DEVELOPMENT OF A PROTEIN-BASED SENSOR ASSAY FOR RAPID CLASSIFICATION OF COMPLEX BIOLOGICAL SAMPLES

by

Timothy Kyle Hamerly

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2016

DEDICATION

To my family and friends. Without your support I would not be who I am today.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Brian Bothner, for all that he has done for me throughout graduate school. Your support and guidance has been instrumental in my success as a student and will continue to be of benefit as I progress through my career. I would also like to thank my graduate committee for providing guidance and advice not only on this dissertation, but also on other projects. I have had many collaborators throughout my time in graduate school, and without their help I would not have been able to accomplish all that I have. Finally, I would like to thank members of the Bothner lab, past and present, for all their help in training me, providing helping hands when needed, or just conversing and making the day go by that much faster.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

Metabolomics, one of the core “omics” fields within the umbrella of systems biology, is the study of the small molecules which can be used to characterize the state of an organism. Metabolites are constantly being transformed inside a cell in direct response to stimuli around them. This makes the metabolome the most dynamic of all the omics fields and is considered to be a direct readout of the cells state at any given time. Although highly informative, the metabolome is inherently difficult to study, with thousands of known metabolites, any of which could be important for classifying a cell into a healthy or diseased state. Techniques such as mass spectrometry are well suited to study the metabolome and have been used to successfully classify cells by identify markers for a given disease state. However, current methods require lengthy analysis times due in part to the complexity of the metabolome. The research presented in this dissertation highlights a new and promising methodology which improves classification and speeds marker discovery. Making use of a protein found in animals which has evolved to selectively bind metabolites, an assay was developed which better classified samples compared to current methods used in the field of metabolomics. This improved classification was achieved with an overall decrease in analysis time. The implementation of this method in the study of complex biological systems would have an immediate impact in academic and medical research.

CHAPTER 1

INTRODUCTION

Metabolomics

Metabolomics belongs to the broader field of study known as system biology and has become a heavily studied omics science in the past decade, as technologies have advanced and our understanding of how the metabolic state predicts the phenotype of an organism. For much of the 19th century biology followed what is known as the central dogma of molecular biology; information flowed in a unidirectional path from gene to mRNA to proteins, with metabolic pathways being the ultimate readout of the information (Hollywood et al. 2006; Jansen et al. 2012). However, it is now known that the different levels of cellular processes are in fact intimately networked with feedback-loops, acting as a dynamic system which is far more complex than originally thought. The ability to study the whole transcriptome or proteome has allowed biologists to gain a global understanding of cellular response to stimuli, rather than studying each gene or protein as an individual entity.

Metabolomics is one of the core omics approaches under the heading of systems biology and is defined as the study and quantification of the small molecules, termed metabolites, which exist in a cell or tissue in a given state (resting, stressed, diseased, etc.) (Goodacre 2005; Hollywood et al. 2006). In recent years, this has been extended to include the metabolites which exist as a result of cellular processes, where cells or tissue

may no longer be present (Urpi-Sarda et al. 2015). The advantage to studying the metabolome compared with other, upstream omics is that the metabolome is far more dynamic; changes take place each second as enzymes consume and produce thousands of molecules in an elegantly orchestrated process. In contrast, a transcriptomic response, leading to changes in mRNA production or degradation often plays out on the time scale of hours (Hollywood et al. 2006; Urpi-Sarda et al. 2015). The dynamic nature of the metabolome has enabled biologists to use it to rapidly determine the state of a cell, be it diseased or healthy, by identifying specific markers of these various states. Using this approach, metabolomics has already been used in a wide variety of studies to discover markers that are now being developed for clinical tests (Davidovic et al. 2011; Patti et al. 2012a; Siuzdak 2014; Urpi-Sarda et al. 2015). Global metabolic differences have allowed biologists to classify cells based on disease progression, or recovery, due in part to the fact that the metabolome is dynamic and readily changing on a second time scale (Bier et al. 2000; Szymanski et al. 2009; Motta et al. 2010). Metabolomics has been shown to be a powerful tool to study biological processes and as technology continues to advance, so too will this field.

LC-MS Based Metabolomics

Although other techniques such as nuclear magnetic resonance (NMR) are powerful tools for studying the metabolome, the most commonly used approach employs mass spectrometry. Specific interest has been given to liquid chromatography coupled to mass spectrometry (LC-MS). Advances in both the chromatography columns used for

separation and the MS instruments now allow biologists to obtain vast amounts of information on a system in a relatively short period of time. Metabolomic profiling of hundreds, and even thousands of metabolites, in a wide range of concentrations, has been used to describe the phenotype of a cell in a given state. The sensitivity of current instruments allows them to detect minute quantities of molecules, enabling them to be used as biomarker discovery tools, or monitor drug metabolism in pharmacology studies (Goodacre 2005; Cuperlovic-culf et al. 2013; Worley and Powers 2013; Johnson et al. 2015). The broad applicability and robustness of MS instruments has enabled a new age of possibilities in the realm of metabolomics, and opens up the field of systems biology.

Mass Spectrometry

The principles of mass spectrometry has been around for 100 years, when Sir Joseph J. Thomson discovered that atomic particles could be manipulated with electric fields (Thomson 1899; Thomson 1907; Thompson 1913). Since Thomson's first experiments the field of mass spectrometry has grown considerable. Improvements in detectors and mass analyzers have greatly improved sensitivity and resulted in better resolution respectively. From a metabolomics standpoint, the introduction of the electrospray ionization (ESI) source in the mid 1980's was critical in allowing a MS to be coupled to an LC (Yamashita and Fenn 1984). With these improvements mass spectrometry has evolved to become the go-to method for rapid analysis of the metabolome.

Although modern mass spectrometers contain a complex array of electronics which enable high sensitivity and high resolution, at the core they are made up of three

parts: an ion source, mass analyzer, and detector. In order for a MS to work, ionized molecules must be generated which can then be manipulated by the instrument and subsequently detected. These ions are produced in the source, which can be one of many different types. However, the most commonly used source for metabolomic data acquisition is the ESI source. When coupled to an LC, the ESI source allows liquid to be directed infused into the high vacuum instrument, and using high voltage and a sheath gas, produce a plume of very small charge droplets known as the Taylor cone (Siuzdak 2006; Gross 2011). Temperature is manipulated in the source to facilitate rapid evaporation of the solvent droplets containing the analytes. This process leads to charge(s) being transferred from the droplet surface to the molecules inside, thus producing a charged species. The charged molecules are then directed toward the high vacuum portion of the instrument using a series of lenses, while the uncharged species simply go to the vacuum and away from the instrument.

After ionization, the charged species move into the mass analyzer, the portion of the instrument responsible for determining the mass of molecules. Mass analyzers used for metabolomics experiments are generally a quadruple time-of-flight (Q-TOF) or a triple quadruple (QQQ), although other mass analyzers may be used (Siuzdak 2006; Gross 2011). A labeled diagram of an ESI Q-TOF MS is shown in Figure 1.1. Pictures of the actual LC ESI Q-TOF MS system used to collect data for this work are shown in Figure 1.2. Quadruple mass analyzers (in both Q-TOF and QQQ) use oscillating radio frequencies in combination with direct current voltages to separate ions based on their mass-to-charge ratio (m/z) before sending them towards the detector (Siuzdak 2006;

Gross 2011). A TOF mass analyzer sends a pulse of ions of various m/z's into a tube which is under high vacuum at the same time with the same amount of kinetic energy. Molecules which have a lower m/z will move through the tube more quickly than higher m/z molecules, therefore separating them. Most TOF mass analyzers have a reflectron at the end of the tube which reverses the direction of ions and doubles the distance they must travel while simultaneously diminishing the spread of flight times for ions of the same m/z, thereby increasing the separation of the ions and increasing the resolving power of the instrument. Resolving power or resolution of mass analyzers is defined as the ability to separate two ions of similar m/z. In both cases, the mass analyzer is used to separate out many m/z's which correlate to different molecules so that they can then be detected individually.

The final major component of a mass spectrometer is the detector, which records the m/z of a molecule as well as the relative intensity of that ion. Detector technology has seen some improvements in the last few decades which have allowed for increases in resolution and sensitivity for the detection of low abundance ions. Most instruments use an electron multiplier to detect ions from the mass analyzer. An electron multiplier is made of a series of aluminum oxide dynodes which when struck by ions from the mass analyzer will emit electrons, which then strike another dynode, emitting more electrons and so on, until the signal of electrons creates a large difference in current which can be detected (Siuzdak 2006; Gross 2011). The most important thing to note about detection in a mass spectrometer is that this process is not inherently quantitative, that is, a signal of X cannot be directly translated into Y grams of a compound being present. However, mass

spectrometers can be used to quantitate ions when a standard curve is generated and compared with the signal in the sample.

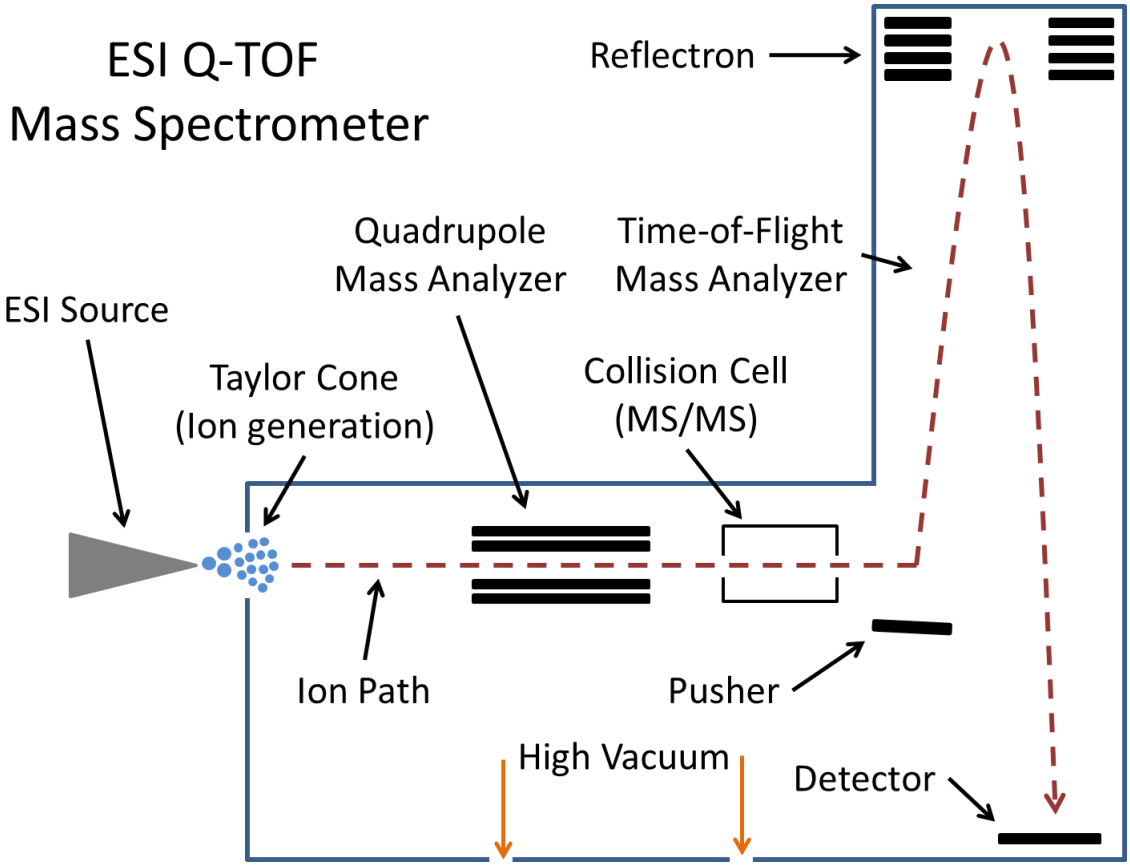

Figure 1.1. Schematic diagram of an ESI Q-TOF MS. Molecules enter the instrument through an electrospray ionization source, which coupled to a high voltage, generates charged droplets. There droplets form a Taylor cone due to the voltage applied to the source. The charged droplets slowly lose solvent until only a single molecule is left, which accepts the charge of the droplet. These charged molecules than move to the first mass analyzer, a quadrupole, which filters ions using a combination of direct voltage and radio frequencies. Ions allow through than move through a collision cell which is used to fragment ions of interest if tandem mass spectra (MS/MS) is require, else the cell is inactive. Ions then move to the time of flight tube which pushes the ions using a high voltage into the flight tube. These ions move through the flight tube at an initial velocity based on their m/z, with smaller m/z molecules moving faster. Upon reaching the top of the tube ions are redirected using a relectron or a series of electric fields which push the ions back the opposite direction. This double the distance ions must travel, thereby increasing the resolution. Finally, ions reach the detector where they produce a signal which is record and displayed by the acquisition software.

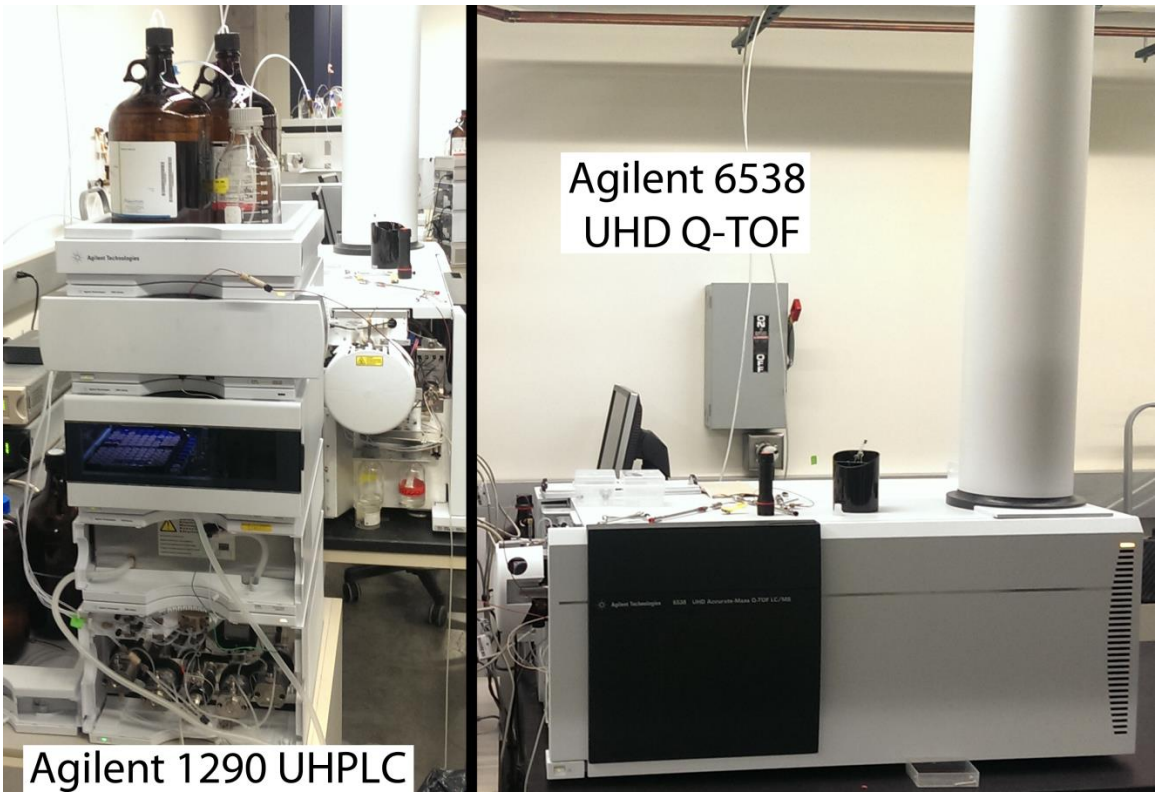

Figure 1.2. Photographs of the LC-MS instrumentation used in this dissertation. On the left is an Agilent 1290 UHPLC which includes two binary solvent pumps, a temperature controlled autosampler, and a temperature controlled column compartment. On the right is an Agilent 6538 ESI UHD Q-TOF mass spectrometer.

Applications of LC-MS Based Metabolomics

Due to the high sensitivity, large dynamic range of detection (ability to detect ions across a broad range of concentrations [10^3-10^6 difference in relative intensity]), tolerance of salts commonly found in biological samples, and high through-put nature of LC-MS, mass spectrometry has become a widely used technique to study the metabolome. The ability to detect 1000's of features in a single LC-MS run has enabled global metabolic profiling which can be used to categorize samples based on profiling or fingerprinting (Patti et al. 2012b). Further investigation of these categorized samples by statistical analysis can identify markers of a given disease or cellular state, leading to

rapid tests for use in clinical applications (Lin et al. 2011; Siuzdak 2014). Extraction of metabolites has been a thrust of recent efforts in the field, driven by advances made in MS technology that allow for detection of extremely low abundance molecules. These efforts have allowed for greater coverage of the metabolome, leading to better understands of the biological systems we study.

Current Challenges

LC-MS based metabolomics has been shown to be beneficial in a great number of studies; however, there are several limitations and challenges with the technique. The idea of finding a biomarker which can accurately and reliably predict the onset or status of a disease has long been the goal of biologists and medical doctors alike. However, the reality is that this is a very complex process, for a multitude of reasons: i) data sets contain a large number of features thanks to the advances in technology, of which most are not changing, ii) once a potential marker has been found, there is no easy way to go from molecular feature, to known identity, especially if no standard for the compound exists, and iii) comparison of data sets can be difficult because the technique is not inherently quantitative, meaning variations in sample size cause drastic differences in the intensities of the features.

The practical side of dealing with large datasets containing 1000's of metabolites generated from these MS-based studies has presented a number of challenges to the field which must be overcome. Dealing with the datasets themselves has become easier thanks to advances made in programs designed at handling LC-MS output (Smith et al. 2006; Pluskal et al. 2010; Tautenhahn et al. 2011). Identification of metabolites has also been

made easier thanks to repositories of molecules with monoisotopic masses (Smith et al. 2005; Tautenhahn et al. 2012). For instance, when looking to profile an organism, databases like the Biocyc server, can be of great help in limiting identifications to those present in a given organism (Karp et al. 2005). Finally, the use of internal standards can make cross comparisons between data sets possible thanks to the reproducibility of modern day MS instruments.

As technology continues to advance LC-MS instrumentation, there exists an equal need to pursue methods which can handle the ever increasing complexity of datasets, reducing it to something meaningful and impactful to biology. Small molecules are thought to be the direct readout for an organism's state, be it healthy, diseased, or somewhere in-between (Johnson et al. 2015). This readout is becoming more thorough as MS technology continues to advance, but there still exists a bottleneck in processing and understanding the datasets (Heinonen et al. 2012; Johnson et al. 2015). This bottleneck exists because many of the metabolites in an organism do not drastically change between states making it difficult to tease out the truly important molecules. Molecular sensors, both synthetic and biologically based have sought to overcome this problem by drastically reducing the number of detected small molecules and looking at only those which are important for classification (Anslyn 2007; Kubarych et al. 2010). Coupling a synthetic or biological sensor to a high-throughput technique such as LC-MS could drastically improve the field of molecular sensing.

