisation mass spectrometry. *International Journal of Health Science*, 2(2), 197-203.
- Wackett, L. P. a. C. D. H. (2001). *Biocatalysis and Biodegradation: Microbial Transformation of Organic Compounds* (1 ed.). Washington, D.C.: ASM Press.
- Wahl, K. L., Wunschel, S. C., Jarman, K. H., Valentine, N. B., Petersen, C. E., Kingsley, M. T., . . . Saenz, A. J. (2002). Analysis of Microbial Mixtures by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Analytical Chemistry*, 74(24), 6191-6199. doi: 10.1021/ac0203847
- Walters, E., McClellan, K., & Halden, R. U. (2010). Occurrence and loss over three years of 72 pharmaceuticals and personal care products from biosolids-soil mixtures in outdoor mesocosms. [Article]. *Water Research*, 44(20), 6011-6020. doi: 10.1016/j.watres.2010.07.051
- White, D. C., Sutton, S. D., & Ringelberg, D. B. (1996). The genus *Sphingomonas*: physiology and ecology. *Current Opinion in Biotechnology*, 7, 301-206.
- Wigginton, K. R., Menin, L., Montoya, J. P., & Kohn, T. (2010). Oxidation of Virus Proteins during UV(254) and Singlet Oxygen Mediated Inactivation. [Article]. *Environmental Science & Technology*, 44(14), 5437-5443. doi: 10.1021/es100435a

- Wittich, R. M., Wilkes, H., Sinnwell, V., Francke, W., & Fortnagel, P. (1992a). Metabolism of dibenzo-p-dioxin by *Sphingomonas* sp. strain RW1. *Appl Environ Microbiol*, 58(3), 1005-1010.
- Wittich, R. M., Wilkes, H., Sinnwell, V., Francke, W., & Fortnagel, P. (1992b). Metabolism of Dibenzo-*para*-Dioxin by *Sphingomonas* sp. Strain-RW1. *Applied and Environmental Microbiology*, 58(3), 1005-1010.
- Wolf, S., Rivera-Aban, M., & Greening, G. E. (2009). Long-Range Reverse Transcription as a Useful Tool to Assess the Genomic Integrity of Norovirus. *Food and Environmental Virology*, 1(3-4), 129-136. doi: 10.1007/s12560-009-9016-7
- Wolf, S., Williamson, W. M., Hewitt, J., Rivera-Aban, M., Lin, S., Ball, A., . . . Greening, G. E. (2007). Sensitive multiplex real-time reverse transcription-PCR assay for the detection of human and animal noroviruses in clinical and environmental samples. *Applied and Environmental Microbiology*, 73(17), 5464-5470. doi: 10.1128/aem.00572-07
- Wunschel, S. C., Jarman, K. H., Petersen, C. E., Valentine, N. B., Wahl, K. L., Schauki, D., . . . White V, E. (2005). Bacterial analysis by MALDI-TOF mass spectrometry: An inter-laboratory comparison. *Journal of the American Society for Mass Spectrometry*, 16(4), 456-462.
- Yeh, H.-Y., Yates, M. V., Chen, W., & Mulchandani, A. (2009). Real-time molecular methods to detect infectious viruses. *Seminars in Cell & Developmental Biology*, 20(1), 49-54. doi: 10.1016/j.semcd.2009.01.012
- Zabrouskov, V., Senko, M. W., Du, Y., Leduc, R. D., & Kelleher, N. L. (2005). New and automated MSn approaches for top-down identification of modified proteins. [Article]. *Journal of the American Society for Mass Spectrometry*, 16(12), 2027-2038. doi: 10.1016/j.jasms.2005.08.004
- Ziv-El, M., Delgado, A. G., Yao, Y., Kang, D. W., Nelson, K. G., Halden, R. U., & Krajmalnik-Brown, R. (2011). Development and characterization of DehaloR boolean AND 2, a novel anaerobic microbial consortium performing rapid dechlorination of TCE to ethene. [Article]. *Applied Microbiology and Biotechnology*, 92(5), 1063-1071. doi: 10.1007/s00253-011-3388-y
- Zoll, G. J., Melchers, W. J. G., Kopecka, H., Jambroes, G., Vanderpoel, H. J. A., & Galama, J. M. D. (1992). General Primer-mediated Polymerase Chain-reaction for Detection of Enteroviruses - Application for Diagnostic Routine and Persistent Infections. *Journal of Clinical Microbiology*, 30(1), 160-165.

Zylstra, G. J., & Gibson, D. T. (1991). Aromatic hydrocarbon degradation: a molecular approach. *Genetic engineering*, *13*, 183-203.

APPENDIX A.