Serum Albumin

Serum albumin (SA) is the most abundant blood protein in the plasma of animals and has been studied extensively for its unique ability to bind a wide range of small molecules. It was first thought that albumin was discovered by the Greek physician Hippocrates, who noted that foamy urine indicated chronic kidney disease and was caused by the presence of albumin in the urine (Peters Jr. 1996). However, it wasn't until the turn of the 20th century when SA crystallization and purification began (Hopkins and Pinkus 1898). In more recent years, SA has been studied from a variety of mammals where 61% of the sequence is conserved among bovine, rat, and human and has the common motif of a repeating series of disulphide bridges (He and Carter 1992). SA has also been shown to be important in maintaining colloidal osmotic blood pressure (Curry et al. 1999).

Structural models based on X-ray crystallography and binding studies have been done to better understand this complex protein (He and Carter 1992; Ho et al. 1993; Curry et al. 1998; Petitpas et al. 2001a; Zunszain et al. 2003). Mammalian SA proteins were found to be heart shaped (Figure 1.3). From these studies, it was found that SA proteins possess multiple binding pockets (between 5 and 7 depending on the ortholog), with each pocket having a different chemical environment (Curry et al. 1999; Petitpas et al. 2001b; Petitpas et al. 2001a; Zunszain et al. 2003; Fielding et al. 2005; Varshney et al. 2010; Nafisi et al. 2011; Jupin et al. 2013; Jupin et al. 2014). Some pockets have charged amino acid residues, while other pockets were deep and made up of primarily non-polar amino acids. Each pocket on the protein appeared to be specialized towards a class of

compounds, be it lipids, metal ions, or aliphatic molecules (Varshney et al. 2010). SA has also been found to transport toxic metabolites from normal bodily functions such as bilirubin, a toxic by product of heme degradation (Curry et al. 1999). Another example of SA for detoxification is in the Chinese cobra which has been shown to protect the snake from succumbing to its own venom (Shao et al. 1993). Stemming from this idea, methods which use SA proteins to remove targeted molecules from the bloodstream have been development (Mitzner et al. 2001). These pockets have also been found to be important for the transport of pharmaceuticals in the blood stream (Petitpas et al. 2001a; Varshney et al. 2010).

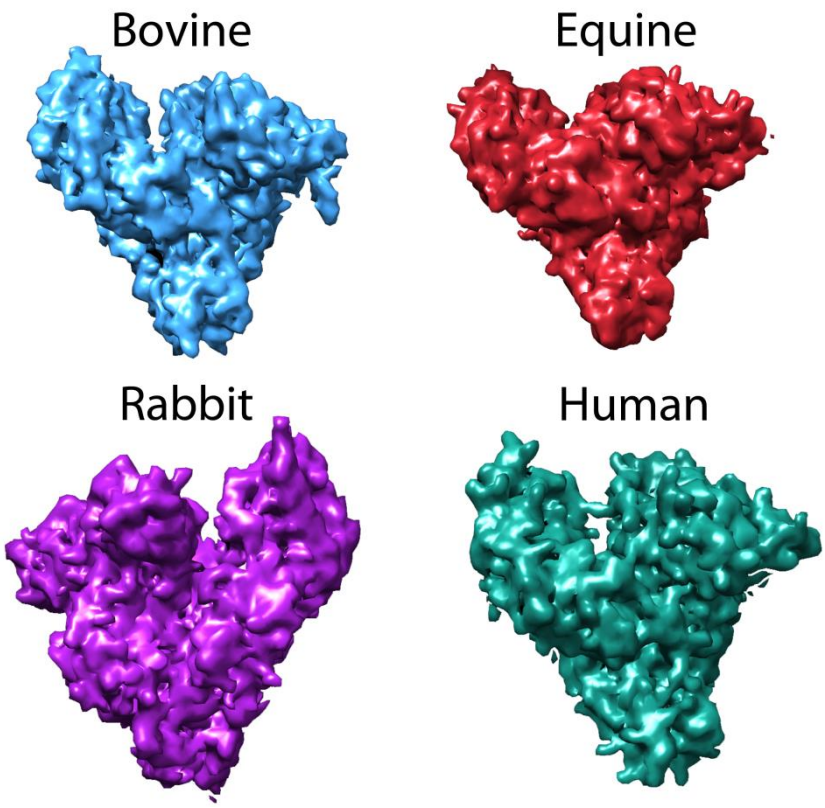

Figure 1.3. Surface model protein structures of four mammalian serum albumins. Models were generated using Chimera's multiscale model function from PBD files (Pettersen et al. 2004). PDB files for each are as follows: Bovine [4F5S], Equine [4f5U], Rabbit [3V09], and Human [4G03] (Bujacz 2012; Majorek et al. 2012)

Many studies have found evidence that pockets among the protein were allosterically connected, where changes in the protein structure brought about by a multitude of factors had effects on the binding of other pockets (Curry et al. 1998; Petitpas et al. 2001b; Zunszain et al. 2003; Varshney et al. 2010). One study on the human ortholog showed that as myristic acid bound to the protein in two pockets a significant conformational change was seen in the protein which opened pockets that were previously not accessible (Curry et al. 1998; Curry et al. 1999). This change in conformation allowed for further binding of fatty acids. Changes in protein structure were also found to be regulated by solution pH. Five different conformations were identified in the bovine ortholog, two of which were physiologically relevant depending on the blood pH (Dockal et al. 2000; Curvale 2009). Further evidence of allosteric regulation has been shown through studies on the binding of fatty acids and drugs. One study showed that the binding of (S)-lorazepam acetate, clonazepam, or (S)-uxepam selectively increased the binding of (S)-warfarin to human SA (HSA) (Fitos et al. 1986; Fitos and Simonyi 1988). Another study showed that binding of a ligand in one site resulted in the instantaneous loss of a dye molecule from another site and was attributed to a change in multiple domains of the protein (Krishnakumar and Panda 2002). These allosteric changes have enabled the controlled release of highly insoluble and toxic chemotherapeutics in the treatment of certain cancers (Larroque et al. 1996).

The important role SA plays in the transport of critical biomolecules and interactions with pharmaceuticals makes it a protein of high interest to both academic and medical scientists. In addition to studies focused on understanding the binding properties

of SA to drugs, studies of SA and the albuminome (metabolites bound to SA) have increased as the protein has become linked to more diseases and disease progression (Fanali et al. 2012). With the ability to bind a wide range of metabolites, each with a different specificity, SA possesses attributes that make it ideal for the use in assays designed to differentiate samples and track disease progression.

Molecular Sensors

Molecular sensing, or supramolecular analytical chemistry, is a relatively new field defined loosely as the recognition of an analyte by a receptor, leading to formation of a complex which can be detected. When in complex, these sensors become fluorescent, typically in the visible light spectrum, thus allowing for detection of an analyte. Some of the earliest sensors were crown ethers which were used to detect metal ions with high specificity based on the size of the pore (Ghosh et al. 1987; Pond et al. 2004). More recently, molecular sensors have been synthesized to target small molecules for use in quantitative assays (Metzger and Anslyn 1998; Wiskur et al. 2001; Anslyn 2007; Umali and Anslyn 2010; Gallagher et al. 2012). Other applications of synthetic sensors include riboswitch control by light activation and viscosity sensors (Haidekker et al. 2006; Sutharsan et al. 2010; Wulffen et al. 2012). Biologists have also made use of proteins as sensors, owning to the fact that they have evolved over time to be bind selectively (Richieri et al. 1993; Braha et al. 1997; Halim et al. 2008).

Synthetic receptors have been around for decades, first used to detect metal ions using crown ethers (Rogers and Wolf 2002; Pond et al. 2004; Anslyn 2007). The

advantage of using synthetic receptors is that they can be tailored to bind an analyte of interest, with high affinity and specificity. Further, the wavelength of light which is emitted when in complex with an analyte can be tuned such that multiple receptors can be used simultaneously (Umali and Anslyn 2010; Gallagher et al. 2012). Additionally, synthetic receptors can be made for use in high throughput assays, where hundreds of samples can be rapidly processed and analyzed (Gallagher et al. 2012). The drawback to using synthetic receptors is that a basic understanding of what is in solution must be known, so that the receptor can be made to bind the analtye(s) of interest. This becomes a problem when looking at samples where the composition is unknown, leaving synthetic receptors to be used only on samples of known composition.

Alternatively, if the goal is to simply differentiate samples without prior knowledge of the composition, another type of sensor can be used. In this case, differences in the fluorescent signal can be observed and related to the change in sample composition. In this instance, synthetic receptors which are specific to an analyte are replaced with a receptor that is capable of binding a large range of molecules. One of the most common types of receptors used for this are proteins, specifically proteins like SA that have evolved to bind an array of molecules. Proteins can be specific to a particular analyte, class of analytes, or non-specific, allowing researchers to tune the specificity of their receptor, without having to do the hard work of synthetizing a specific receptor.

Previous reports showed that dye molecules which have been bound to SA can be used to differentiate a variety of samples (Adams and Anslyn 2009; Kubarych et al. 2010). When bound to the protein, a change in fluorescent signal was seen as the protein

was loaded with small molecules. This was attributed to the allosteric regulation of the protein, which induce conformational changes that altered the fluorescence of attached dye molecules (Adams and Anslyn 2009). In a separate experiment, the pockets of the SA protein were loaded with a reporter dye, which was quenched was bound in the pockets. However, upon mixing with a sample of small molecules which bound more tightly to the pockets, the dye was released and the fluorescence signal returned (Kubarych et al. 2010). In both experiments, SA was found to be highly successful in differentiating the samples from one another, with little to no prior knowledge of the molecular composition. However, unlike with synthetic receptors the actual molecules which were "sensed" by SA were unknown.

Given these problems, it would be ideal if there was an assay which could rapidly differentiate samples of unknown composition and also be capable of identifying the compounds responsible for the difference. An example of this would be a combination of the protein based sensor work already published, coupled to an analytical technique such as LC-MS which has the ability to rapidly identify 1000's of molecules.

Research Goals

The goal of my dissertation research was to develop an assay that is capable of both classifying samples and identifying the resulting metabolites responsible for the classification. Previous work showed that SA can be used as a sensor which binds a subset of molecules from samples and is capable of classifying them, but could not determine the exact molecular composition (Adams and Anslyn 2009; Kubarych et al.

2010). Although this work showcased the use of SA to classify samples, it was not of use in determining which particular molecules were used for classification. A way around this is to build receptors designed to bind a certain class of compounds, but this limits their use to samples of known composition (Metzger and Anslyn 1998; Anslyn 2007; Gallagher et al. 2012). An alternative is that the molecules bound to SA could be eluted and analyzed by LC-MS. While this alternative does not provide an immediate classification like a fluorescence based assay using synthetic receptors would, it would still enable classification and, more importantly, enable the identification of molecules which cause the classification.

Two aims were established in order to reach the goal of this research; i) develop a protocol for treatment of samples with SA for LC-MS analysis, and ii) test the protocol on a variety of samples to determine the limitations of the assay. A large number of variables become apparent when trying to develop an assay such as this. One such variable is the incubation times of sample with SA to facilitate binding and subsequently the removal of molecules, both of which must be optimized. In order to keep the assay rapid, these times must be as short as possible, while ensuring that enough material binds and is subsequently removed to allow classification. Another variable is the approach for denaturing and removing molecules from SA, with choices of chaotropic agent based or solvent based being the most convenient. Finally, optimization of the LC-MS method must be done, keeping the acquisition time to a minimum while still enabling sufficient separation of molecules. Each of the steps developed in the assay must be optimized to produce the desired result of a rapid assay that is capable of classifying samples.

Once the first aim has been realized, the second will be to test the assay on a variety of samples to determine how broadly applicable the assay will be to the scientific community. For instance, the assay may only work for one specific application, but is better than any other method currently implored. On the other hand, the assay may be useful for a wide variety of samples, but have different applications for them. It may be used as a simple classification tool for specific samples and used for biomarker discovery in others. By testing a wide range of samples, the specific uses will be found, or the broad applications probed for further research in the future. In addition to determining the applicability of the assay in a wide range of sample types, it will be important to determine if there are special protocols for different sample types, a pretreatment of the samples prior to being assayed. As an example, it is already known that SA is pH sensitive, so samples which are outside the working range of SA will need to have their pH adjusted before they can be assayed.

Although sensors are commonly used for a variety of applications, the use of a protein based sensor coupled to LC-MS is limited. The application of SA in conjunction with a fluorescence probe showed great promise as a molecular sensor, but little information was gained about the sample composition. Using LC-MS to detect the compounds which are bound to SA instead of relying on a fluorescence probe, information on the differences in molecular composition between samples can be obtained.

References

Adams MM, Anslyn E V (2009) Differential sensing using proteins: exploiting the cross-reactivity of serum albumin to pattern individual terpenes and terpenes in perfume. J Am Chem Soc 131:17068–17069. doi: 10.1021/ja908319m

Anslyn E V (2007) Supramolecular analytical chemistry. J Org Chem 72:687–99. doi: 10.1021/jo0617971

Bier M, Bakker BM, Westerhoff H V (2000) How yeast cells synchronize their glycolytic oscillations: a perturbation analytic treatment. Biophys J 78:1087–93. doi: 10.1016/S0006-3495(00)76667-7

Braha O, Walker B, Cheley S, et al (1997) Designed protein pores as components for biosensors. Chem Biol 4:497–505. doi: 10.1016/S1074-5521(97)90321-5

Bujacz A (2012) Structures of bovine, equine and leporine serum albumin. Acta Crystallogr Sect D Biol Crystallogr 68:1278–1289. doi: 10.1107/S0907444912027047

Cuperlovic-culf M, Culf AS, Morin PJ, Touaibia M (2013) Application of Metabolomics in Drug Discovery , Development and Theranostics. Curr Metabolomics 1:41–57.

Curry S, Brick P, Franks NP (1999) Fatty acid binding to human serum albumin: new insights from crystallographic studies. Biochim Biophys Acta 1441:131–40.

Curry S, Mandelkow H, Brick P, Franks N (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat Struct Biol 5:827–35. doi: 10.1038/1869

Curvale R (2009) Buffer capacity of Bovine serum albumin.(BSA). J Argent Chem Soc 97:174–180.

Davidovic L, Navratil V, Bonaccorso CM, et al (2011) A metabolomic and systems biology perspective on the brain of the Fragile X syndrome mouse model. Genome Res 21:2190–2202. doi: 10.1101/gr.116764.110

Dockal M, Carter DC, Rüker F (2000) Conformational Transitions of the Three Recombinant Domains of Human Serum Albumin Depending on pH. J Biol Chem 275:3042–3050. doi: 10.1074/jbc.275.5.3042

Fanali G, di Masi A, Trezza V, et al (2012) Human serum albumin: from bench to bedside. Mol Aspects Med 33:209–90. doi: 10.1016/j.mam.2011.12.002

Fielding L, Rutherford S, Fletcher D (2005) Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems. Magn Reson Chem 43:463–70. doi: 10.1002/mrc.1574

Fitos I, Simonyi M (1988) Selective effect of clonazepam and (S)-uxepam on the binding of warfarin enantiomers to human serum albumin. J Chromatogr 450:217–220.

Fitos I, Tegyey Z, Simonyi M, et al (1986) Stereoselective binding of 3-acetoxy-, and 3-hydroxy-1, 4-benzodiaze- pine-2-ones to human serum albumin. Selective allosteric interaction with warfarin enantiomers. Biochem Pharmacol 35:263–269.

Gallagher LT, Heo JS, Lopez MA, et al (2012) Pattern-based discrimination of organic acids and red wine varietals by arrays of synthetic receptors. Supramol Chem 24:2:37–41.

Ghosh S, Petrin M, Maki AH, Sousa LR (1987) Dependence of the triplet state properties on the orientation of metal ion perturbers in naphthalene-crown ether metal ion complexes. I. External heavy atom effect. J Chem Phys 87:4315–4323. doi: 10.1063/1.452890

Goodacre R (2005) Metabolomics – the way forward. Metabolomics 1:1–2. doi: 10.1007/s11306-005-1111-7

Gross JH (2011) Mass Spectrometry - A Textbook, 2nd edn. Springer

Haidekker M a, Brady TP, Lichlyter D, Theodorakis E a (2006) A ratiometric fluorescent viscosity sensor. J Am Chem Soc 128:398–9. doi: 10.1021/ja056370a

Halim AAA, Kadir HA, Tayyab S (2008) Bromophenol blue binding as a probe to study urea and guanidine hydrochloride denaturation of bovine serum albumin. J Biochem 144:33–8. doi: 10.1093/jb/mvn036

He XM, Carter DC (1992) Atomic Structure and Chemistry of Human Serum Albumin. Nature 358:209–215.

Heinonen M, Shen H, Zamboni N, Rousu J (2012) Metabolite identification and molecular fingerprint prediction through machine learning. Bioinformatics 28:2333–41. doi: 10.1093/bioinformatics/bts437

Ho JX, Holowachuk EW, Norton EJ, et al (1993) X-ray and Primary Structure of Horse Serum Albumin (Equus Caballus) at 0.27-nm resolution. Eur J Biochem 212:205–212.

Hollywood K, Brison DR, Goodacre R (2006) Metabolomics: Current technologies and future trends. Proteomics 6:4716–4723. doi: 10.1002/pmic.200600106

Hopkins FG, Pinkus SN (1898) Observations on the Crystallization of Animal Proteids. J Physiol 23:130–136. doi: 10.1016/S0008-6215(00)00094-X

Jansen JJ, Szymańska E, Hoefsloot HCJ, et al (2012) Between Metabolite Relationships: an essential aspect of metabolic change. Metabolomics 8:422–432. doi: 10.1007/s11306-011-0316-1

Johnson CH, Ivanisevic J, Benton HP, Siuzdak G (2015) Bioinformatics: The Next Frontier of Metabolomics. Anal Chem 87:147–156. doi: 10.1021/ac5040693

Jupin M, Michiels PJ, Girard FC, et al (2014) NMR metabolomics profiling of blood plasma mimics shows that medium- and long-chain fatty acids differently release metabolites from human serum albumin. J Magn Reson 239:34–43. doi: 10.1016/j.jmr.2013.11.019

Jupin M, Michiels PJ, Girard FC, et al (2013) NMR identification of endogenous metabolites interacting with fatted and non-fatted human serum albumin in blood plasma: Fatty acids influence the HSA-metabolite interaction. J Magn Reson 228:81–94. doi: 10.1016/j.jmr.2012.12.010

Karp PD, Ouzounis C a, Moore-Kochlacs C, et al (2005) Expansion of the BioCyc Collection of Pathway/Genome Databases to 160 Genomes. Nucleic Acids Res 33:6083–6089. doi: 10.1093/nar/gki892

Krishnakumar SS, Panda D (2002) Spatial relationship between the prodan site, Trp-214, and Cys-34 residues in human serum albumin and loss of structure through incremental unfolding. Biochemistry 41:7443–7452. doi: 10.1021/bi025699v

Kubarych CJ, Adams MM, Anslyn E V (2010) Serum albumins as differential receptors for the discrimination of fatty acids and oils. Org Lett 12:4780–4783. doi: 10.1021/ol101906g

Larroque C, Pelegrin A, Van Lier JE (1996) Serum albumin as a vehicle for zinc phthalocyanine: photodynamic activities in solid tumour models. Br J Cancer 74:1886–90.