PROTEOMIC PROFILING OF THE DIOXIN-DEGRADING BACTERIUM
SPHINGOMONAS WITTICHII RW1

Dioxins are some of the most widely studied and prevalent anthropogenic environmental pollutants. Common sources include backyard burning, incineration of plastics, and chlorine bleaching of pulp in paper mills (Steenland & Deddens, 2003). Documented health effects include acute and chronic effects, including chloracne, various types of cancers, reproductive diseases, circulatory and respiratory diseases, as well as diabetes (Bertazzi et al., 2001). The traditional approach to environmental remediation includes a host of physical and chemical methods, depending on the characteristics of the polluted site and the extent of contamination present (Kulkarni, Crespo, & Afonso, 2007). Bioremediation, i.e., the use of biogenic materials and organisms for environmental cleanup, has also been proposed, including phytoremediation using plants (Campanella, Bock, & Schroder, 2002), and microbial degradation using primarily bacteria and fungi (Halden, Halden, & Dwyer, 1999a). Bioremediation is an attractive strategy, as it can destroy the pollutant rather than transferring it from one environmental compartment to another. It also can be less expensive than physical strategies (Wackett, 2001). Common bioremediation strategies include the addition of nutrients, degradative microorganisms or both.

Sphingomonas wittichii RW1 is a microorganism of great interest to the bioremediation community for its ability to biotransform a large number of toxic polychlorinated dioxins and to utilize both non-chlorinated dibenzo-*p*-dioxin and nonchlorinated dibenzofuran as a growth substrate and sole source of carbon and energy (Halden, Halden, et al., 1999a).

One of the major challenges in bioaugmentation strategies relying on the addition of non-native microbes is to ensure their viability and degradative activity toward the target compounds. Monitoring bioremediation is critical to ensure the efficacy of the process and the reduction of contaminant mass to acceptable levels. Traditionally, the most important characteristics investigated for microorganisms used in bioremediation were their ability to transform the substrate, the rate of substrate removal, and the resulting metabolites (Halden & Dwyer, 1997; Halden, Tepp, Halden, & Dwyer, 1999; Wackett, 2001; Wittich, Wilkes, Sinnwell, Francke, & Fortnagel, 1992a). To optimize the bacterial degradation of pollutants, it is important to understand how these organisms function during growth on recalcitrant substrates and which factors influence their degradative abilities. This includes analyzing not only the degradative pathways (Armengaud, Happe, & Timmis, 1998b; Basta, Keck, Klein, & Stolz, 2004b; Bunz & Schmidt, 1997), but also the peripheral processes and mechanisms that are involved in taxis (i.e., directed motion in a chemical gradient), uptake, and transport during exposures to specific substrates. Analysis of DNA and RNA (Goncalves et al., 2006) can shed light on an organism's metabolic potential; however, these measurements poorly correlate to actual protein expression profiles (Nie, Wu, Culley, Scholten, & Zhang, 2007). Therefore, global analyses of protein expression profiles may be a more informative tool for understanding the physiological mechanisms of biodegradation. In addition to identifying important degradative enzymes in a variety of important microbes (S. I. Kim, Song, Kim, Ho, & Oh, 2003; K. E. Nelson et al., 2002; Singh, 2006), proteomic

studies have opened the door to a better understanding of system-wide changes in response to differing substrates (Singh & Nagaraj, 2006).

The imperative to perform proteomic analyses is particularly true for *S. wittichii* RW1 because the enzymes in the dioxin degradation pathway are encoded on different loci throughout the genome (Armengaud, Happe, & Timmis, 1998c), certain elements in the pathway are located on a plasmid (Basta et al., 2004a), and there may be alternative pathways at work (Seah et al., 2007). The present study builds on previous work (Halden, Colquhoun, & Wisniewski, 2005a) and utilized difference gel electrophoresis (DIGE) coupled with mass spectrometry (MS) to exploit recently gathered RW1 genome data (Miller et al., 2010). When used together, these tools yield information on the response of cells of *S. wittichii* RW1 to dioxin exposure, and the bacterium's degradative activity toward this recalcitrant compound.

The aim of this study was to investigate system-wide changes in protein expression during growth on dibenzofuran, a non-toxic surrogate for dibenzo-*p*-dioxin, as compared to non-selective growth media. Acetate was selected as the non-selective alternate substrate, as growth on this compound was observed to influence expression of select proteins, including the dioxin dioxygenase (Bunz & Cook, 1993b). Thus, any changes measured in response to cells grown on dibenzofuran should represent cell-wide effects related specifically to the growth substrate and not to unanticipated extraneous effects. This work constitutes the first global assessment of protein expression by *S. wittichii* RW1.

Materials and Methods

Culture maintenance. Cultures of *S. wittichii* strain RW1 (100 mL to 1.0 L) were grown to mid log phase at 30°C in M9 phosphate buffered minimal medium (pH 7.05) supplemented with either dibenzofuran crystals or 50 mM acetate as growth substrates. Saturated dibenzofuran medium contained approximately 3 – 5 mg/L of the selective growth substrate in the dissolved phase. Cells were grown overnight on dibenzofuran and acetate as sole carbon sources to an optical density of 0.4 – 0.6 absorbance units (λ 560 nm). Following biomass processing, protein levels in the samples were in the order of 75 – 200 μ g/mL. Protein concentrations were normalized prior to analysis by concentration and resuspension in DIGE sample preparation buffer. Culture purity was confirmed by the streak plate method using Luria Bertani medium supplemented with 1.5% agar.

Protein extraction and cleanup. Cultures were centrifuged at 3,000 – 5,000 x g for 10 minutes at 4°C. Harvested biomass was washed, spun again, and the resultant pellet suspended in a small volume of 100 mM ammonium bicarbonate (pH ~7.0). This microbial suspension was then sonicated under cooling with ice, using a microtip sonicator (Fisher Scientific, Pittsburgh, PA) in a sequence of three 10-second bursts delivered in thirty second intervals. The sonicated cells were then immediately centrifuged at 10,000 x g for 10 minutes at 4°C. The supernatant was collected and the protein purified by trichloroacetic acid (TCA)/acetone precipitation. Briefly, 8 parts of 10% TCA in acetone (-20°C) were added per volume of supernatant and, following mixing on a vortex, the

resultant dilution was incubated at -20°C overnight. Following centrifugation (10,000 x g, 10 minutes, 4°C), harvested biomass was washed in cold acetone for 10 minutes at -20°C. Following a subsequent centrifugation, the pellet was resuspended in sample preparation buffer (7 M urea, 2 M thiourea, 2% CHAPS, 0.2% DTT, 0.02% bromophenol blue) and stored at -20°C until analyzed. Protein concentrations were measured using the bicinchoninic acid assay (Pierce, Rockford, IL) following dilution to reduce the concentration of interfering agents.

DIGE Labeling. Twenty-five µg of crude cell lysates of RW1 biological replicates grown on dibenzofuran (n = 3) and acetate (n = 3) were labeled using Cy dyes (GE Healthcare) as described elsewhere (Unlu, Morgan, & Minden, 1997). Briefly, samples were adjusted to 1 µg/µL using sample preparation buffer and the pH checked. Subsequently, 0.25 µL of 1 pmol/µL Cy dyes were added to samples for 30 minutes in the dark on ice. To stop the labeling reaction, 0.5 µL of 10 mM lysine was added to the samples, which were mixed and incubated on ice for 10 minutes prior to storage at -20°C until analysis.

2D-DIGE. Unless stated otherwise, all procedures were carried out in the dark or minimal light to protect the integrity of the fluorescent dyes. Samples were randomized to reduce the effect of dye bias and in-gel variations. A global pool consisting of fractions of each sample was labeled as outlined above using Cy 2 and added to each sample as an internal standard. One Cy 3 and one Cy 5 labeled sample were added to each gel, as defined by the experimental randomizing procedures. To each sample, an additional 175 µg of unlabeled sample were added; the volume was increased to a total of 450 µL using sample

preparation buffer. The reducing agent DTT (dithiothreitol; 1.3 mg per tube) and IPG (immobilized pH gradient) buffer (0.5%) were added, and the samples incubated and mixed in the dark at room temperature for approximately 1 h. Samples were then applied to 24 cm pH 4 – 7 IPG strips (GE Healthcare) and focused for 60 kVh using the following protocol: 12 h rehydration at 30 V; 1 h step and hold at 500 V; 7 hour gradient to 1,000 V; step and hold at 1,000 V for 1 hour; gradient to 8,000 V for 3 h; step and hold at 8,000 V until 60 kVh. Strips were then reduced and equilibrated using 10 mg/mL DTT (15 min) followed by 25 mg/mL iodoacetamide (15 min). The IPG strips were overlaid on 24 x 26 cm 8 – 16% gradient Tris-HCl pH 8.8 pre-cast gels (NextGen Sciences, Ann Arbor, MI) cast between low fluorescing glass plates. Approximately 1 mL of agarose was applied to fix the gels and a Cy 2 labeled molecular weight marker was applied adjacent to the acidic side of the strip. The gels were then run 1 – 2 W per gel overnight (~22 – 24 hours) at 20°C until the marker dye ran off the gel. Gels were then imaged with a Typhoon 9400 scanner and processed using DeCyder v6.5 (GE Healthcare) BVA batch processor tool. Gels were post-stained using a silver stain as described previously (Shevchenko, Wilm, Vorm, & Mann, 1996). Images were uploaded to DeCyder (version 6.5) and spurious image objects (water spots, streaks and mismatches) were identified and excluded from further analysis. Following allocation of changed proteins, individual spots were manually inspected and excluded from analysis if they fell outside acceptable parameters for peak height, area, and slope.

Gel picking and protein digestion. Pick lists were generated by selecting proteins whose expression was statistically changed in the two growth conditions ($p < 0.05$) following digital image analysis using Decyder, and the corresponding spots were automatically picked using an Ettan Spot Picker (GE Healthcare) with Ettan Spot Pick Software v.1.1. Spots were delivered in 100 mM ammonium bicarbonate to a 96-well plate and digested using established protocols. Briefly, gel pieces were sequentially dried using three exchanges of 100% acetonitrile followed by a 10-minute SpeedVac (Savant) drying. Gel pieces were rehydrated in 40 μ L of 10 ng/ μ L trypsin in 100 mM NH_4HCO_3 on ice for 45 minutes. The supernatant was removed and replaced with 100 mM NH_4HCO_3 and digested at 37 °C overnight. Peptides were then extracted using 50% acetonitrile/0.1% TFA (trifluoroacetic acid) for 30 minutes at 37°C. The peptides were microextracted using Omix C18 tips (Varian, Palo Alto CA) following the manufacturers instructions, and then deposited on a stainless-steel target plate in a matrix consisting of 10 mg/mL 2,5-dihydroxybenzoic acid.