Lin L, Huang Z, Gao Y, et al (2011) LC-MS based serum metabonomic analysis for renal cell carcinoma diagnosis, staging, and biomarker discovery. J Proteome Res 10:1396–405. doi: 10.1021/pr101161u

Majorek KA, Porebski PJ, Dayal A, et al (2012) Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. Mol Immunol 52:174–182. doi: 10.1016/j.molimm.2012.05.011

Metzger A, Anslyn E V. (1998) A Chemosensor for Citrate in Beverages. Angew Chemie Int Ed 37:649–652. doi: 10.1002/(SICI)1521-3773(19980316)37:5<649::AID-ANIE649>3.0.CO;2-H

Mitzner SR, Stange J, Klammt S, et al (2001) Extracorporeal detoxification using the molecular adsorbent recirculating system for critically ill patients with liver failure. J Am Soc Nephrol 12:S75–S82.

Motta A, Paris D, Melck D (2010) Monitoring real-time metabolism of living cells by fast two-dimensional NMR spectroscopy. Anal Chem 82:2405–11. doi: 10.1021/ac9026934

Nafisi S, Bagheri Sadeghi G, PanahYab A (2011) Interaction of aspirin and vitamin C with bovine serum albumin. J Photochem Photobiol B Biol 105:198–202. doi: 10.1016/j.jphotobiol.2011.09.002

Patti G, Yanes O, Shriver L (2012a) Metabolomics implicates altered sphingolipids in chronic pain of neuropathic origin. Nat Chem … 8:232–234. doi: 10.1038/NCHEMBIO.767

Patti GJ, Tautenhahn R, Siuzdak G (2012b) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 7:508–16. doi: 10.1038/nprot.2011.454

Peters Jr. T (1996) All About Albumin. Academic Press, San Diego, California

Petitpas I, Bhattacharya AA, Twine S, et al (2001a) Crystal structure analysis of warfarin binding to human serum albumin: anatomy of drug site I. J Biol Chem 276:22804–9. doi: 10.1074/jbc.M100575200

Petitpas I, Gru T, Bhattacharya AA, Curry S (2001b) Crystal Structures of Human Serum Albumin Complexed with Monounsaturated and Polyunsaturated Fatty Acids. J Mol Biol 314:955–960. doi: 10.1006/jmbi.2001.5208

Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1602–12.

Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11:395. doi: 10.1186/1471-2105-11-395

Pond SJK, Tsutsumi O, Rumi M, et al (2004) Metal-ion sensing fluorophores with large two-photon absorption cross sections: Aza-crown ether substituted donor-acceptor-donor distyrylbenzenes. J Am Chem Soc 126:9291–9306. doi: 10.1021/ja049013t

Richieri G V, Anel a, Kleinfeld a M (1993) Interactions of long-chain fatty acids and albumin: determination of free fatty acid levels using the fluorescent probe ADIFAB. Biochemistry 32:7574–80.

Rogers CW, Wolf MO (2002) Luminescent molecular sensors based on analyte coordination to transition-metal complexes. Coord Chem Rev 233-234:341–350. doi: 10.1016/S0010-8545(02)00023-1

Shao J, Shen H, Havsteen B (1993) Purification, characterization and binding interactions of the Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP. Biochem J 293:559–66.

Siuzdak G (2014) Mass spectrometry-based metabolomics as a unique biochemical approach for understanding disease pathogenesis. FASEB J 28:350.3.

Siuzdak G (2006) The Expanding Role of Mass Spectrometry in Biotechnology, 2nd edn. MCC Press, San Diego, California

Smith C a, O'Maille G, Want EJ, et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–751.

Smith CA, Want EJ, O'Maille G, et al (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Anal Chem 78:779–787.

Sutharsan J, Lichlyter D, Wright NE, et al (2010) Molecular rotors: Synthesis and evaluation as viscosity sensors. Tetrahedron 66:2582–2588. doi: 10.1016/j.tet.2010.01.093

Szymanski J, Jozefczuk S, Nikoloski Z, et al (2009) Stability of metabolic correlations under changing environmental conditions in Escherichia coli--a systems approach. PLoS One 4:e7441. doi: 10.1371/journal.pone.0007441

Tautenhahn R, Cho K, Uritboonthai W, et al (2012) An accelerated workflow for untargeted metabolomics using the METLIN database. Nat Biotechnol 30:826–828. doi: 10.1038/nbt.2348

Tautenhahn R, Patti GJ, Kalisiak E, et al (2011) metaXCMS: Second-Order Analysis of Untargeted Metabolomics Data. Anal Chem 83:696–700.

Thompson JJ (1913) Rays of Positive Electricity and Their Applications to Chemical Analysis. Longmans, Green and Co., London

Thomson JJ (1907) On Rays of Positive Electricity. Philos Mag 13:561–575. doi: 10.1080/14786440709463633

Thomson JJ (1899) On the Masses of the Ions in Gases at Low Pressures. Philos Mag 48:547–567. doi: 10.1080/14786449908621447

Umali AP, Anslyn E V (2010) A general approach to differential sensing using synthetic molecular receptors. Curr Opin Chem Biol 14:685–692. doi: 10.1016/j.cbpa.2010.07.022

Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, et al (2015) Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends. Curr Cardiovasc Risk Rep 9:12. doi: 10.1007/s12170-015-0440-y

Varshney A, Sen P, Ahmad E, et al (2010) Review Article Ligand Binding Strategies of Human Serum Albumin : How Can the Cargo be Utilized ? Chirality 22:77–87. doi: 10.1002/chir

Wiskur SL, Ait-Haddou H, Lavigne JJ, Anslyn E V (2001) Teaching old indicators new tricks. Acc Chem Res 34:963–72.

Worley B, Powers R (2013) Multivariate Analysis in Metabolomics. Curr Metabolomics 1:92–107.

Wulffen B, Buff MCR, Pofahl M, et al (2012) Caged glucosamine-6-phosphate for the light-control of riboswitch activity. Photochem Photobiol Sci 11:489–92. doi: 10.1039/c1pp05242j

Yamashita M, Fenn JB (1984) Electrospray ion source. Another variation on the free-jet theme. J PhysChem 88:4451–4459. doi: 10.1021/j150664a002

Zunszain P a, Ghuman J, Komatsu T, et al (2003) Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. BMC Struct Biol 3:1–9. doi: 10.1186/1472-6807-3-6

CHAPTER 2

METHODS AND DEVELOPMENT

General Method

The primary goal was to develop a general workflow for treatment of solutions to prepare for LC-MS analysis with minimal variation for different samples types. The method, herein referred to as the protein sensor assay (PSA), has three basic steps: binding, washing, and eluting. Sample preparation prior to the binding step is needed and can vary from simple pre-filtering to remove large debris, to more complex preparation requiring multiple steps. Analysis of the eluted compounds as described here was done using LC-MS. This is by no means the only viable technology for compound detection, but is the focus here because of the sensitivity and wide-range of molecule type that are observed. A diagram showing the steps in the PSA is shown, with each step described in the following sections (Figure 2.1).

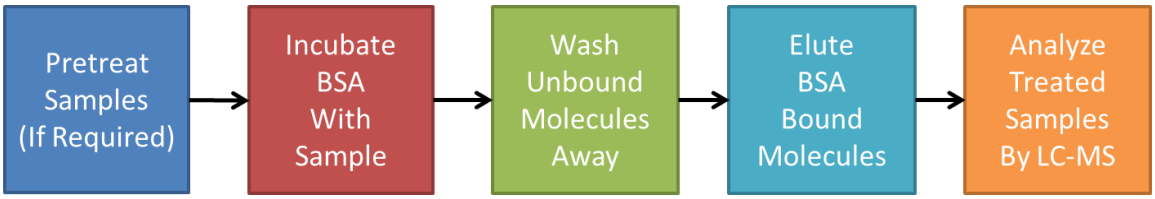

Figure 2.1. Block diagram showing the steps involved in the treatment of samples with the PSA prior to LC-MS analysis. The first step in the assay is pretreatment of samples such as adjusting the pH of wine or suspending dried metabolite extracts from cells. The second step of the assay is to incubate the sample with a solution of BSA. The third step is to remove unbound metabolites, while the forth step is to elute the metabolites bound to BSA. Finally, metabolites which were eluted from BSA are analyzed by LC-MS.

Cleaning of Spin Filters

Prior to use with the PSA, molecular weight cut-off (MWCO) spin filters were cleaned to remove preservatives and residuals of the manufacturing process. Water (500 μL) was placed onto the top portion of the spin filter and then centrifuged at 9,500 x g for 5 minutes to clean the membrane. A second wash was then carried out with the same conditions. Tests of the flow through after one and two washes showed that two wash steps was sufficient to remove background signal as assessed by LC-MS. Washing of filters was not done more than 6 hours prior to PSA analysis of samples, as longer periods of time lead to drying of the membrane, which compromised function as protein was observed to flow through.

Preparation of Protein

Protein for the PSA was prepared in batches and stored until use. Lyophilized BSA (Sigma-Aldrich, St. Louis, MO, item code A7030, 50g, lyophilized powder, pH 7, ≥98% pure as assessed by gel electrophoresis, essentially fatty acid free, essentially globulin free, protease free) was dissolved in a 10 mM HEPES buffer (pH 7.0) to a concentration of approximately 30 mg/mL and washed over 30 kDa MWCO spin filters. Filters were spun at 3000 x g for 10 minutes, washed with clean 10 mM HEPES buffer, pipetted to suspend protein bound to the filter membrane, and spun again. This was repeated twice more to remove non-specific and weakly bound molecules stemming from manufacturing and processing. Clean BSA was removed from the spin filters with fresh HEPES buffer and stored at a concentration greater than 30 mg/mL at -20 °C. For use in

the PSA, stock BSA was thawed on ice, then diluted with 10 mM HEPES buffer (pH 7.0, 250 mM NaOH) to the final working concentration of 0.1 mg/mL.

PSA Treatment of Samples

A three step procedure was developed for the PSA which involves binding of metabolites to BSA, washing away non-specific material, and eluting metabolites bound to the protein. Samples originating from a variety of different backgrounds have been tested with the PSA, and the specific details for preparing each are outlined later in this chapter. In general, samples must be dissolved in a liquid for PSA treatment, with the liquid containing a majority of water. Organic solvents may be used in low concentrations to suspend dried material, but must be kept to a minimum to avoid premature denaturation of the protein.

Liquid samples were mixed with BSA, generally in equal volumes, and allowed to equilibrate for 5 minutes. This step allows for the binding of metabolites to BSA. An incubation time of 5 minutes was found to be enough time to allow metabolites to bind to the protein and kept the overall time required for the assay to a minimum. After 5 minutes, the spin filters were centrifuged for 5 minutes at 9,500 x g to remove the liquid containing metabolites which did not bind to BSA. To ensure all non-specific material had been removed an additional washing step with 10 mM HEPES buffer (pH 7.0, 250 mM NaCl) was added, with the same conditions as above. In both cases, the flow through was discarded before moving on.

Elution of metabolites from BSA was carried out by denaturing the protein, which decreases protein-metabolite interactions. A solution of 70% aqueous methanol was

washed over the spin filter three separate times to obtain maximum elution of metabolites. The first wash of 200 µL was followed by a 5 minute spin at 9,500 x g. Two more washes of 100 µL each with a 5 minute spin at 9,500 x g were done and the resulting supernatant was dried using a speed-vac, and stored at -80 °C until analysis by LC-MS

Analysis of PSA Treated Samples

Samples treated using the PSA were suspended in a 50% aqueous methanol solution and analyzed using an Agilent 1290 UPLC system connected to an Agilent 6538 Q-TOF Mass Spectrometer. Separation of molecules was accomplished using a Kinetex 1.7 µm C18 150 mm x 2.1 mm column kept at 50 °C with a flow rate of 600 µL/min. A two solvent system was used to elute molecules from the column; solvent A consisted of 0.1% formic acid in water, while solvent B was 0.1% formic acid in acetonitrile. An elution gradient was as follows: 2% B for 2 minutes, increasing to 95% B over 10 minutes, held at 95% for 2 minutes, and then returned to 2% for 1 minute, with a total run time of 15 minutes. The mass spectrometer was operated in positive ion mode, with a cone voltage of 3500V and fragmentor voltage of 120V. Drying gas temperature was 350 °C with a flow of 8 L/min and the nebulizer was set to 60 psig. Spectra were collected at a rate of 2.52/s with a mass range of 50 to 1000 m/z, starting after the first minute to avoid excess salt building up on the source. Data acquired from the mass spectrometer was converted to mzML data format using the Agilent Data Reprocessor software.

Sample Preparation

Although the general methodology for the PSA was carried out for every sample, as outlined above, specific alterations of the method were needed depending on the actual sample to prepare it for PSA treatment. Preparation methods are outlined below for each different type of sample that was tested with the PSA. Methods are generally in order as they appear in following chapters.

Wine

Two problems were identified when treating wine with the PSA: i) large debris from the manufacturing process was left over and needed to be removed to avoid clogging spin filters, and ii) the pH of the wine was around 3.5-4.0 and had to be adjusted to 7.0 before treatment with the PSA. With respect to the first problem, the manufacturing process leaves behind relatively large debris, the cork can break down, and large proteins exist, all of which can clog the spin filter used in the PSA. To remove the larger, visible debris, wine samples were processed through a 0.22 μm syringe filter. For protein removal, the filtered wine was washed over a 10 kDa MWCO spin filter and spun at 9,500 x g. The supernatant was removed from the filter, and the wine was stored until PSA treatment.

Prior to PSA treatment, a second pre-processing step to adjust the pH to 7.0 was done. This step was necessary as SA proteins are sensitive to pH, having different confirmations based on the pH environment it is in (Dockal et al. 2000). To account for this, the pH was adjusted to 7.0 using 1 M NaOH. In doing so, the color of the wine

changed from red to green. To streamline the process in future experiments and alleviate the need for a pH probe, an experiment was done with a variety of wines to identify a volume of NaOH which would bring the pH of these wines as close to 7.0 as possible. A volume of 10 μL of 1 M NaOH was found to bring the wines within 0.1 pH unit of 7.0, and used in all future experiments.

Whiskey

As with the wine samples, large debris left over from the manufacturing process needed to be removed from the whiskey samples prior to PSA treatment. Additionally, the alcohol content in whiskey ranges from 60% when first distilled, to 40% in the final product, which is not compatible with the SA. The high alcohol content causes the protein used in the PSA to partially denature and can remove weakly bound molecules prematurely during the assay, or prevent binding altogether. Large debris was first removed by centrifugation, 20,000 x g for 15 minutes, followed by filtration through a 10 kDa MWCO spin filter. The filtered supernatant (100 μL) was then diluted with water, to lower the alcohol content to less than 15% (3:1 water to whiskey). The final alcohol content of the diluted whiskey was at the same level as with the wine samples, meaning it would not interfere with the protein in the PSA. Furthermore, the final volume of the diluted whiskey (400 μL) when mixed with protein for the PSA (100 μL) could all be washed over the MWCO spin filters, meaning there was not loss of metabolites, only a dilution.

Metabolite Extracts and Biological Fluids

Biological samples (e.g., extracted from cells or fluids like urine) required minimal preparation for treatment with the PSA. Metabolites were extracted from cells using a water and methanol based method and stored dry at -80 °C. These samples were re-suspended in 50% aqueous methanol prior to PSA treatment. After samples were mixed with SA protein the resulted solution contained 25% methanol; this concentration did not impact SA protein function. Samples prepared from a biological fluid like urine were stored as frozen solution, therefore they required only filtration through a 30 kDa MWCO spin filter to remove residual protein that would interfere with PSA analysis by clogging the 3 kDa MWCO spin filters.

Blood Plasma

Plasma samples are rich in metabolites but also contain high concentrations of protein, the most abundant of which is SA protein at roughly 50 mg/mL. Given that plasma is primarily dominated by SA, two different analyses were done from plasma. The first analysis was to strip metabolites bound to the SA from the animal and then bind them to laboratory prepared BSA, as in the traditional PSA. The second analysis stripped the metabolites bound to the endogenous protein and measure them directly by LC-MS. The first analysis required that the metabolites be removed from proteins naturally found in the plasma before being mixed with laboratory BSA in the normal PSA. This was accomplished by mixing 100 μL of plasma with 800 μL of cold methanol, vortex mixing the sample and incubating the sample at -80 °C for 30 minutes. Following this, the sample was spun at 20,000 x g for 5 minutes and the resulting supernatant was dried with

a speed-vac and stored at -80 °C. To use the PSA on dried metabolites from plasma, the sample was re-suspended in 100 µL of 50% aqueous methanol and treated like a cell extract as described above. Endogenous PSA analysis was done by taking 100 µL of plasma, mixing it with 800 µL of 10 mM HEPES buffer (pH 7.0, 250 mM NaOH), vortexing. From the resulting diluted solution, 10 µL was washed over a 3 kDa MWCO spin filter and taken through the PSA as normal starting at the washing step.

Method Development and Optimization

Progression of the PSA from initial conception to the current iteration involved a significant amount of optimization from sample preparation through data acquisition. Through this progress, each step of the PSA was optimized for efficiency while retaining a high level of quality and reproducibility between samples. A number of parameters were altered in order to obtain the goal of a fast and reproducible assay including: protein conditions, method for denaturing the protein, spin filter size and spin speed, and LC column and run condition. Each step was evaluated individually through extensive testing, and the process for determining the best parameters is described below.

The first iteration of the PSA was initially developed to get a baseline method which could be altered moving forward, and was by in large successful, but could be improved upon. The first trial of the PSA used 30 mg/mL of BSA in a 10 mM HEPES buffer (pH 7.0, no salt), washed over 3 kDa MWCO spin filters and centrifuged at roughly 14,000 x g. The protein was denatured using a 6M guanidine HCl solution. Two major problems presented themselves almost immediately: i) the protein was too

concentrated for the spin filters and ended up clogging them, and ii) the filters were being spun too fast, causing them to break and leak protein through. This resulted in either no metabolites being detected by the mass spectrometer when the filters became clogged, or if the filters broke than large quantities of protein were found, which causes permanent damage to the column. To address these problems, a series of experiments which altered the protein concentration and the spin speed were done. Careful analysis showed that a protein concentration of less than 1 mg/mL were best, with a concentration of 0.1 mg/mL giving the best tradeoff for signal of metabolites with shorter spin times. For protein concentrations of 0.1 mg/mL, speeds of 9,500 x g for centrifugation resulted in no filters being broken and allowed for quick 5 minute spins to be done. Slower speeds required longer spin times, with no gain in the signal of metabolites.

Another set of experiments aimed at reducing the frequency of spin filter clogging was simultaneously done, focusing on different filters sizes. The original method developed for the PSA used 3 kDa MWCO spin filters. MWCO filters with larger cutoff sizes are also available from the same company and were explored. The protein used in the PSA, bovine serum albumin, has a mass of roughly 66 kDa, with other orthologs having similar weight. Filter sizes of 10 kDa and 30 kDa were tested with varying concentrations of BSA to ascertain if larger filters could be used with higher protein concentrations to increase the signal of metabolites selected with the PSA. Not surprising, using larger cutoff filters had no effect on the amount protein which could be loaded onto them. Just as before, protein concentrations greater than 1 mg/mL clogged the filters. Additionally, the 30 kDa MWCO spin filters allowed protein through the filters, while

the 10 kDa allowed small proteins, and larger peptides through the filter, both of which can damage the LC column used to separate metabolites. Given these results, 3 kDa MWCO spin filters were chosen for all further experiments, with a low concentration of protein to be used (0.1 mg/mL for traditional PSA experiments, and less than 1 mg/mL for plasma based endogenous analysis).

Protein denaturation is a common practice in biochemistry, with applications for removing proteins from samples or facilitating analysis by gel electrophoresis. A variety of common methods used in the field of protein work were explored to denature the protein using both chaotropic agents and solvent based systems. The first denaturants used were chaotropic agents such as guanidine-HCl and urea at high concentrations. The problem with the chaotropic agents was that the high concentrations needed to denature SA protein had downstream repercussions during LC-MS analysis. This was evident with guanidine-HCl, where high concentration required relatively large volumes of solvent (100 μL) to re-suspend the metabolites which were eluted from the protein, causing low signal of metabolites in the mass spectrometer. Additionally, injection of large quantities of chaotropic agents on to the LC column used for separation caused large retention time shifts of metabolites and dirtied the mass spectrometer source, requiring it to be cleaned more often.

An alternative method of protein denaturation which is commonly used in biochemical labs is done using organic solvents which are mass spectrometer compatible and can be removed with simple drying. Several organic solvents were tested to determine which would work best for denaturing SA protein including alcohols

(methanol and isopropanol) and acetonitrile at various concentrations in water. The maximum concentrations of each were chosen based on the spin filters which could only tolerate up to 70% alcohol and 20% acetonitrile. In general, higher concentrations of alcohol or acetonitrile gave better results as assed by increased signals in the mass spectrometer, suggesting more denaturation took place. Comparison of PSA samples which had been denatured with 70% methanol, 70% isopropanol, and 20% acetonitrile showed similar peak profiles and signal intensities, although methanol produced slightly more intense peaks followed by isopropanol and finally acetonitrile (Figure 2.2). These results indicated that there was little difference between one solvent over another when they are used at the maximum concentration tolerated by the spin filters, likely because the protein is either fully denatured in each solvent at those concentrations, or that they all denature the protein to the same extent. A literature search revealed that at the concentrations used in this experiment, the alcohols have the same effect on protein denaturation (Herskovits et al. 1970). Given these results, methanol was chosen as the ideal solvent for denaturation as it was also the solvent used to re-suspend PSA treated samples for LC-MS analysis, decreasing the number of different solvents used throughout the assay. Comparison of PSA treated samples where protein was denatured by 6M guanidine-HCl and 70% methanol showed similar peak profiles but the intensity was drastically reduced in the guanidine-HCl samples (Figure 2.3). This was due to the large volume of solvent required to re-suspend the sample compared to the methanol treated sample (100 μL for guanidine-HCl and 20 μL for 70% methanol).

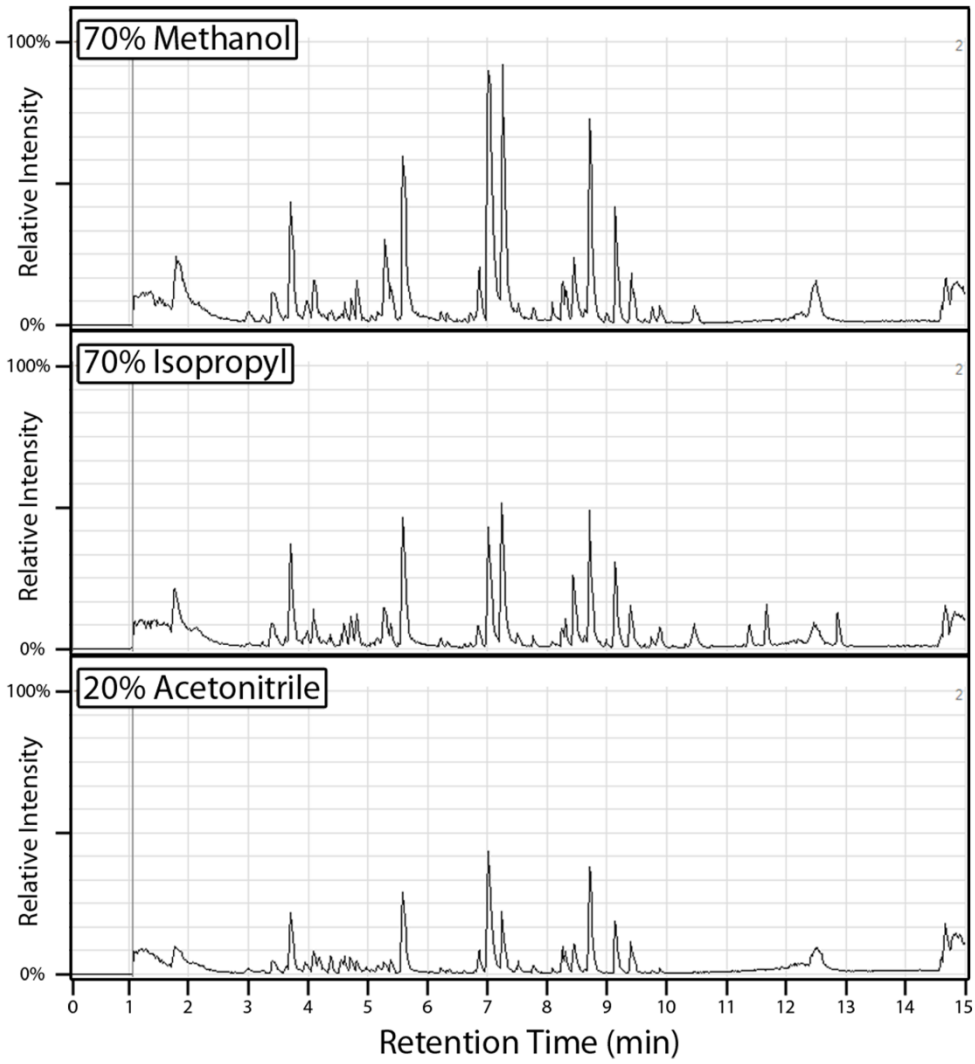

Figure 2.2. Comparison of base peak chromatograms (BPCs) for three solvents used to extract metabolites from SA. Metabolites extracted from SA using three different solvents to denature the protein and remove the metabolites were compared for their ability to extract the maximum number of metabolites with the greatest signal as assessed by LC-MS. BPCs for each solvent at the maximum concentration that is compatible with the spin filters used in the assay are shown. Each solvent produced roughly the same trace, showing that the extracts work similar. However, there is a difference in the intensity of the peaks when comparing the solvents, where 70% methanol produces the most intense peaks, followed by isopropanol, and finally acetonitrile. Due to this, methanol was chosen as the solvent for future work.

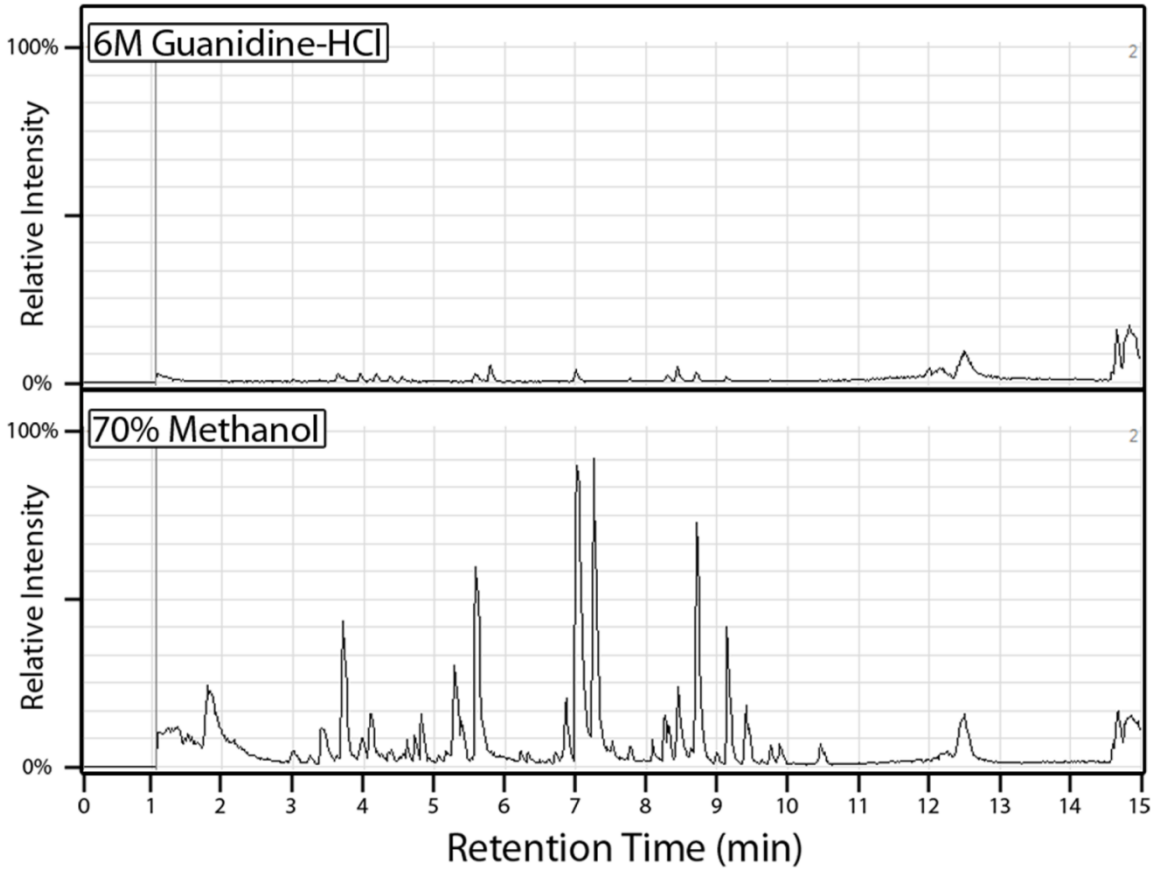

Figure 2.3. Comparison of BPCs for guanidine-HCl and 70% methanol used to extract metabolites from SA. Metabolites extracted from SA using chaotropic agents and solvents to denature the protein and remove the metabolites were compared for their ability to extract the maximum number of metabolites with the greatest signal as assessed by LC-MS. Comparison of the BPC for one chaotropic agent (guanidine-HCl) and one solvent (70% methanol) shows that the solvent based method produces a more intense spectrum than that of the chaotropic agent. In all cases, similar peak profiles were produces, but the solvent based methods produced more intense spectra when compared to chaotropic agents based methods.

A key advantage of the PSA over a traditional LC-MS experiment is that the time required for data acquisition is significantly reduced. However, this reduction in time required optimization of the chromatography parameters. A traditional LC-MS experiment typically requires 30 minutes of chromatography; a balance between having enough time to separate the metabolites from one another, while not being so long that

fluctuations in the instrument from day to day would become a problem due to data acquisition taking a week or more. The length of column plays a significant role in determining acquisition time, where a shorter column can have shorter run times but at the same time sacrifices separation efficiency. Conversely, a longer column can provide better separation, but requires longer run time to analyze one sample. A reversed phase (RP) column was chosen because the molecules expected to bind to SA are semi-polar in nature. This result was confirmed when investigations of untreated samples and samples treated with the assay revealed a shift in elution profile from polar molecules in untreated to semi polar after treatment (Hamerly and Bothner 2015). An additional reason for choosing a RP column is that they have excellent compound resolving power (separation). This is due to the high pressures they can withstand which allows for smaller particles to be used inside the columns, increasing the number of interactions a compound has with the stationary phase.

To find the best RP column suited for the PSA, samples were analyzed using one 15 cm column (Phenomenex C18), one 10 cm column (Agilent C18), and one 5 cm column (Agilent C18), all C18 (RP) stationary phases. All three columns used a 2-95 %B gradient, where A was water and B was acetonitrile, both with 0.1% formic acid. The run times for the 10 and 15 cm columns were 15 minutes, while the 5 cm column had a run time of 5 minutes. The 5 cm column was not long enough to separate the metabolites well enough for analysis from PSA treated samples and was not tested further. The difference between the 15 cm and 10 cm columns (aside from length) was subtle, but important in their performance. Both columns were C18 RP columns, but the type of material used in

the columns was different, resulting in the Phenomenex column giving better separation than the Agilent column (Figure 2.4). The type of particles in the stationary phase beads plays an important role in separating closely related compounds in chromatography, and the difference between the two columns tested was that the Phenomenex column uses a core-shell for its particles, while the Agilent column uses fully porous particles. The advantage of using the core-shell particles is that the resulting particles are more homogenous in size compared to traditionally fully porous particles, resulting in better separation of metabolites and less peak broadening (Gritti and Guiochon 2012).

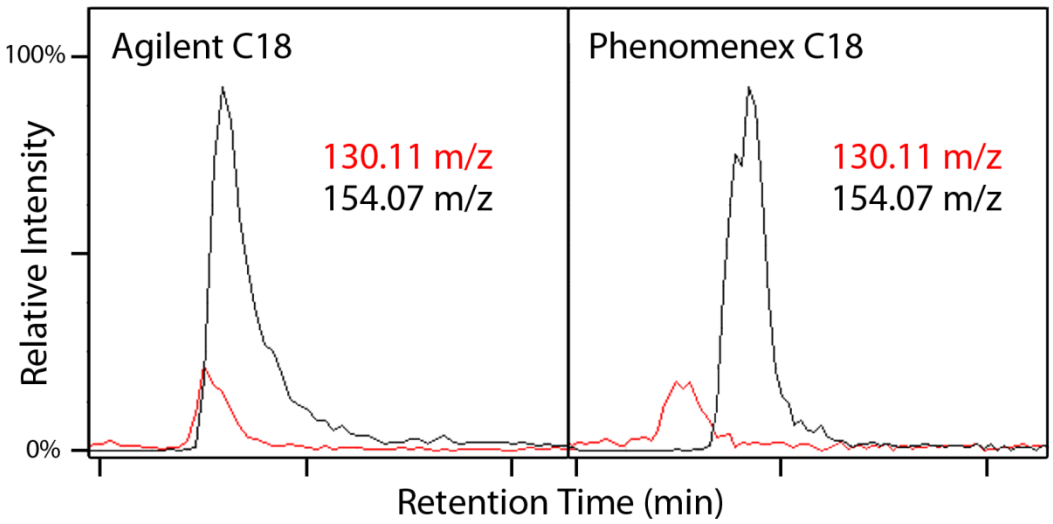

Figure 2.4. Comparison of Agilent C18 and Phenomenex C18 columns for the separation of metabolites extracted using the PSA. Samples treated using the PSA were than analyzed by LC-MS using two columns that varied by column length (10 cm for the Agilent C18 and 15 cm for the Phenomenex C18). Extracted ion chromatograms (EICs) for two features show that Agilent C18 column cannot resolve the peaks and are overlapping, while the Phenomenex C18 separates the two features with baseline resolution. The separation is a result of the longer column length and shows that with the same solvent system and elution gradient, the longer column produces better separation of similar polarity molecules.

References

Dockal M, Carter DC, Rüker F (2000) Conformational Transitions of the Three Recombinant Domains of Human Serum Albumin Depending on pH. J Biol Chem 275:3042–3050. doi: 10.1074/jbc.275.5.3042

Gritti F, Guiochon G (2012) Facts and Legends on Columns Packed with Sub-3-µm Core-Shell Particles. LCGC North Am 30:586–595.

Hamerly T, Bothner B (2015) Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis. Appl Biochem Biotechnol 1–13.

Herskovits TT, Gadegbeku B, Jaillet H (1970) On the Structural Stability Denaturation of Proteins. J Biol Chem 245:2588–2598.

CHAPTER 3

BOVINE SERUM ALBUMIN AS A MOLECULAR SENSOR FOR THE DISCRIMMINATION OF COMPLEX METABOLITE SAMPLES

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Timothy Hamerly

Contributions: Prepared wine samples, treated samples with PSA, acquired LC-MS data on PSA treated samples and untreated wine samples, analyzed data, generated figures, wrote manuscript.

Co-Author: Joshua Heinemann

Contributions: Provided help for LC-MS acquisition, valuable insight.

Co-Author: Monika Tokmina-Lukaszewska

Contributions: Pre-processed urine samples, acquired LC-MS data on untreated urine samples, insight into the model for shock.

Co-Author: Elizabeth R. Lusczek

Contributions: Gathered urine samples for shock model.

Co-Author: Kristine E. Mulier

Contributions: Gathered urine samples for shock model.

Co-Author: Greg J. Beilman

Contributions: Oversight on shock model, sample collection, insight on data, edited manuscript.

Contribution of Authors and Co-authors Continued

Co-Author: Brian Bothner

Contributions: Pivotal in assay development, provided insight and interpretation of results, edited manuscript.