MS and database searching. Mass spectra were acquired using a Voyager DE-STR matrix-assisted laser desorption/ionization time-of-flight MS (Applied Biosystems, Foster City, CA) in positive reflector mode with delayed extraction using the following parameters: laser energy, 1400 arbitrary units; mass range, 500 – 5,000 Da; 120 nsec delay, 100 laser shots per spectrum. External calibration was conducted using a standard peptide mixture (bradykinin, insulin B chain, P14R, and ACTH), and internal calibration was carried out using trypsin autolysis peaks. Data were processed in Data Explorer v1.1 (Applied Biosystems,

Foster City, CA) using noise reduction (2 standard deviations) and peak deisotoping. Peak masses were searched using the Mascot online search engine (<http://www.matrixscience.com>) with the following settings: Database, NCBI entire database (5.6 million entries); no missed cleavages; monoisotopic peaks; no fixed modifications; variable modification of methionine oxidation; error tolerance of 150 ppm. Protein identifications were mapped back to the gel using Decyder v6.5. Database and literature searches were used to further characterize and classify the proteins identified by MALDI-TOF MS. Where ambiguous names were encountered, BLASTp searches (Altschul et al., 1997) were used to identify homologous proteins from orthologous species.

Results

Image analysis of 24 cm 2D-DIGE gels loaded with protein of *S. wittichii* RW1 cells grown on either dibenzofuran or acetate revealed 937 unique spots. Differential in-gel analysis of individual gels determined gel-specific parameters for selection criteria and allowed visual examination of changes between growth on the two substrates (Figure A.1). Of the 937 identified spots, 595 were matched between all the gels used to statistically compare the quantitative abundance of proteins. Statistical analysis compared triplicate biological observations for each condition, normalized to the internal pooled standard (Figure A.2).

Crude cell lysates from *S. wittichii* RW1 grown on dibenzofuran showed that, of all proteins observed, 23 proteins were modulated in response to changes in culture conditions. These candidate biomarkers of metabolic activity and

phenotype were observed in at least 6 of 9 DIGE images and were modulated as follows: 17 showed an apparent increase and 6 an apparent decrease (Figure A.2). These proteins, along with 21 proteins selected due to their high abundance in both growth conditions, were further analyzed and identified using mass spectrometry (Figure A.3). A Mascot search of the entire NCBI database using mass spectral data generated by peptide mass fingerprinting identified 23 of the 44 proteins (52%). All protein identifications corresponded to the genome of *S. wittichii* strain RW1. Among the 17 proteins upregulated during growth on dibenzofuran, 8 were successfully identified (Table A.1). Among the 6 proteins downregulated during growth on dibenzofuran, 3 were successfully identified

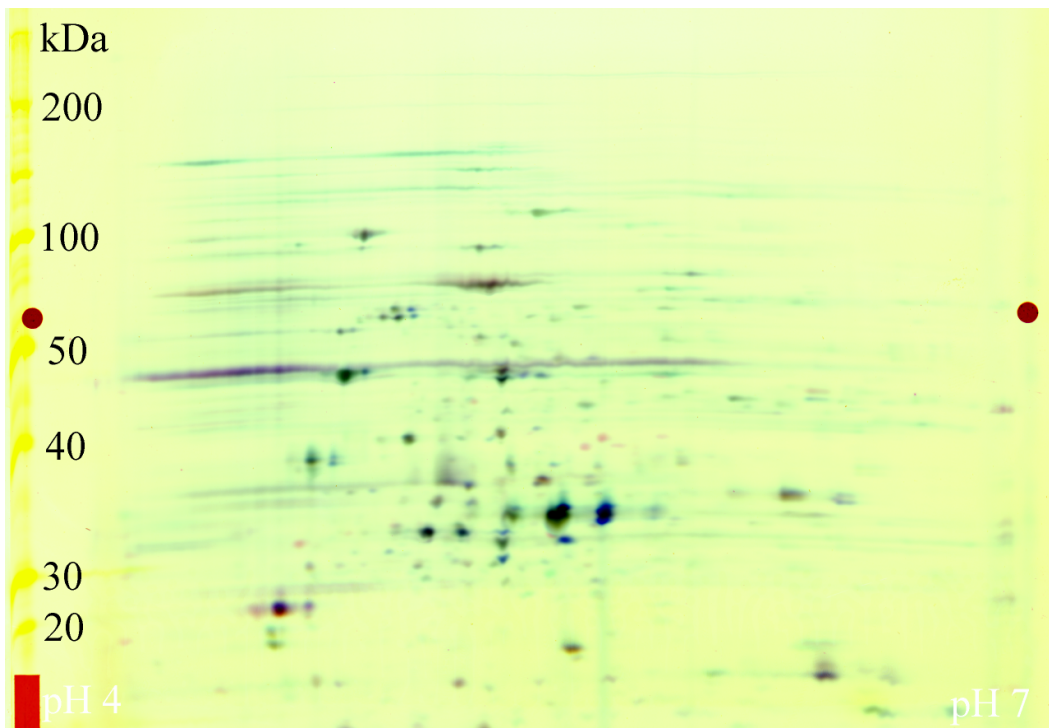


Figure A.1. A representative 2D-DIGE gel of RW1 showing the contrast in protein expression between cells grown on acetate and dibenzofuran. Proteins that are expressed equally for the two conditions appear white, differentially expressed proteins are shown in green (higher on acetate) and red (higher on dibenzofuran).

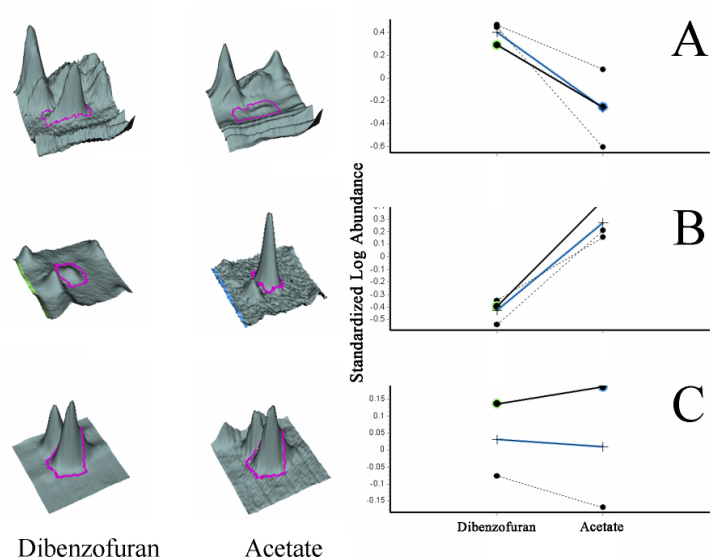


Figure A.2. Statistics and spot information for three representative spots from DIGE analysis of the RW1 crude cell proteome. Following visual examination of the spot characteristics and matching parameters, spots were identified as increased (A, gi|148555952 glyoxalase/bleomycin resistance protein/dioxygenase), decreased (B, unidentified protein) or unchanged (C, gi|148553776 OmpA/MotB domain protein).

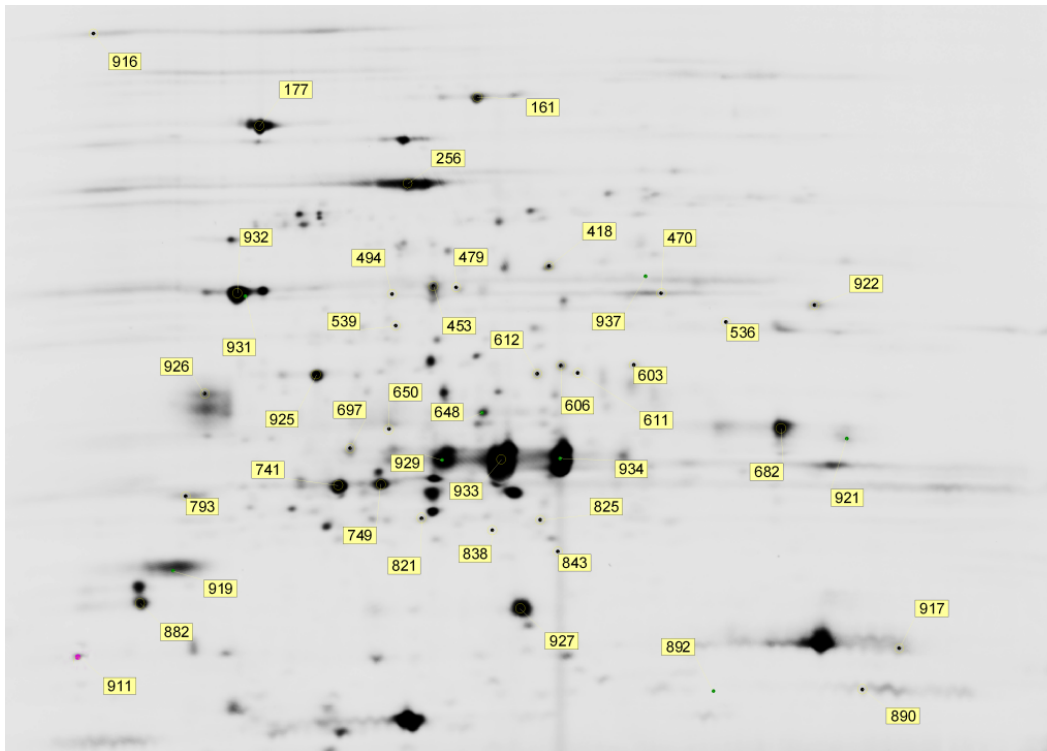


Figure A.3. A scanned and cropped 24 cm gel showing annotations for spots selected for further analysis. The spots were selected due to (i) their relative increase or decrease during growth on the selected substrate or (ii) their high abundance in both samples. The spot numbers are identifiers for the quantitative information and identity (see Tables A.1-3). The approximate pH range (horizontal) is 4 – 7, and the approximate molecular weight (vertical) is 110 kDa to 10 kDa.

(Table A.2). An additional 12 proteins were identified whose expression level remained unchanged regardless of culture conditions (Table A.3).