Manuscript Information Page

Authors: Timothy Hamerly, Joshua Heinemann, Monika Tokmina-Lukaszewska, Elizabeth R. Lusczek, Kristine E. Mulier, Greg J. Beilman, and Brian Bothner

Journal: Analytica Chemica Acta

Status of Manuscript:
___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
__X_ Published in a peer-reviewed journal

Published by El Sevier

Issue in which manuscript appears: Volume 818, pages 61–66, 2014

BOVINE SERUM ALBUMIN AS A MOLECULAR SENSOR FOR THE DISCRIMMINATION OF COMPLEX METABOLITE SAMPLES

Timothy Hamerly,† Joshua Heinemann,† Monika Tokmina-Lukaszewska,† Elizabeth R. Lusczek,‡ Kristine E. Mulier,‡ Greg J. Beilman,‡ and Brian Bothner*†

† Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717
‡ Department of Surgery, University of Minnesota, Minneapolis, MN 55433

*Running title: Protein sensor discriminates complex metabolite samples

To whom correspondence should be addressed: Brian Bothner, Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717, Phone: 406-994-5270, FAX: 406-994-4807, E-mail: bbothner@chemistry.montana.edu

Abstract

The potential for using serum albumin (SA) as a broadly applicable molecular sensor was explored in an effort to develop a method for rapid analysis of complex metabolite samples. SA is a protein present at high concentration in blood, which transports a diverse set of compounds including fatty acids, hormones, and drugs. The effectiveness of the bovine ortholog (BSA) as a molecular sensor was tested by analyzing the pool of small molecules bound to the protein after a brief incubation with complex fluids of biological origin. As an initial test, three varietals of red wine were readily distinguished. Further analysis using four varietals of white wine also showed clear separation. In a second analysis using urine, animals in hemorrhagic shock were separated from a group of comparably treated controls. A time course analysis showed that recovery from injury could also be followed using the assay. This finding is significant as there currently is no method or biomarker for predicting the onset of shock.

Comparison of samples was based on Liquid Chromatography Mass Spectrometry (LCMS) analysis of compounds selectively bound by BSA. Analysis of the samples after protein selection revealed a significant reduction in complexity and clear separation of groups by Principle Component Analysis (PCA). These results show the potential for using cargo-carrying proteins as molecular sensors for screening complex samples without the need for prior knowledge of sample composition or concentration and may streamline elucidation of biomarkers.

Introduction

The ability to quickly and reliably differentiate complex samples is a continuing challenge for both analytical chemists and biologists. This problem has been addressed using both chemical (synthetic ligands and instrumentation) and biological (antibody and aptamer) approaches (Anslyn 2007; Adams and Anslyn 2009; Kubarych et al. 2010; Umali and Anslyn 2010; Gallagher et al. 2012). These methods generally require prior knowledge of the molecule(s) of interest and rely on a single or at most a few specific molecules for sample differentiation. However, differentiation between disease and healthy states in a human, or pinpointing the origin of an agricultural product are more complex. In such cases, the relative abundance of a group of components may be important for unambiguous differentiation. The protein serum albumin (SA) is known to interact with a wide-range of small molecules with varying affinity and has shown potential for use as a sensor for characterizing complex samples (Adams and Anslyn 2009; Kubarych et al. 2010). Serum albumin is a highly abundant protein in the blood of

all mammals (Peters Jr. 1996). It has been relatively well characterized owing to its role in the transport of fatty acids, hormones, and drugs through the blood stream (Spector et al. 1969; Sjoholm et al. 1979; Varshney et al. 2010).

Efforts to understand the basis for small molecule binding by SA have revealed a set of sites for cargo loading (Adams and Anslyn 2009; Varshney et al. 2010; Kubarych et al. 2010; Nafisi et al. 2011). The ability to bind a diverse array of small molecules using specific binding sites is a common feature of all SA proteins. It has also been shown that binding at one site can have allosteric effects on the other sites (Varshney et al. 2010). These properties have led to the use of SAs in differential sensing of small organic molecules in complex mixtures such as terpenes in perfume and fatty acids in oils (Bhattacharya et al. 2009; Adams and Anslyn 2009; Kubarych et al. 2010). These initial studies were conducted using a fluorescent dye molecule which is noncovalently bound to SA, and upon the addition of analyte, is released into solution resulting in a change in fluorescence. While clearly powerful, a limitation of this approach is that it often requires replicate analyses with different fluorescent reporters and there is no way to determine the molecule(s) responsible for different signals using the assay. The mechanism of fluorescence signal modulation was believed to be allosteric changes in the SA caused by small molecule interaction events (Adams and Anslyn 2009). Based on this work, we hypothesized that the allosteric effects of the small molecules on the protein were independent of the indicator dye; therefore SA alone should be able to discriminate mixtures of small molecules by integrating binding and allosteric properties. To test this hypothesis, we designed a set of experiments to see if bovine serum albumin (BSA) could

discriminate complex samples of biological origin. Here in we report the use of BSA as a molecular sensor which can reliably differentiate wine varietals as well as identify animals entering hemorrhagic shock from urine samples.

The production and properties of wine have long been known. The fermentation process was likely discovered by the ancient Egyptians, followed soon after by awareness of the antimicrobial properties due to the ethanol content (Soleas et al. 1997; Mcgovern 2000). The Greeks knew of its health benefits and it became a staple in their diets (Mcgovern 2000; Bisson et al. 2002). Given the long history and culture of wine production, analysis and comparison of different varietals is a relatively recent development (Umali and Anslyn 2010). Anthrocyanins, phenols, and resveratrol have been heavily studied because of interest in their antioxidant properties. Relative abundance of these compounds has been investigated as markers to help discriminate varietals, albeit with marginal success (Villiers et al. 2004; Martí et al. 2004; Cuadros-Inostroza et al. 2010; Jaitz et al. 2010). Further advancements in the differentiation of wines using headspace mass spectrometry (HS-MS) and synthetic receptors have been made over the past decade (Martí et al. 2004). These techniques, although promising, have yet to see wide spread application, largely due to practical limitations. For example, HS-MS runs are dominated by common volatile compounds such as ethanol, making it hard to detect less abundant molecules of higher diagnostic value (Martí et al. 2004). Synthetic receptors are versatile, but due to the need for different receptors for each class of compounds, they can be cost prohibitive for rapid and high-throughput differentiation of complex mixtures, such as wine (Umali and Anslyn 2010; Gallagher et al. 2012).

Increased interest in a rapid method to verify valuable and limited varietals is a further driving force for using wine as a test case.

Hemorrhagic shock, a result of acute blood loss following injury, is the third leading cause of death in the United States and the leading cause of death for people under 40 (Kauvar and Wade 2005). It is a priming event for the development of Multiple Organ Dysfunction Syndrome and the outcomes of hemorrhagic shock include increased mortality, length of hospital stay, and cost of treatment (Boulanger et al. 2007). Currently there is no clinical test that can be used to screen patients for progression to shock. Therefore, as a second test of the BSA protein sensor assay, urine samples from a large animal model for hemorrhagic shock were used. Hemorrhagic shock is defined as a condition produced by the rapid and severe loss of intravascular volume, resulting in decreased oxygen delivery to tissues, cell hypoxia, organ damage and often death (Taylor et al. 2004; Gutierrez et al. 2004; Scribner et al. 2010; R. Lusczek et al. 2011; Cherkas 2011). Shock is most often a result of trauma and quickly becomes irreversible if left untreated (Gutierrez et al. 2004; R. Lusczek et al. 2011). The cause of death from shock is a result of cell hypoxia, in which the cell can no longer make ATP via oxidative phosphorylation and cannot keep up with the energy demands (Taylor et al. 2004). Efforts over the past decade aimed at finding a biomarker or test which could predict onset of shock in a clinical setting have yet to yield such results (Skarda et al. 2007; Kumar et al. 2011). Analysis of serum, muscle, and urine samples following a time course after injury have added greatly to our understanding of the physiology of shock (Scribner et al. 2010; R. Lusczek et al. 2011; Kumar et al. 2011). However, to our

knowledge, a straight-forward and robust assay to predict hemorrhagic shock does not exist.

Experimental

Materials

Three varietals of red wine were purchase off the shelf: Cabernet Sauvignon (vintage 2003), Merlot (vintage 2003), and Pinot Noir vintage 2008), all produced by Barefoot Cellars (Modesto, CA). White wine samples were collected at a wine tasting event, being collect immediately after uncorking in clean 15 mL tubes filled so that no head space remained. Used in the study were two productions of Chardonnay (Januik Winery, Woodinville, WA and Abeja Winery, Walla Walla, WA), one Viogner/Chardonnay blend (Qupe Wines, Santa Barbara, CAa) and one Grenache/Roussane/Viogner blend (Nostre Pais, Languedoc Roussillon, France) and two types of Sauvignon Blanc (Dog Point Vineyard, Marlborough, New Zealand and Whitehall Lane Winery, St. Helena, CA). Bovine serum albumin (greater than 98% agarose gel electrophoresis pure) was purchased from Sigma-Aldrich (St. Louis, MO). Mmolecular weight cutoff spin filters and syringe filters were purchased from Pall (Port Washington, NY). The buffer agent 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) and Sodium Chloride were purchased from VWR (Radnor, PA) at a purity of 99% or greater. All solvents were purchased in HPLC grade; water from Avantor (Center Valley, PA) and methanol and acetonitrile from EMD Chemicals Inc.

(Gibbstown, NJ). Formic acid (98% GR ACS) for use as an ion pairing agent was purchased from EMD Chemicals Inc. (Gibbstown, NJ).

Preparation of BSA for Assay

BSA was extensively washed to remove any molecules that were left bound to the protein after manufacturers processing. A solution of BSA in 10 mM HEPES buffer at a pH of 7.0 was washed over 3kD molecular weight cutoff (MWCO) spin filters and rinsed further with 10 mM HEPES buffer. The BSA was then re-suspended in 250 mM NaCl, 10 mM HEPES buffer at a concentration of 30 mg mL^{-1} and frozen until analysis. All spin filters were washed thoroughly prior to use with water to remove any contaminants or preservatives.

Treatment of Wine Samples with BSA Assay

Red wine samples were collected 30 minutes and 24 hours after opening, clarified with a 0.22 μm syringe filter, and frozen. Before analysis, samples were further clarified with a 10 kD MWCO spin filter and the pH of each was adjusted to 7.0 using 1M NaOH, before being mixing with 0.1 mg of BSA and allowed to interact for five minutes. The samples were then washed over a 3kD MWCO spin filter and were centrifuged for 5 minutes. The samples were then rinsed with a 10 mM HEPES solution to further remove nonspecific material and centrifuged again for 5 minutes. The BSA was then washed and spun three separate times with 70% methanol to elute retained compounds and the supernant was collected after each spin. Samples were evaporated to dryness in a speed-vac and kept at -80 °C until liquid chromatography mass spectrometry (LCMS) analysis.

Additionally, experimental blanks were used at each step in the method as controls for protein, solvent, and filter contamination. The graphical abstract shows a schematic of the method.

White wine samples were all collected at the same time after opening, then treated with the BSA assay and analyzed exactly the same as the reds.

Treatment of Urine Samples with BSA Assay

Urine samples collected from pigs during a controlled experiment for hemorrhagic shock were prepared in a similar manner as described for wine samples.(Skarda et al. 2007) Urine samples from animals which had undergone shock, and a control set which did not, were mixed with BSA and washed over 3kD MWCO spin filters. The samples were washed with 10 mM HEPES solution, and the metabolites were eluted from BSA with 70% methanol. Samples were then dried using a speed-vac and stored at -80 °C until LCMS analysis.

Mass Spectrometry of Samples

Prior to LCMS analysis, samples were re-suspended in methanol:water (50:50). A Kinetex 1.7 μm C18 150 mm x 2.1 mm column (Phenomenex, Torrance, CA) kept at 45 °C was used for LC separation with a flow rate of 500 μL min^{-1}. Solvent A consisted of 0.1% formic acid in water, while solvent B was 0.1% formic acid in acetonitrile. The elution gradient consisted of 2% B for 2 minutes (with the first minute going to waste to avoid contaminating the source with excess salt), to 95% B over 10 minutes, held at 95% for 2 minutes, and then returned to 2% for 1 minute, with a total run time of 15 minutes

using an Agilent 1290 UPLC (Agilent, Santa Clara, CA) system connected to an Agilent 6538 Q-TOF Mass Spectrometer (Agilent, Santa Clara, CA). Mass spectrometry analysis was conducted in positive ion mode, with a cone voltage of 3500V and fragmentor voltage of 120V. Drying gas temperature was 350 °C with a flow of 12 L min^{-1} and the nebulizer was set to 60 psig. Spectra were collected at a rate of 2.52 per second with a mass range of 50 to 1000 m/z.

Data Analysis of LC-MS

Molecular features from each run were extracted using Masshunter Software (Agilent; Santa Clara, CA) with export to XCMS for analysis (Smith et al. 2006; Tautenhahn et al. 2008; R Core Team 2012). Samples were first grouped with the default settings applied, followed by retention time correction with the family set to symmetric. A second grouping was performed on the retention time corrected data, with a bandwidth of 10. A difference report was then generated for each set of samples and the data was piped directly into metaXCMS for analysis (Tautenhahn et al. 2011). metaXCMS takes the output of XCMS in the form of a comma separated value file and attempts to align compound results from multiply XCMS outputs. The end result is a new file, which contains aligned compounds from all XCMS outputs. The data output from metaXCMS (molecular features in the form of m/z and their associated intensities) was analyzed using principle component analysis (PCA) with the R package PCA methods and visualized with the R package rgl (Adler and Murdoch 2012). PCA involves an orthogonal data transformation which maps features to a new coordinate system such that the greatest variance between data points can be observed, and is often used in the

analysis of LCMS data sets (Jolliffe 2002; Kim et al. 2009). Further analysis using agglomerative hierarchical clustering was used to group the white wine varietals and calculate the dissimilarity index using XLSTAT (Addinsoft 2013).

Results and Discussion

Analysis of Red and White Wines Treated with BSA Assay Using PCA and Hierarchical Clustering

The protein sensor-based assay developed here involves binding, washing and elution of small molecules to BSA, followed by LCMS analysis (Figure 3.1).

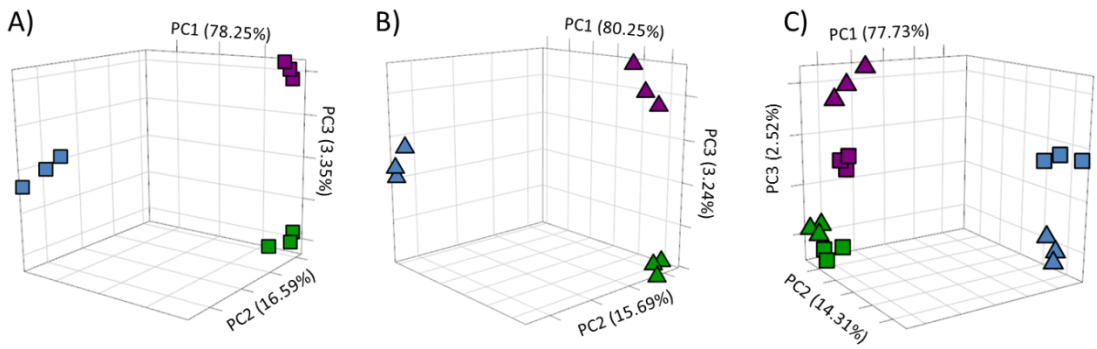

Figure 3.1. Differentiation of BSA treated red wine varietals. PCA plots showing the separation of red wines based on varietal at 30 minutes (A) and 24 hours (B) after opening. Plot C shows the separation of varietals at both 30 minutes (squares) and 24 hours (triangles) together. In all plots; cabernet (green), merlot (purple), pinot noir (blue).

As a proof of principle, three red wine varietals were analyzed with the protein sensor assay. Analysis of the small molecules bound to BSA using LCMS revealed a clear differentiation of varietals. Principle Component Analysis (PCA) on the three wines (cabernet sauvignon, merlot, and pinot noir) 30 minutes after opening are shown in Figure 3.1A. The contribution to separation for each component is shown on the axes,

with the first principle component (PC1) showing the greatest separation. The varietals were well separated, whereas technical replicates were closely clustered. Figure 3.1B shows that the wines could also be differentiated 24 hours after opening. To test if the method was robust enough to maintain discrimination of the varietals while including a temporal variable, data from all of the samples were combined; Figure 3.1C shows the PCA plot of the three varietals, at 30 minutes and 24 hours after opening. Not only did the varietals remain separated, but the individual time points within each varietal were also separated from one another, but to a lesser degree than the varietals themselves. This indicated that not only do varietals show unique characteristics, but also that the process of letting a wine breath imparts measurable and consistent change. These data clearly show that the method is reproducible and has the sensitivity to distinguish red wine based on small molecule profiles.

To further test the ability of our assay to differentiate wine, a more challenging set of white wines were analyzed. In this case, the same varietal from different regions and blended wines were examined. Using the same method as for the red wines, a 3D PCA plot of the white wines showed good separation for the complete set (Figure 3.2). The closest grouping was among technical replicates as expected. Next nearest neighbors were between the same varietal from different wineries. The two different Chardonnays and Sauvignon Blancs were best separated by principle component one (PC1), which accounts for 76.26% of the variance between samples. An interesting point is that the two blended wines sit at the utmost reaches of the plot, clearly being separated by PC1 alone, but sitting in the middle of PC2. This suggests that the blending of varietals during

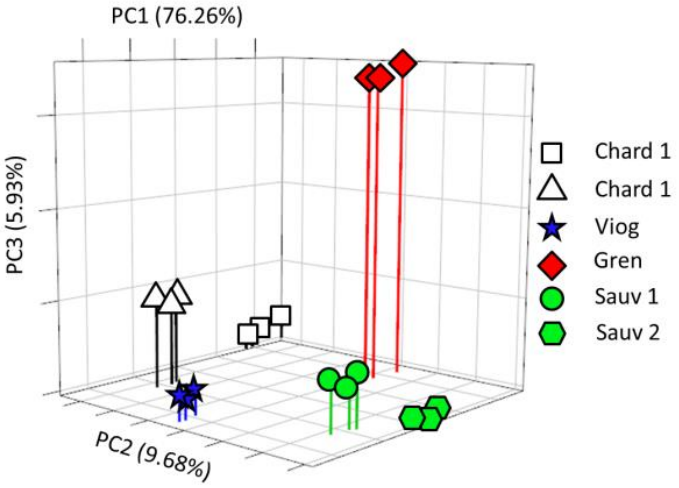

Figure 3.2. Differentiation of BSA treated white wine varietals. PCA plot showing the separation of white wines. Two bottles of Chardonnay (Chard 1 – Januik Winery and Chard 2 – Abeja Winery), one bottle of Viogner/Chardonnay (Viog – Qupe Wines), one bottle of Grenache/Roussane/Viogner (Gren – Nostre Pais), and two bottles of Sauvignon Blanc (Sauv – Dog Point Vineyard and Sauv – Whitehall Lane Winery)

production may have averaged the metabolites contributing to loadings in that component. The blends were well separated along PC2 and PC3. As a whole, using three principle components showed that the assay was effective for the white wines.