Of the 7 identified proteins increased during growth on dibenzofuran, 3 were directly related to the dibenzofuran degradation pathway (Figure A.4); the others were involved in downstream metabolic processes (catechol 1,2-dioxygenase, adenylyhomocysteinase), cell growth (elongation factor Ts) and cell protection (cold shock DNA-binding domain protein, alkyl hydroperoxide reductase). The three identified proteins whose expression was decreased (fumarylacetoacetate hydrolase, TonB-dependent receptor, and acyl-CoA dehydrogenase) are involved in biosynthesis, catabolism and transport. The unchanged proteins represented basic cell functions, although biosynthesis, catabolism and transport proteins dominated the identities.

Discussion

DIGE and 2D electrophoresis are an accepted strategy for mining microbial proteomes for biomarkers related to a number of processes (Giometti, 2006; Hufnagel & Rabus, 2006; Mazzoli et al., 2007). The complete protein content of *S. wittichii* RW1 consists of approximately 5,000 putative proteins from the bacterial chromosome and two megaplasms. Using simple extraction and purification techniques followed by DIGE, over 500 protein spots were resolved on a large (24-cm) 2-dimensional gel and matched between the three biological replicates, representing approximately 10% of the entire protein

content. Of these 500 proteins, 23 were found to be regulated in response to growth condition changes.

Table A.1. Proteins identified as being increased in *S. wittichii* RW1 during growth on dibenzofuran when compared to growth on acetate. The spots were identified on a minimum of 6 gel images.

Master Number ^a	Protein Accession ^b	Gene Locus	Protein Name	Average Ratio ^c	T-test ^d
919	148556489	Swit_3587	Alkyl hydroperoxide reductase/thio specific antioxidant/Mal allergen	1.52	0.099
937	148550856	Swit_4897	Dioxin dioxygenase, alpha subunit	1.59	0.38
922	148553900	Swit_0977	Catechol 1,2-dioxygenase	1.8	0.078
921	148555809	Swit_2901	Putitative cold shock DNA-binding domain	1.89	0.078
539	148555586	Swit_2674	Adenosylhomocysteina se	1.97	0.062
418	148553385	Swit_0461	Elongation factor Ts	2.10	0.05
934	115279619	Swit_3055	<i>meta</i> -cleavage pathway hydrolase	2.54	0.055
603	148555952	Swit_3046	Glyoxalase/Bleomycin resistance protein/dioxygenase	3.88	0.031

^a Arbitrary identifier for spot location (see Figure 2); ^b NCBI gi| number;

^c Using internal standard as 1.0, >1 is an increase and <1 is a decrease in abundance;

^d Student's *t*-test comparing spot intensity for acetate vs. dibenzofuran grown cells

Table A.2. Proteins identified as decreased in *S. wittichii* RW1 during growth on dibenzofuran as compared to acetate. The spots were identified on a minimum of 6 gel images.

Master Number ^a	Protein Accession ^b	Gene Locus	Protein Name	Average Ratio ^c	T-test ^d
821	148551036	Swit_5089	Fumarylacetoacetate hydrolase	-1.99	0.063
470	148553574	Swit_0650	Acyl-CoA dehydrogenase domain	-1.74	0.031
161	148550568	Swit_5129	TonB dependent receptor	-1.59	0.086

^a Arbitrary identifier for spot location (see Figure A.2)

^b NCBI gi| number

^c Using internal standard as 1.0, >1 is an increase and <1 is a decrease in abundance

^d Student's *t*-test comparing spot intensity for acetate vs. dibenzofuran grown cells

Table A.3. Highly abundant proteins identified as unchanged in *S. wittichii* RW1 during growth on dibenzofuran or acetate. The spots were identified on a minimum of 6 gel images.

Master Number ^a	Protein Accession ^b	Gene Locus	Protein Name	Average Ratio ^c	T-test ^d
916	148556048	Swit_3144	TonB-dependent receptor	-1.33	0.23
929	148553835	Swit_0912	4-oxalocrotonate decarboxylase	-1.26	0.25
256	148556278	Swit_3376	Chaperonin GroEL	-1.21	0.5
453	148553821	Swit_0898	Phenylpropionate dioxygenase, ferredoxin reductase subunit	-1.17	0.5
741	148553833	Swit_0910	Alpha/beta hydrolase fold	-1.08	0.9
933	148555961	Swit_3055	<i>meta</i> -cleavage product hydrolase	-1.03	0.85
932	148553776	Swit_0853	OmpA/MotB domain	-1.00	0.93
749	148553834	Swit_0911	4-oxalocrotonate decarboxylase	1.09	0.3
927	148555704	Swit_2794	Opacity protein and related surface antigen-like protein	1.09	0.82
648	148550878	Swit_4921	3-keto-5-aminohexanoate cleavage enzyme	1.15	0.45
177	148555643	Swit_2731	Aconitate hydratase 1	1.23	0.44
479	148550877	Swit_4920	FAD-dependent pyridine nucleotide-disulphide oxidoreductase	1.57	0.5

^a Arbitrary identifier for spot location (see Figure A.2); ^b NCBI gi| number;

^c Using internal standard as 1.0, >1 is an increase and <1 is a decrease in abundance; ^d Student's *t*-test comparing spot intensity for acetate vs. dibenzofuran grown cells

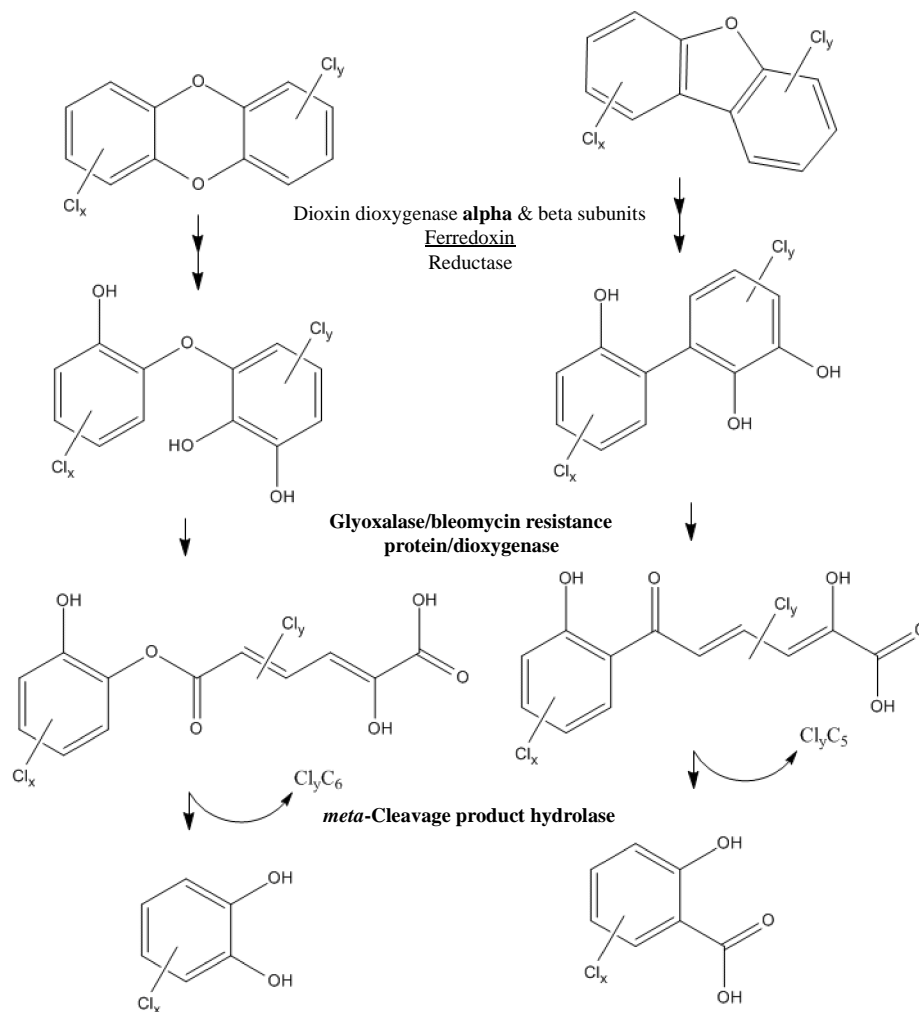


Figure A.4. The dioxin (left) and dibenzofuran (right) degradation pathways. Enzymes catalyzing individual reactions are written in the center. Bold typeface indicates a protein that was found to be increased in expression. The ferredoxin (underlined) is located in an operon whose expression was found to be decreased. Cl_yC_5 and Cl_yC_6 represent potentially chlorinated aliphatic moieties.

This study identified a number of proteins that are related to dioxin/dibenzofuran degradation (e.g., dioxin dioxygenase, meta-cleavage product hydrolase, 2,3-dihydroxybiphenyl 1,2-dioxygenase). Other proteins were identified that showed increases in abundance but whose role was not directly

related to the dibenzofuran degradation pathway. The increase in the presence of antioxidants such as alkyl hydroperoxide reductase suggest that there is an increasing stress upon the bacterial cell during growth on dibenzofuran, perhaps due to a change in catabolism resulting in an increase in endogenous peroxide generation (Seaver & Imlay, 2001). Increases in a cold-shock DNA-binding protein may be further evidence of an increased cellular stress (Susin, Baldini, Gueiros-Filho, & Gomes, 2006). However, proteins of the cold shock family and related ones are also known to have transport and protein processing roles (Schiene-Fischer, Habazettl, Schmid, & Fischer, 2002).

Among the proteins in the dioxin degradation pathway, the most prominent on the gel was the meta-cleavage product hydrolase. This identification was produced from two adjacent spots, likely representing an artifact due to the protein's extremely high expression or a reflection of the presence of multiple isoforms or a modified enzyme. *S. wittichii* RW1 has three known isoforms of this meta-cleavage product hydrolase (Seah et al., 2007). The one identified in the present study is the product of the Swit_3055 locus, a gene also known as DxnB2 (Seah et al., 2007). Its identification in this study corroborates previous findings (Seah et al., 2007). Unlike the dioxin dioxygenase, this gene is found on the chromosome.