Based on these observations, a second approach to data analysis was applied with the intent of estimating the relationship between samples. Agglomerative hierarchical clustering with Euclidean distance was used to determine the distances between the wines and to test if bottles of the same varietal are more closely related than bottles of different varietals. Figure 3.3 shows the results of the clustering analysis. As expected, samples of the same varietal were grouped closely together, forming tight clusters with very short branch lengths. Both bottles of Sauvignon Blanc cluster under the same branch; similarly, both bottles of Chardonnay were placed in a cluster together on the same branch. The

clustering results show that the Viogner/Chardonnay (Viog) sample is the most different and that the Grenache/Roussane/Viogner (Gren) is in-between the Chardonnays and Sauvignon Blancs. Just as with the red wines, the treatment of white wines with the protein sensor assay demonstrated clear discrimination of the samples.

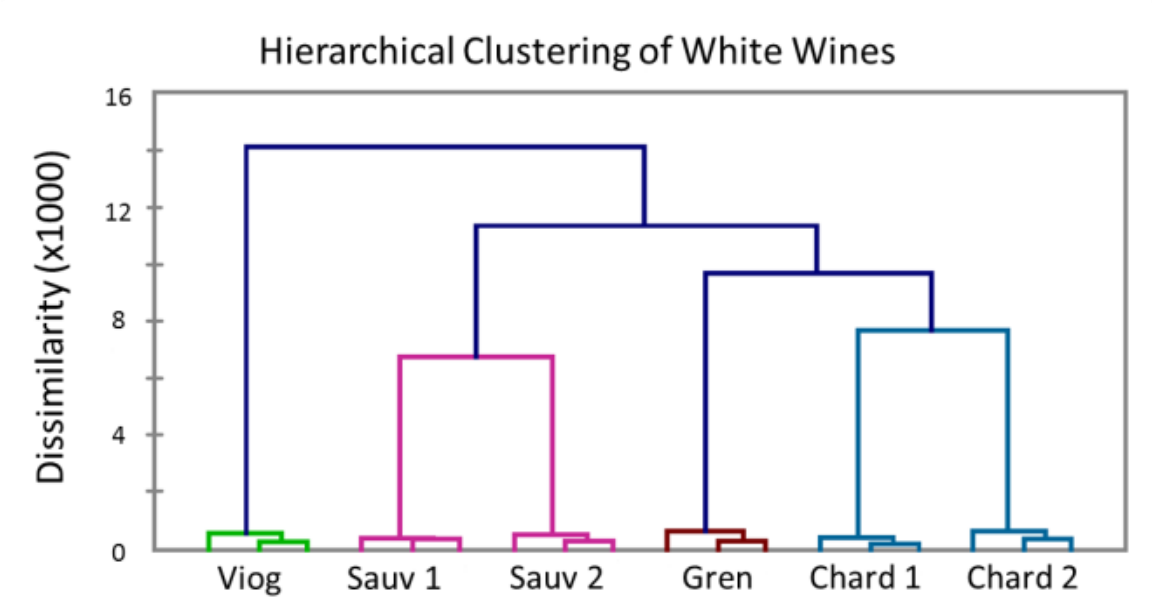

Figure 3.3. Clustering of BSA treated white wine varietals. Agglomerative Hierarchical Clustering dendrogram showing the grouping of white wines. Two bottles of Chardonnay (Chard 1 – Januik Winery and Chard 2 – Abeja Winery), one bottle of Viogner/Chardonnay (Viog – Qupe Wines), one bottle of Grenache/Roussane/Viogner (Gren – Nostre Pais), and two bottles of Sauvignon Blanc (Sauv 1 – Dog Point Vineyard and Sauv 2 – Whitehall Lane Winery).

Analysis of Pig Urine in a Clinically Relevant Large Animal Model Using the BSA Assay

As a final trial for our protein sensor assay, urine samples from pigs that were part of a clinically relevant study looking at the physiological response to hemorrhagic shock were tested (R. Lusczek et al. 2011). Analysis of metabolites in urine is complicated by large variations in volume, concentration, and biological variation making this a

challenging test of assay performance. PCA of samples from the control group (injured but not in shock) and a shock group (injured with loss of blood) (n=3) are presented in Figure 3.4. At the beginning of the time course, 0 minutes, the groups cannot be separated. At the first time point after trauma (45 minutes) the two groups were well separated.

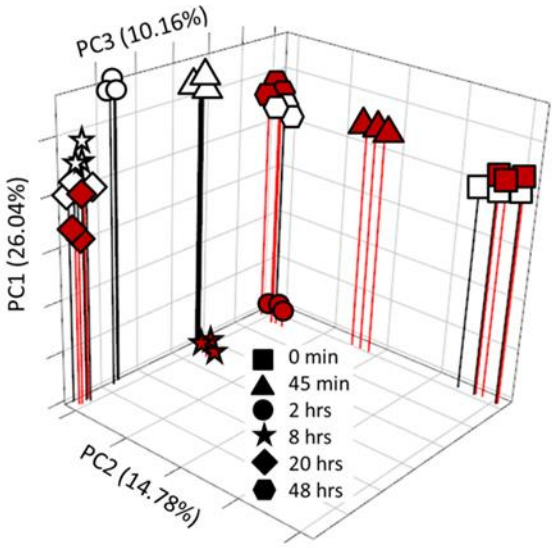

Figure 3.4. Analysis of animals entering and recovering from shock with BSA treatment. PCA plot of BSA protein sensor data from animals in hemorrhagic shock (red) or control (white), where the lines are to help orient the reader as to where the data points are in relationship to the XY plane. Both groups were followed for 48 hours, the time of recovery. The large separation of groups at 45 minutes and 2 hours are the most important from the perspective of emergency medical care.

The control and hemorrhagic shock groups were well separated at 2 and 8 hours post event as well. As the animals progressed through recovery, separation decreased and by 48 hours post injury the two groups could no longer be differentiated. From the standpoint of emergency medicine, being able to distinguish patients 45 minutes and 2 hours post trauma into those that are likely to progress to hemorrhagic shock and those

that are not, is a significant outcome. Once again, the BSA protein sensor was able to distinguish between complex solutions of metabolites, this time in a clinically relevant large animal model. Further analysis of this data is ongoing in our lab to identify the compound(s) responsible for differentiation, as they are potential biomarkers for the onset of hemorrhagic shock. A significant point was the ability of the assay to cluster animals at a given time point, yet separate between time points even though urine output varied as much as 50 fold and was not directly correlated with concentration (data not shown). This

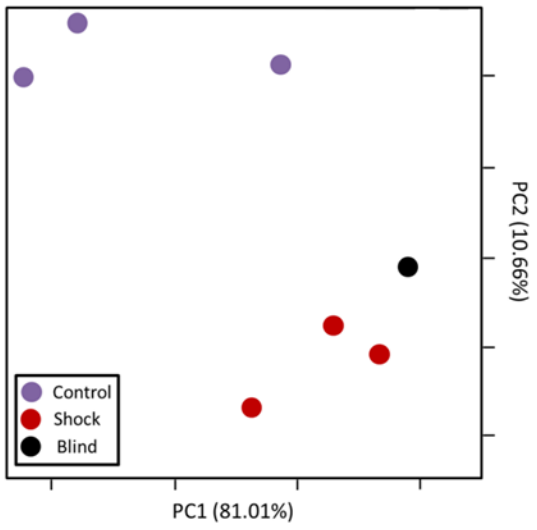

Figure 3.5. Blind test of urine samples at 2 hours post shock. PCA plot showing the separation of control and shock urine samples, with the addition of a previously unanalyzed shock sample. The plot clearly shows that the black dot (blind shock sample) clusters with the other shock samples.

is an important factor to consider because large variations in urine output and concentration complicate any quantitative analysis using signal intensity and would be expected in the case of humans. This also indicated that BSA was able to selectively bind

small molecules across a range of starting concentrations. No method to date exists to rapidly identify the onset of shock in an organism as well as the assay described here.

As a final test of the method, we analyzed a urine sample collected from a shock experiment conducted nearly a year after the original group. Data from the additional sample, which was randomly selected from the shock group, was added and the PCA analysis repeated. Figure 3.5 shows that the additional "blind" sample grouped closest to the shock animals as expected if the protein sensor was working. A single sample was used in this test to mimic actual use of the assay for categorizing a "new" sample in an emergency situation, which would involve comparing data from a patient to previously analyzed control and shock samples.

Examination of Results from BSA Treated Samples and Comparison with Untreated Samples

To examine the value added by using the protein sensor assay, analysis of the metabolite samples without the BSA step was conducted. Aliquots of red wine were processed just as before except the BSA binding step was left out. A PCA plot of the molecular features showed that differentiation of the varietals was now tenuous in the case of the merlot and cabernet, and differentiation based on time after uncorking was no longer possible (Figure 3.6A). Analysis of the urine samples without the BSA resulted in an even more dramatic change in the results. Individual time points of control and stress animals were intermixed and in places, multiple times points clustered together. This is in sharp contrast to the PCA plot shown in Figure 3.4, in which protein sensor treated urine samples were clearly discernible from one another at a given time point, and the stress and control samples show separation from each other throughout the time course. A

check of the features making up each principle component (loadings) with and without BSA treatment revealed that the features used for separation were different. This indicates that improved performance is due to compound selectivity by BSA.

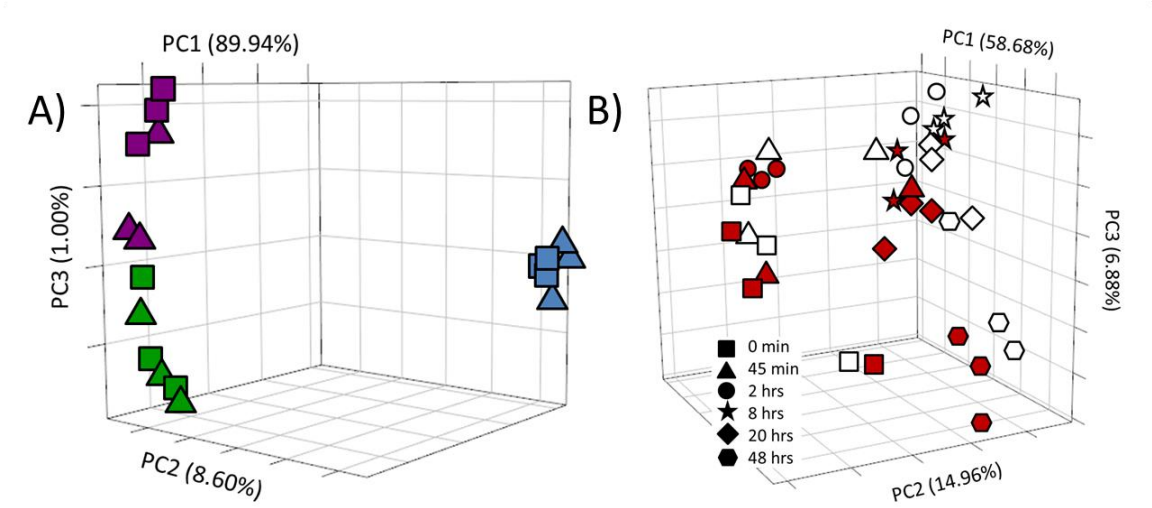

Figure 3.6. Analysis of untreated red wine and urine samples. A) PCA plots of red wine varietals 30 minutes (squares) and 24 hours (triangles) after opening with cabernet (green), merlot (purple), pinot noir (blue). B) PCA plots of urine samples without BSA treatment with hemorrhagic shock animals (red) and control (white). Data shows the entire 48 hour time course.

The small molecule complexity of samples changed drastically when treated with BSA. Urine samples were complex and highly variable before treatment having an average of 3245 ± 48% features. After BSA treatment, 471 features were found to be common in all samples. The complexity of wine was similarly reduced, with untreated wine containing an average of 1241 ± 18% features which was reduced to 396 common features after BSA treatment. This reduction of features could play an important role in biomarker discovery by reducing sample complexity, removing noise, and simplifying data analysis. In addition, analysis time was significantly reduced. An LC run of greater

than 30 minutes was used for the untreated samples whereas a 15 minute run was used after BSA treatment.

Conclusion

Using two disparate examples, it has been shown that BSA can be used as a biosensor to differentiate samples containing a complex mixture of small biomolecules. The method that was developed could separate red wines by varietal and time after uncorking, as well as closely related white wines. It was also shown that urine could be used to categorize animals into control and hemorrhagic shock groups in a clinically relevant large animal model. In the case of the urine samples, metabolite concentration varied widely, yet the method performed well. The initial hypothesis behind this work was that BSA would selectively bind a range of metabolites by integrating binding and allosteric effects. At this point the role of allostery in the process is unclear; however the binding was selective, when gauged by the number of compounds detected by LCMS, reducing the sample complexity by as much as a third. The specific nature of the bound compounds is the subject of ongoing research. Clear strengths of this assay are that it is straight-forward, rapid, and should be adaptable to a number of detector technologies.

Acknowledgements

The authors would like to thank Dr. Philip Sullivan for helpful discussions and critical reading of the manuscript. The Murdock Charitable Trust and NIH 5P20RR02437 of the Cobre program for support of the mass spectrometry facility at MSU, BB receives

support from Office of Naval Research grant #N00014-06-01-1016 and the Nation Science Foundation MCB 1022481.

References

Adams MM, Anslyn E V (2009) Differential sensing using proteins: exploiting the cross-reactivity of serum albumin to pattern individual terpenes and terpenes in perfume. J Am Chem Soc 131:17068–17069. doi: 10.1021/ja908319m

Addinsoft (2013) XLSTAT.

Adler D, Murdoch D (2012) rgl: 3D visualization device system (OpenGL).

Anslyn E V (2007) Supramolecular analytical chemistry. J Org Chem 72:687–99. doi: 10.1021/jo0617971

Bhattacharya B, Nakka S, Guruprasad L, Samanta A (2009) Interaction of bovine serum albumin with dipolar molecules: fluorescence and molecular docking studies. J Phys Chem B 113:2143–2150. doi: 10.1021/jp808611b

Bisson LF, Waterhouse AL, Ebeler SE, et al (2002) The present and future of the international wine industry. Nature 418:696–699. doi: 10.1038/nature01018

Boulanger L, Joshi A V., Tortella BJ, et al (2007) Excess Mortality, Length of Stay, and Costs Associated with Serious Hemorrhage among Trauma Patients: Findings from the National Trauma Data Bank. Am Surg 73:1269–1274.

Cherkas D (2011) Traumatic hemorrhagic shock: advances in fluid management. Emerg Med Pract 13:1–19.

Cuadros-Inostroza A, Giavalisco P, Hummel J, et al (2010) Discrimination of wine attributes by metabolome analysis. Anal Chem 82:3573–3580. doi: 10.1021/ac902678t

Gallagher LT, Heo JS, Lopez MA, et al (2012) Pattern-based discrimination of organic acids and red wine varietals by arrays of synthetic receptors. Supramol Chem 24:2:37–41.

Gutierrez G, Reines HD, Wulf-Gutierrez ME (2004) Clinical review: hemorrhagic shock. Crit Care 8:373–381. doi: 10.1186/cc2851

Jaitz L, Siegl K, Eder R, et al (2010) LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem 122:366–372. doi: 10.1016/j.foodchem.2010.02.053

Jolliffe JT (2002) Principle Component Analysis, 2nd ed. Springer

Kauvar DS, Wade CE (2005) The epidemiology and modern management of traumatic hemorrhage: US and international perspectives. Crit Care 9:S1–9. doi: 10.1186/cc3779

Kim K, Aronov P, Zakharkin SO, et al (2009) Urine metabolomics analysis for kidney cancer detection and biomarker discovery. Mol Cell Proteomics 8:558–570. doi: 10.1074/mcp.M800165-MCP200

Kubarych CJ, Adams MM, Anslyn E V (2010) Serum albumins as differential receptors for the discrimination of fatty acids and oils. Org Lett 12:4780–4783. doi: 10.1021/ol101906g

Kumar A, Remmel RP, Mann HJ, Beilman GJ (2011) Drug metabolism in hemorrhagic shock: pharmacokinetics of selective markers of cytochrome-P450 2C9, 2D6, and 3A4 enzyme activities in a porcine model. J Surg Res 167:e231–243. doi: 10.1016/j.jss.2010.06.040

Martí MP, Busto O, Guasch J (2004) Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing. J Chromatogr A 1057:211–217. doi: 10.1016/j.chroma.2004.08.143

Mcgovern P (2000) Origins and Ancient History of Wine. Psychology Press

Nafisi S, Bagheri Sadeghi G, PanahYab A (2011) Interaction of aspirin and vitamin C with bovine serum albumin. J Photochem Photobiol B Biol 105:198–202. doi: 10.1016/j.jphotobiol.2011.09.002

Peters Jr. T (1996) All About Albumin. Academic Press, San Diego, California

R Core Team (2012) R: A language and environment for statistical computing.

R. Lusczek E, Nelson T, Lexcen D, et al (2011) Urine Metabolomics in Hemorrhagic Shock: Normalization of Urine in the Face of Changing Intravascular Fluid Volume and Perturbations in Metabolism. J Bioanal Biomed 03:38–48. doi: 10.4172/1948-593X.1000041

Scribner DM, Witowski NE, Mulier KE, et al (2010) Liver metabolomic changes identify biochemical pathways in hemorrhagic shock. J Surg Res 164:e131–139. doi: 10.1016/j.jss.2010.07.046

Sjoholm I, Ekman B, Kober A, et al (1979) Binding of Drugs to Human Serum Albumin: XL. Mol Pharmacol 16:767–777.

Skarda DE, Putt KS, Hergenrother PJ, et al (2007) Increased poly(ADP-ribose) polymerase activity during porcine hemorrhagic shock is transient and predictive of mortality. Resuscitation 75:135–44. doi: 10.1016/j.resuscitation.2007.02.020

Smith CA, Want EJ, O'Maille G, et al (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Anal Chem 78:779–787.

Soleas GJ, Diamandis EP, Goldberg DM (1997) Wine as a biological fluid: history, production, and role in disease prevention. J Clin Lab Anal 11:287–313.

Spector AA, John K, Fletcher JE (1969) Binding of long-chain fatty acids to bovine serum albumin. J Lipid Res 10:56–67.

Tautenhahn R, Boettcher C, Neumann S (2008) Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504.

Tautenhahn R, Patti GJ, Kalisiak E, et al (2011) metaXCMS: Second-Order Analysis of Untargeted Metabolomics Data. Anal Chem 83:696–700.

Taylor JH, Beilman GJ, Conroy MJ, et al (2004) Tissue energetics as measured by nuclear magnetic resonance spectroscopy during hemorrhagic shock. Shock 21:58–64. doi: 10.1097/01.shk.0000101674.49265.93

Umali AP, Anslyn E V (2010) A general approach to differential sensing using synthetic molecular receptors. Curr Opin Chem Biol 14:685–692. doi: 10.1016/j.cbpa.2010.07.022

Varshney A, Sen P, Ahmad E, et al (2010) Review Article Ligand Binding Strategies of Human Serum Albumin : How Can the Cargo be Utilized ? Chirality 22:77–87. doi: 10.1002/chir

Villiers A De, Vanhoenacker G, Majek P, Sandra P (2004) Determination of anthocyanins in wine by direct injection liquid chromatography–diode array detection–mass spectrometry and classification of wines using discriminant analysis. J Chromatogr A 1054:195–204. doi: 10.1016/j.chroma.2004.07.087

CHAPTER 4

ADDING METRICS TO THE AGING OF WHISKEY USING A PROTEIN SENSOR ASSAY

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Timothy Hamerly

Contributions: Prepared whiskey samples, treated samples with PSA, acquired LC-MS data, analyzed data, generated figures, wrote manuscript.

Co-Author: Brian Bothner

Contributions: Provided insight and interpretation of results, edited manuscript.

Manuscript Information Page

Authors: Timothy Hamerly, Brian Bothner

Status of Manuscript:
__X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

ADDING METRICS TO THE AGING OF WHISKEY USING A PROTEIN SENSOR ASSAY

The following work is currently in progress to be submitted for publication

Timothy Hamerly,† and Brian Bothner*†

† Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717

*To whom correspondence should be addressed:
Brian Bothner
Department of Chemistry and Biochemistry
Montana State University
Bozeman, MT 59717
Phone: 406-994-5270
FAX: 406-994-4807
E-mail: bbothner@chemistry.montana.edu

Abstract

We have developed a protein-based sensor assay that is both highly selective and broadly applicable for differentiating complex biological samples from one another. In this work, this protein based assay was applied to a set of whiskey samples produced by a craft distillery. The assay involves incubating samples with serum albumin protein, followed by analysis of bound molecules by mass spectrometry. The results of this analysis showed that whiskey samples aged for different lengths of time could be separated using unsupervised statistical analyses and a visible trajectory of the aging process was apparent. In a separate experiment designed to determine if subtle changes made during processing could be detected, partially aged whiskey was transfer to new barrels and subsequently tested. In this case, only subtle differences were present between the samples, yet they were easily separated, indicating that the assay is sensitive to subtle

changes between samples. Our protein-based assay shows great promise as a tool to help distilleries maintain consistent whiskey maturation across batches, track aging, and detect adulteration.

Introduction

The process of fermenting cereal grains into ethanol, and subsequently alcoholic beverages, has a long and storied tradition. It is thought that the first alcoholic beverage produced by fermentation came about in China, nearly 9,000 years ago (Russell and Stewart 2014). The ancient Egyptians and later the Greeks refined the process and began to understand the benefits of consuming alcoholic beverages (Soleas et al. 1997; Mcgovern 2000). However, it was not until the 1500's when the distillation of fermented cereal grains began to grow in England, Ireland, and Scotland, and the production of whiskey was born (Russell and Stewart 2014). Since then, whiskey has grown into a world market, with production in North America, Europe, and Asia (Russell and Stewart 2014). As whiskey demand grew, so did production, and along with this came an increase in adulteration (Burns 2012; Russell and Stewart 2014). To amend this, countries put in place measures designed to standardize production and quality (High and Coppin 1988; Burns 2012; Russell and Stewart 2014). The rise of craft distilleries in the last decade is changing the landscape of the whiskey market and raising old questions (High and Coppin 1988; Headley and Hardy 1992; Pontes et al. 2006; Sundale Research 2013). From a business standpoint, a method for speeding up the aging process of whiskey has sparked a boom in approaches and businesses focused on maturing whiskey (Singleton

1962). This has heightened the need for analytical methods that can provide metrics that allow distilleries to track the aging process and identify adultered or suboptimal product.

Recently, the utility of a protein-based assay for differentiating wine varietals and other complex biological fluids was presented (Hamerly et al. 2014a; Hamerly and Bothner 2015). The assay uses a highly abundant protein in mammalian blood, serum albumin (SA), as a sensor that binds subsets of small molecules. Analysis of the selectively bound molecules by liquid chromatography mass spectrometry (LC-MS) improves the ability to differentiate and categorize samples. The assay takes advantage of the biological role of SA as a transport protein in blood, picking up and delivering a cargo of fatty acids, drugs, metal ions, and other small molecules (Reynolds et al. 1968; Spector et al. 1969; Spector et al. 1971; Halim et al. 2008; Bhattacharya et al. 2009; Fanali et al. 2012). In a subsequent publication it was demonstrated that the small molecules can be concentrated by as much as 100-fold when treated using the protein sensor assay (PSA) (Hamerly and Bothner 2015). This work also showcased the assays ability to differentiate subtle differences in cells recovering from oxidative stress, where traditional LC-MS failed.

With these successes in mind, a series of experiments were undertaken to determine whether the PSA could be of use in the production and aging of whiskey. To this end, experiments were set up to test the value of the PSA in tracking larger changes brought about through the aging of whiskey and if the approach was sensitive to the subtle changes brought about by alternative aging processes. Samples of whiskey were obtained from a local craft distillery and treated with the PSA before analysis by LC-MS.

Results showed that the PSA was able to differentiate whiskey samples, both with respect to time and barrel type. Although more testing is needed, the work presented here could have implications for craft distillers and larger manufacturers alike, where the PSA could be used to give a metric to the quality control process in whiskey making.

Methods

Materials

Bovine serum albumin (greater than 98% agarose gel electrophoresis pure) was purchased from Sigma-Aldrich (St. Louis, MO). Molecular weight cutoff spin filters and syringe filters were purchased from Pall (Port Washington, NY). The buffer agent 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) and Sodium Chloride were purchased from VWR (Radnor, PA) at a purity of 99% or greater. All solvents were purchased as HPLC grade; water from Avantor (Center Valley, PA), methanol and acetonitrile from EMD Chemicals Inc. (Gibbstown, NJ). Formic acid (98% GR ACS) for use as an LC ion pairing agent was purchased from EMD Chemicals Inc. (Gibbstown, NJ). Wheat bourbon and rye bourbon whiskies were supplied by Willie's Distillery Inc. (Ennis, MT). Samples of whiskey were taken from barrels, placed in Falcon tubes and stored at -80 °C until analysis.

Preparation of BSA

BSA was prepared for use in the PSA exactly as previously described (Hamerly et al. 2014a; Hamerly and Bothner 2015). Briefly, a solution of BSA at 30 mg mL^{-1} in 10 mM HEPES buffer (pH 7.0) was washed over a 3kD molecular weight cutoff (MWCO)

spin filter. The protein was washed three times with HEPES buffer to remove small molecules and contaminants remaining which would interfere with the assay. Following washing, BSA was re-suspended 10 mM HEPES buffer (pH 7.0) at a concentration of roughly 30 mg mL^{-1} and frozen at -20 °C until analysis. Prior to analysis, the BSA stock was diluted to 0.1 mg mL^{-1} with 250 mM NaCl, 10 mM HEPES (pH 7.0) buffer. Molecular weight cut-off (MWCO) spin filters were washed with deionized water prior to use to remove any preservatives as recommended by the manufacturer.

PSA Treatment of Whiskey Samples

Prior to PSA treatment, whiskey samples were pre-filtered with a 10 kDa MWCO spin filter to remove proteins and particulates from the manufacturing process. A filtered whiskey aliquot of 100 μL was combined with 300 μL of H_2O and 100 uL of BSA at 0.1 mg/mL in HEPES buffer and allowed to equilibrate for 5 minutes. The solution was then washed over a 3 kDa MWCO spin filter and spun at 9500 x g for 5 minutes. The supernatant was removed and the filter was then washed with a solution of 10 mM HEPES buffer (pH 7.0, 250 mM NaCl) to remove nonspecific material. Following this, 70% MeOH was added to the top of the spin filters and centrifuged three times to remove metabolites bound to BSA. The supernatant containing metabolites which were bound was than dried using a speed-vac and stored at -80 °C. To control for containments stemming from protein, solvent, and filters throughout the PSA process, experimental blanks were conducted.

Mass Spectrometry Analysis of PSA Treated Whiskey Samples

Analysis by LC-MS was done as described in detail previously (Hamerly et al. 2014a; Hamerly and Bothner 2015). Briefly, samples were re-suspended in 50% aqueous methanol (MeOH) and injected onto a Kinetex 1.7 μm C18, 150 mm x 2.1 mm column (Phenomenex, Torrance, CA) kept at 50 °C. An Agilent 1290 UPLC (Santa Clara, CA) system with a flow rate of 600 μL min^{-1} was used for LC separation. The solvent system consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile for solvents A and B, respectively. A solvent program of 2% B for 2 minutes, to 95% B over 10 minutes, with a hold at 95% B for 2 minutes, before finally returning to 2% B for 1 minute, for a total run time of 15 minutes was used. For MS acquisition, an Agilent 6538 Q-TOF Mass Spectrometer (Santa Clara, CA) was used, operating in positive ion mode, with a cone voltage of 3500V and fragmentor voltage of 120V. Drying gas (N_2) temperature was 350 °C with a flow of 12 L min^{-1} and the nebulizer was set to 60 psig. Spectra were collected at a rate of 2.5 Hz with a mass range of 50 to 1000 m/z.

LC-MS Data Analysis

A pipeline for analysis of LC-MS data from PSA treated samples was modified from previously described work from our lab (Hamerly et al. 2014a; Hamerly et al. 2014b; Hamerly and Bothner 2015). Briefly, data was converted to mzML file format using Agilent Masshunter Software (Santa Clara, CA), and aligned using MZmine 2.14. Data files converted from Masshunter were imported into MZmine as raw data files. A crop filter was applied to each data file to remove the first minute of each run, where the

LC stream was sent to waste. Features were extracted as centroid and lists of features with a minimum elution time span of 0.1 minutes and a minimum height of 5000 a.u. were generated. Peak lists were retention time normalized and then aligned into one list with a max tolerance of 0.02 m/z or 30.0 ppm to define the same peak across runs. This aligned peak list was then gap filled with the same tolerances to find missing peaks. This list was then exported as a .csv file for statistical analysis. An ANOVA was done using Excel and features which were found to be significant (fold >1.5, p value <0.05) between any two samples were retained for further statistical analysis. The resulting lists of significant features were analyzed using principal component analysis (PCA) and visualized using the R package rgl, while hierarchical clustering was done using XLStat (Adler and Murdoch 2012; R Core Team 2012; Addinsoft 2013). Bubble plots were generated from significant features for each pairing of rye whiskey using the built in plot function in R (R Core Team 2012).

Results and Discussion

The success of our prior work on the classification of wine, urine, and intracellular metabolites prompted us to try the PSA in the analysis of whiskey. As an initial test, a set of wheat based whiskey ranging in age from 0 to 9 months was treated with the PSA and analyzed by LC-MS. This test served two purposes, i) to develop a method for treating whiskey with the PSA and ii) determine the viability of the assay for further whiskey analysis. Samples needed to be pre-filtered with a 10 kDa MWCO spin filter and diluted with water to lower the ethanol content, before being incubated with

BSA. The pre-filtering was necessary to remove particulates and large biomolecules to prevent clogging of the spin filter used in the PSA. Dilution of the pre-filtered whiskey was required to avoid denaturing the BSA upon mixing with the sample. This was accomplished by diluting 100 μL of pre-filtered whiskey 3-fold in water. Data containing details on the molecular weight, retention time, and intensity of each compound in the PSA treated samples was collected using LC-MS. This information was imported into MZmine for retention time and feature alignment, and finally statistical analysis to compare the intensity of compounds between samples. To this end, ANOVA analysis was used to identify features that were significantly different between samples (fold change >1.5, p value <0.05).

Features found to be statistically significant in the ANOVA were used as input for principal component analysis (PCA) so that similarities and differences in samples could be visually assessed. PCA is an unsupervised method, meaning it provides an unbiased analysis using all of the input data (Worley and Powers 2013). A 2-dimensional PCA plot of data from the aging test is shown in Figure 4.1. Technical replicates (same marker) were tightly clustered, while samples from each age group (different markers) were well separated. A logical trajectory beginning with the freshly distilled samples (squares) through to the most mature (pentagons) was observed. A dashed line has been added as a visual guide. The result of this analysis shows that the PSA can be used to track the aging process of whiskey. Importantly, the ethanol content of whiskey as it ages drops, but did not appear to affect the PSA and its ability to differentiate samples. This means the method developed for treating whiskey with the PSA is robust enough to work across a

wide range of ethanol concentrations, with the protein remaining intact and capable of binding molecules in a fashion such that it can be used to effectively differentiate time points/samples.

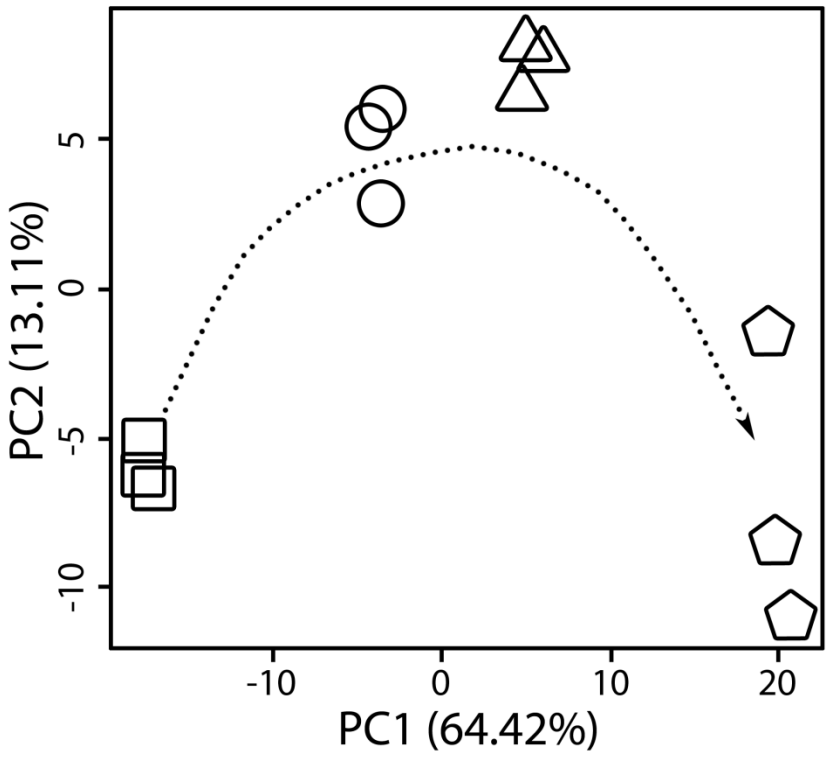

Figure 4.1. PCA plot of wheat bourbon whiskey samples treated with the PSA before LC-MS analysis. Samples represent four different stages of aging. Replicate samples from distillation (day 0, squares), day 174 (circles), day 180 triangles, and day 240 (pentagons). A dashed line has been added as a visual aid showing the trajectory of ageing.

PSA sensitivity to variations in the aging process was tested by analyzing partially matured samples transferred to different barrels for finishing. A rye based whiskey was distilled and placed into a 53 gallon barrel to develop. After two years, portions were transferred into two new 10 gallon barrels, originating from two different manufacturers.

The original stock barrel and the two new barrels were then corked and left to age for an additional 4 months before sampling. The distillery conducted this as part of product testing aimed at developing new tastes and for assessing if barrel type changed the ageing process. Our goal was to assess whether the PSA could provide a metric for comparison. PCA was used to visualize the data from PSA treated samples. Once again, tight clustering of replicates was observed while the different samples were distinctly separated (Figure 4.2a).

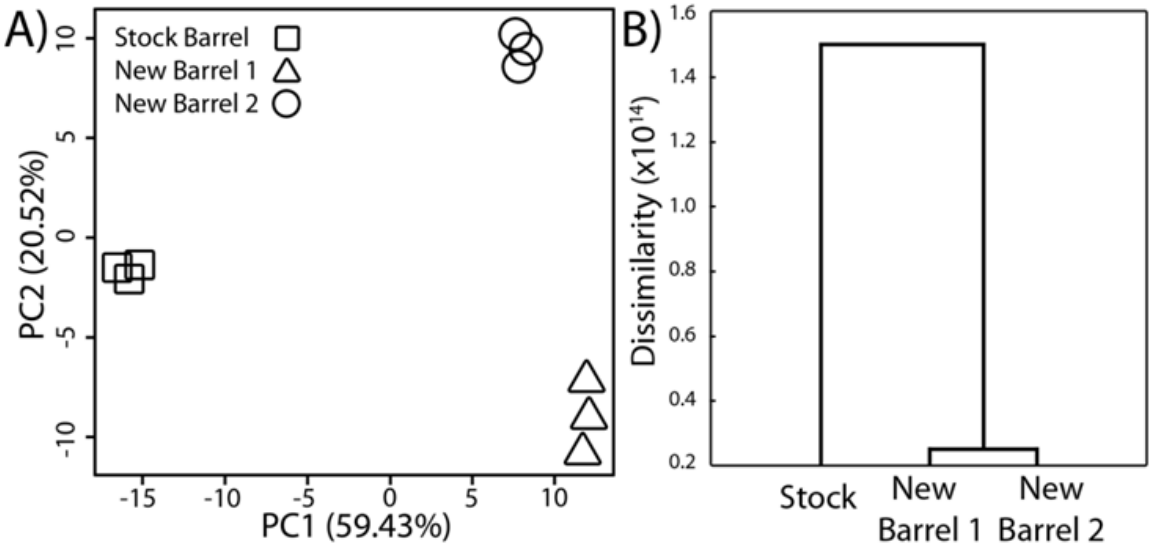

Figure 4.2. Influence of the barrel on whiskey maturation. a) PCA plot of rye bourbon whiskey samples treated with the PSA before LC-MS analysis. Samples are from the same batch of whiskey, which was aged in the stock barrel for 24 months, after which portions were transferred to new barrels of a different make. All three barrels were allowed to age for an additional 4 months. b) AHC of the three barrels of rye bourbon whiskey. The y-axis shows the dissimilarity, where length of the vertical lines are a measure of dissimilarity.

The samples transferred to the new barrels were separated from the stock barrel by PC1, which is composed of the features making the largest contribution to differences

between samples. The samples placed in the new barrels were only separated from one another by PC2, which accounted for roughly one-third of the variation attributed to PC1 (60% in PC1 vs 20% in PC2). This indicates a similar, relatively large change between the new barrel samples and the stock, while a more subtle change differentiates the former. This subtle change is better understood when the data is displayed using agglomerative hierarchical clustering (AHC); a bottom up approach where each sample starts on its own, being merged as the tree is built (Day and Edelsbrunner 1984). Vertical lines indicate dissimilarity of samples in the tree. By AHC, the two new barrels had only a small amount of dissimilarity compared with the stock barrel, which was placed on a separate branch (Figure 4.2b). Together, these results show that the PSA is sensitive to both the small and large changes seen between whiskeys in different barrels, when aging time is kept constant.