The glyoxalase/bleomycin resistance protein/dioxygenase identified in this study is also annotated as a 2,3-dihydroxybiphenyl-1,2-dioxygenase located on the chromosome at the Swit_3046 locus. Again, there are multiple isoforms of this enzyme found both on the chromosome and the megaplasmids. The KEGG

dioxin degradation pathway identifies Swit_4182 as the dihydroxybiphenyl dioxygenase involved in biphenyl metabolism and Swit_4902 as the trihydroxybiphenyl dioxygenase in both dioxin and dibenzofuran metabolism. These genes have all been annotated as glyoxalase/bleomycin resistance protein/dioxygenase. The increased expression in response to dibenzofuran suggests that the Swit_3046 dioxygenase plays a more important role in dibenzofuran degradation in vivo.

The high degree of redundancy in the dioxin and dibenzofuran degradation pathways, i.e. the presence of multiple ring-hydroxylating alpha and beta subunits, glyoxalase/bleomycin resistance protein/dioxygenases, and meta-cleavage product hydrolases, remains to be explained. One possibility is that the various isoforms have different affinities for chlorinated metabolites that would result from chlorinated dioxins and furans. Further experiments are needed to fully distinguish the roles of these enzymes in *S. wittichii* RW1 degradation pathways.

Although not directly implicated in dioxin degradation, the fumarylacetoacetate hydrolase is also of interest because the gene encoding this protein (Swit_5089) flanks the ferredoxin Fdx1 (Swit_5088) that has been identified as part of the electron supply chain supporting dioxin dioxygenase activity. Of the multiple isoforms of this enzyme, Fdx1 was found to function in vitro with the dioxin dioxygenase (Armengaud & Timmis, 1997). The electron supply chain also contains two isofunctional reductases (Bunz & Cook, 1993a). Neither the ferredoxin itself nor the reductases could be identified. In previous

studies, the reductase was present as a much smaller fraction of the soluble cell proteome than either the ferredoxin or the dioxin dioxygenase (Bunz & Cook, 1993a), so gel-based methods may not be sensitive enough to detect this protein. If transcription of the ferredoxin is linked to the other genes at that locus, as is predicted (Armengaud & Timmis, 1997), the decreased expression in response to dibenzofuran suggests that another ferredoxin is more important in the dioxin degradation pathway *in vivo*. Further studies are needed to confirm this hypothesis.

The detection of the dioxin dioxygenase alpha subunit and related enzymes in both acetate- and dibenzofuran-grown cells is potentially of importance for the field of bioremediation because it suggests an avenue of biostimulation. When utilizing *S. wittichii* RW1 as a bioremediation agent, it may be possible to induce the expression of the dioxin degradation pathway using acetate. Induction of the dioxin degradation pathway has not been observed when *S. wittichii* RW1 is grown on glucose or rich medium (Halden et al., 2005a), and growth in a complex environmental medium (landfill leachate) was correlated with a decrease in copy number of the gene encoding the dioxin dioxygenase alpha subunit (Hartmann, Badalamenti, Krajmalnik-Brown, & Halden, 2012). Previous studies using *S. wittichii* RW1 to transform chlorinated dioxins in soil or fly ash have observed a progressive decrease in degradative activity (Halden, Halden, et al., 1999b) or viable cells (Nam et al., 2005), respectively. The addition of acetate may generate sufficient relevant protein biomass to catalyze the

successful degradation of dioxin and dioxin-like compounds in environments bioaugmented with *S. wittichii* RW1.

Proteomic technology has emerged in microbiology more rapidly than in other fields for several reasons. The relatively small genomes code for relatively limited proteomes featuring no or very limited post-translational modifications compared to higher organisms (Gupta et al., 2007). Furthermore, microbes are easily controlled and manipulated in the laboratory, both during growth and gene expression. These factors will continue to drive biomarker discovery in microbial proteomes, including phenotypic biomarkers informing on the degradative activity of biomass produced for bioaugmentation of contaminated environments. Furthermore, the field of bioremediation can benefit from methods suitable for monitoring microbial biomarkers in field samples to inform on progress in site bioremediation. This study highlights a number of proteins that were changed in response to dibenzofuran exposure, opens the door to a greater understanding of how *S. wittichii* RW1 performs and regulates the degradation of dioxins, and suggests ways to enhance the biodegradation of dioxins.

APPENDIX B.
PERMISSIONS

This dissertation contains previously published work in which I am listed as first author along with my co-authors. All of my co-authors have given their permission to use these articles in my dissertation.

Permission has also been obtained to reprint copyrighted material. "Challenges of Detecting Bioterrorism Agents in Complex Matrices" has been licensed from Springer under the license number 2907781134193. Permission to reproduce "Identification of Putative Biomarkers for Toluene Degrading *Burkholderia* and Pseudomonads by MALDI-TOF MS and PMF" was obtained from Dr. Kazumitsu Ueda, the editor in chief at the Japan Society for Bioscience, Biotechnology, and Agrochemistry. Under the American Society for Microbiology author rights, authors retain the rights to reuse their articles in full in a dissertation or thesis; thus, I have the right to use "Quantitative PCR for Tracking the Megaplasmid-Borne Biodegradation Potential of a Model Sphingomonad." "Analytical methods for the detection of viruses in food by example of CCL-3 bioagents" has been licensed from Springer under the license number 2907781468899.