The ability of the PSA to differentiate whiskey samples which have similar taste profiles, according to the distiller, prompted us to undertake a more thorough analysis of the features which were significantly different between the barrels based on ANOVA. Bubble plots displaying molecules with significant differences in pairwise analyses of whiskey barrels were generated to help understand the nature of the molecular species which are distinct in each barrel (Figure 4.3). For each plot, retention time and m/z for molecules are displayed on the x- and y-axis respectively; where the size of the circle indicates the fold change of a given feature and the hue signifies the p-value (larger, darker circles have a higher fold change and smaller, more significant p-value). The first

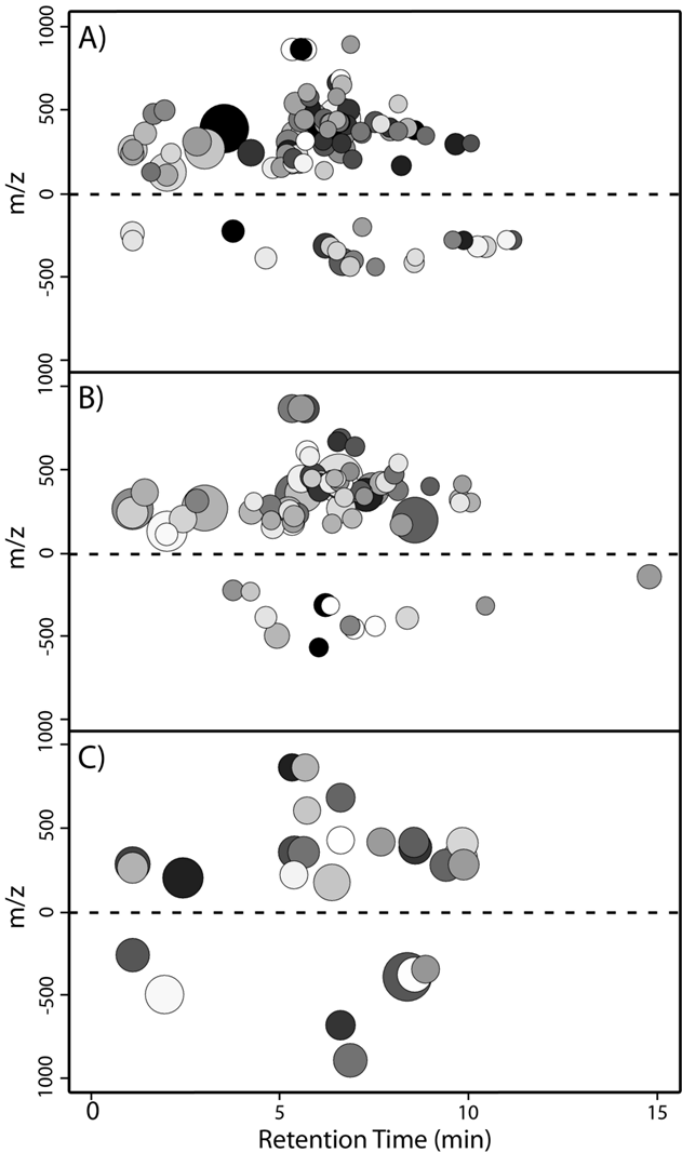

Figure 4.3. Bubble plots generated from the significant features (fold >1.5, p-value <0.05) between each pairing of rye based whiskey samples. For each plot, x-axis represents retention time of the given feature from reverse-phase chromatography, while the y-axis shows the m/z. Size of each bubble represents fold change, and hue represents the p-value, where a larger, darker, circle indicates a greater fold change and smaller p-value. a) Comparison of new barrel 1 and the stock barrel; above the dashed line means more in new barrel 1, below dashed line means more in stock barrel. b) Comparison of new barrel 2 and the stock barrel; above the dashed line means more in new barrel 2, below dashed line means more in stock barrel. c) Comparison of new barrel 1 and new barrel 2; above the dashed line means more in new barrel 1, below dashed line means more in new barrel 2.

two panels show the comparison of the stock barrel to each new barrel, while the third panel shows the comparison of the two new barrels. Most of the features shown in these plots appear to elute in the early to middle part of the chromatogram, which on a reverse phase column indicates they are more polar in nature. In comparing the stock barrel with each of the new barrels, a majority of significantly different features were observed to increase in the new barrel samples (circles above the dashed line) (Figure 4.3a and 4.3b). This suggests that transfer to a new barrel introduced or enhanced the development of specific components. However, there are molecules that decrease in abundance as well (circles below the dashed line), suggesting the new barrels may in general speed removal of specific compounds.

Although the majority of the changes were between the stock and new barrel samples, differences were also detected between the two new barrel samples (Figure 4.3c). The features in this bubble plot, both above and below the dashed line, are compounds which are characteristic of each new barrel. Of the 80-100 features which are more intense in the comparison of the new barrels to the stock barrel, only a select few are retained when comparing the new barrels with one another. These select few are likely what give rise to subtle difference in flavor profile described by the distiller, and visualize in the PCA and AHC. The rest of the features which showed significant change between the stock and new barrels (above the dashed line in 4.3a and 4.3b) are the same compounds, and thus do not show up in the comparison of the two new barrels (panel 4.3c). Some of these compounds were only detected in samples from the new barrels, however, the majority were merely enhanced by exposure to the new barrels.

A plausible explanation for why introduction to a new barrel caused significant change is that the fresh surface for interaction facilitates changes in the whiskey. Contacting a fresh surface may allow the whiskey in the new barrels to leech compounds from the wood and to have compounds depleted by absorbance of the wood. The number of features that were different between the two new barrels and the stock barrel was roughly 100. The number of different features between the two new barrels was just over 25. The small number of difference in the new barrels when compared to one another, Is consistent with the PCA, which shows them having almost no difference in PC1, with differentiation relying on PC2, which accounted for less of the total variance. It is also consistent with AHC which indicated that the two new barrels are highly similar, while there is significant dissimilarity between the stock barrel and the new barrels. The results of this analysis show that transfer to new barrels increases the rate of change, suggesting this as a method for speeding the process of aging. However, the nature of this change and if is related to true ageing or represents a repeat of the initial process when first barreled will require further testing.

As a whole, the PSA performed well in the differentiation of whiskey and could be used to provide information in the form of a fingerprint useful in tracking the aging process or investigating differences in flavor profile imparted by specific barrels. Aging is the key process that takes clear, bitter distillate and transforms it into the polished product ready for consumption and retail. The process requires skilled craftsman with years of experience to ensure a consistent and high quality product. Our analysis shows that the PSA can track the aging process and could be used in combination with a master

distiller or as a stand-alone assay to guide production and signal completion. The PSA was also sensitive to changes brought on by different barrels which could be of interest to those seeking alternative strategies in the production process or as an analytical tool to assess aging methodology. The sensitivity of the PSA suggests that it could be adapted as a straight-forward test for counterfeit or adultered product on a broad scale.

Acknowledgements

The authors would like to thank Willie's Whiskey, Ennis, MT for providing the samples. The authors would also like to thank Dr. Phil Sullivan for providing help in obtaining the samples and for thoughtful discussions throughout this work. Tim Hamerly would like to recognize support from the Kopriva Graduate Fellowship at MSU. This work was supported by the National Science Foundation, MCB102248 through an award to Brian Bothner. The Proteomics, Metabolomics, and Mass Spectrometry facility at MSU received support from the Murdock Charitable Trust and NIH 5P20RR02437 of the CoBRE program.

References

Addinsoft (2013) XLSTAT.

Adler D, Murdoch D (2012) rgl: 3D visualization device system (OpenGL).

Bhattacharya B, Nakka S, Guruprasad L, Samanta A (2009) Interaction of bovine serum albumin with dipolar molecules: fluorescence and molecular docking studies. J Phys Chem B 113:2143–2150. doi: 10.1021/jp808611b

Burns E (2012) Bad Whisky. In: 3rd ed. Neil Wilson Publishing, pp 33–34

Day WHE, Edelsbrunner H (1984) Efficient algorithms for agglomerative hierarchical clustering methods. J Classif 1:7–24. doi: 10.1007/bf01890115

Fanali G, di Masi A, Trezza V, et al (2012) Human serum albumin: from bench to bedside. Mol Aspects Med 33:209–90. doi: 10.1016/j.mam.2011.12.002

Halim AAA, Kadir HA, Tayyab S (2008) Bromophenol blue binding as a probe to study urea and guanidine hydrochloride denaturation of bovine serum albumin. J Biochem 144:33–8. doi: 10.1093/jb/mvn036

Hamerly T, Bothner B (2015) Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis. Appl Biochem Biotechnol 1–13.

Hamerly T, Heinemann J, Tokmina-Lukaszewska M, et al (2014a) Bovine serum albumin as a molecular sensor for the discrimination of complex metabolite samples. Anal Chim Acta 818:61–66. doi: 10.1016/j.aca.2014.01.058

Hamerly T, Tripet BP, Tigges M, et al (2014b) Untargeted metabolomics studies employing NMR and LC–MS reveal metabolic coupling between Nanoarcheum equitans and its archaeal host Ignicoccus hospitalis. Metabolomics 11:895–907. doi: 10.1007/s11306-014-0747-6

Headley LM, Hardy JK (1992) Adulterants In Whiskey - Detection And Identification By Principal Component Analysis. J Food Sci 57:980–984.

High J, Coppin C a (1988) Wiley and the Whiskey Industry: Strategic Behavior in the Passage of the Pure Food Act. Bus Hist Rev 62:286–309.

Mcgovern P (2000) Origins and Ancient History of Wine. Psychology Press

Pontes MJC, Santos SRB, Araújo MCU, et al (2006) Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. Food Res Int 39:182–189. doi: 10.1016/j.foodres.2005.07.005

R Core Team (2012) R: A language and environment for statistical computing.

Reynolds J, Herbert S, Steinhardt J (1968) The Binding of Some Long-Chain Fatty Acid Anions and Alcohols by Bovine Serum Albumin. Biochemistry 7:1357–1361.

Russell I, Stewart G (2014) Whiskey: Technology, Production and Marketing, 2nd edn. Elsevier Ltd

Singleton VL (1962) Aging of wines and other spiritous products, acceleration by physical treatments. Hilgardia 32:319–392. doi: 10.3733/hilg.v32n07p319

Soleas GJ, Diamandis EP, Goldberg DM (1997) Wine as a biological fluid: history, production, and role in disease prevention. J Clin Lab Anal 11:287–313.

Spector a a, Fletcher JE, Ashbrook JD (1971) Analysis of long-chain free fatty acid binding to bovine serum albumin by determination of stepwise equilibrium constants. Biochemistry 10:3229–32.

Spector AA, John K, Fletcher JE (1969) Binding of long-chain fatty acids to bovine serum albumin. J Lipid Res 10:56–67.

Sundale Research (2013) State of the Industry: Craft Distilleries and Distilled Spirits in the U.S. Bay Shore, NY

Worley B, Powers R (2013) Multivariate Analysis in Metabolomics. Curr Metabolomics 1:92–107.

CHAPTER 5

ANALYSIS OF WINE USING THE PROTEIN SENSOR ASSAY

Contribution of Authors and Co-Authors

Manuscript in Chapter 5

Author: Timothy Hamerly

Contributions: treated samples with PSA, acquired LC-MS data, analyzed data, generated figures, wrote manuscript.

Co-Author: Brian Bothner

Contributions: Provided insight and interpretation of results, edited manuscript.

Manuscript Information Page

Authors: Timothy Hamerly, Brian Bothner

Status of Manuscript:
_ X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
___ Published in a peer-reviewed journal

CHAPTER 5

ANALYSIS OF WINE USING THE PROTEIN SENSOR ASSAY

Introduction

The use of a protein sensor assay (PSA) as an analytical method to classify samples of differing origin has shown great promise, particularly when compared to traditional analysis where samples were not treated using the assay. Previous work demonstrated that the PSA could be used to classify bottles of wine based on grape varietal or time post uncorking (See chapter 3 of this dissertation, Hamerly et al. 2014). In this previous experiment, three different varietals of wine from the same vineyard (Barefoot) were left to breathe over 24 hours, sampled at various time points throughout, and treated with the PSA prior to liquid chromatography coupled to mass spectrometry (LC-MS) analysis. The results showed that without PSA treatment, wine varietals could be separated using principal component analysis (PCA), but the individual time points were intermixed.

Although these results were promising, a concern with this experiment was that the wine varietals themselves were distinctly different, having only the vineyard in common and thus region of growth, as a constant. It was therefore not surprising that the wines were separated from one another without the use of the PSA. On the other hand, the ability to differentiate time points when a traditional analysis could not clearly shows that the PSA is sensitive to subtle differences in samples. These results prompted an experiment which used a larger set of data to be undertaken. A set of wines from

Bordeaux grape varietals selected by an expert wine sommelier to ensure distinct differences in their taste profiles. In addition to bottles of pure varietals, two blends were also selected to provide an additional challenge to the PSA. As with the previous experiment, each bottle was also left to stand open for 24 hours to further test the capability of the PSA to differentiate samples with only subtle changes brought on by breathing.

Another potential use for the PSA in the food and agricultural business, including wine and its production, is the testing for potentially harmful compounds such as pesticides. Pesticides are commonly sprayed onto grape vines and soil to control for disease and pests, with the goal of having the least impact on the product and the environment (Bisson et al. 2002; Garbi et al. 2010). In addition to pesticides used on the grapes themselves, toxic contaminants stemming from the cork can also be introduced into wine (Khan and Sinha 1996; Cabras and Angioni 2000; Dasgupta et al. 2010). Methods aimed at monitoring pesticides and contaminants are difficult, typically requiring complicated extraction procedures, advanced instrumentation, or both. Using extensive 2-dimensional gas chromatography (GC) allows for the detection of over 150 different pesticides, but at the cost of two hour analysis times per sample (Dasgupta et al. 2010; Garbi et al. 2010). More rapid methods relay on triple quadrupole mass spectrometers which can improve sensitivity and reduce analysis time compared to GC methods. However, to employ multiple reaction monitoring (MRM) used in these methods, a predetermined set of compounds must be selected (Sannino et al. 2004; Goto et al. 2006). Another challenge in analyzing pesticides, regardless of the method used, is

that they can be at concentrations which are orders of magnitude below the main components of wine (Dasgupta et al. 2010). My previous reports have shown that the PSA actually pulls low abundance molecules out, allowing their detection compared to samples analyzed by LC-MS only (Hamerly et al. 2014; Hamerly and Bothner 2015). Given that pesticide contamination of food products is normally at low levels, I reasoned that the PSA would be a good tool to use in testing for chemicals in wine. To test this idea, an experiment was conducted to look for the presence of pesticides in red and white wines treated with the PSA and without treatment. Not only did the PSA pull out pesticides that were below the limit of detection in the analysis by LC-MS, but the PSA performed better than the traditional method by identifying more pesticides.

Classification of Bordeaux Wines

To evaluate the full capabilities of the PSA with regards to the classification of wine, a series of red wines were analyzed by LC-MS after treatment with the assay. These wines were selected by a master sommelier, with the goal of testing wine produced from varietals of Bordeaux grapes. Bordeaux grapes are often considered the best for making red wines by many connoisseurs. In France, red wines produced under the name Bordeaux always contain cabernet sauvignon or merlot grapes; with other varietals blended in small portions depending on the wine makers own tastes. Outside of the Bordeaux region of France, bottles may be labeled based the most predominant grape provided it contains at least 75% of that varietal. The taste profiles of these wines are therefore dominated by the more abundant grape, with subtle characteristics of other

varietals also infused into them. Bottles used for this experiment originated from vineyards in Washington, USA and Argentina. The infusion of multiple varietals which brings about subtle similarities in the taste profile will make it potentially difficult for the PSA to differentiate them. Additionally, bottles labeled as blends were also included in this experiment and contained grapes of the same varietals as the so called pure bottles. As with the previous experiment, each wine was allowed to breathe for 24 hours to determine if the PSA was sensitive to the subtle changes that take place during the breathing process after uncorking.

For this experiment, a set of five bottles labeled as pure varietals and two bottles labeled as blends were treated with the PSA and analyzed by LC-MS. Data from PSA treated wine samples was visually assessed by PCA. Bottles labeled as pure varietals included a cabernet sauvignon, cabernet franc, malbec, merlot, and petit verdot. The bottles labeled as blended wines were as follows: blend 1 contained 50% merlot, 29% syrah, 15% cabernet sauvignon, 6% petit verdot and blend 2 contained 29% cabernet sauvignon, 23% sangiovese, 18% syrah, 15% merlot, 12% cabernet franc. Samples of each wine were collected and flash frozen immediately after uncorking (0 minutes post uncorking). Aliquots of the wines were also left out on the bench top to breathe and samples were collected and flash frozen at 30 minutes and 24 hours post uncorking. Each sample was processed and treated with the PSA as described previously and data was acquired by LC-MS with duplicate injections (Hamerly et al. 2014).

The first analysis of the data was done on individual bottles at the three time points (0 minutes, 30 minutes, and 24 hours after uncorking), looking to see temporal

resolution. In a previous report, the PSA was capable of separating time points with great success, and this result was further tested with each of the bottles in this experiment. For each bottle in this experiment PCA plots were generated and showed that the three time points were separated from one another (Figure 5.1). While most bottles were well resolved between each time point, some of the bottle did not display a significant change between analyses at 0 minutes and 30 minutes post uncorking. One bottle in particular had virtually no separation of the 0 and 30 minute times points based on PCA (Figure 5.1F). For this wine, blend 2, the two time points were intermixed with one another, but separated from the 24 hour time point. Given that the other six bottles were resolved when compared at 0 minutes 30 minutes, albeit subtly in some cases, it suggests that this particular bottle did not change significantly after uncorking for only 30 minutes. At the later time point (24 hours) there was a significant change. This result could be significant in showing that certain bottles of wine do not need to breathe before drinking as little change is seen after 30 minutes of breathing as assessed by PCA. In contrast, the result could also indicate that certain bottles of wine require longer time to breath before drinking.

Having discerned that the PSA was indeed capable of separating individual wine bottles over a time course of breathing, the ability of the PSA to discriminate wine bottles from one another at specific time points was evaluated. For each time point, PCA plots of the seven bottles were generated (Figure 5.2). At each time point the wines were well separated from one another, demonstrating that the PSA was also capable of classifying varietals at a given time post uncorking. Examination of the PCA plot for the wines after

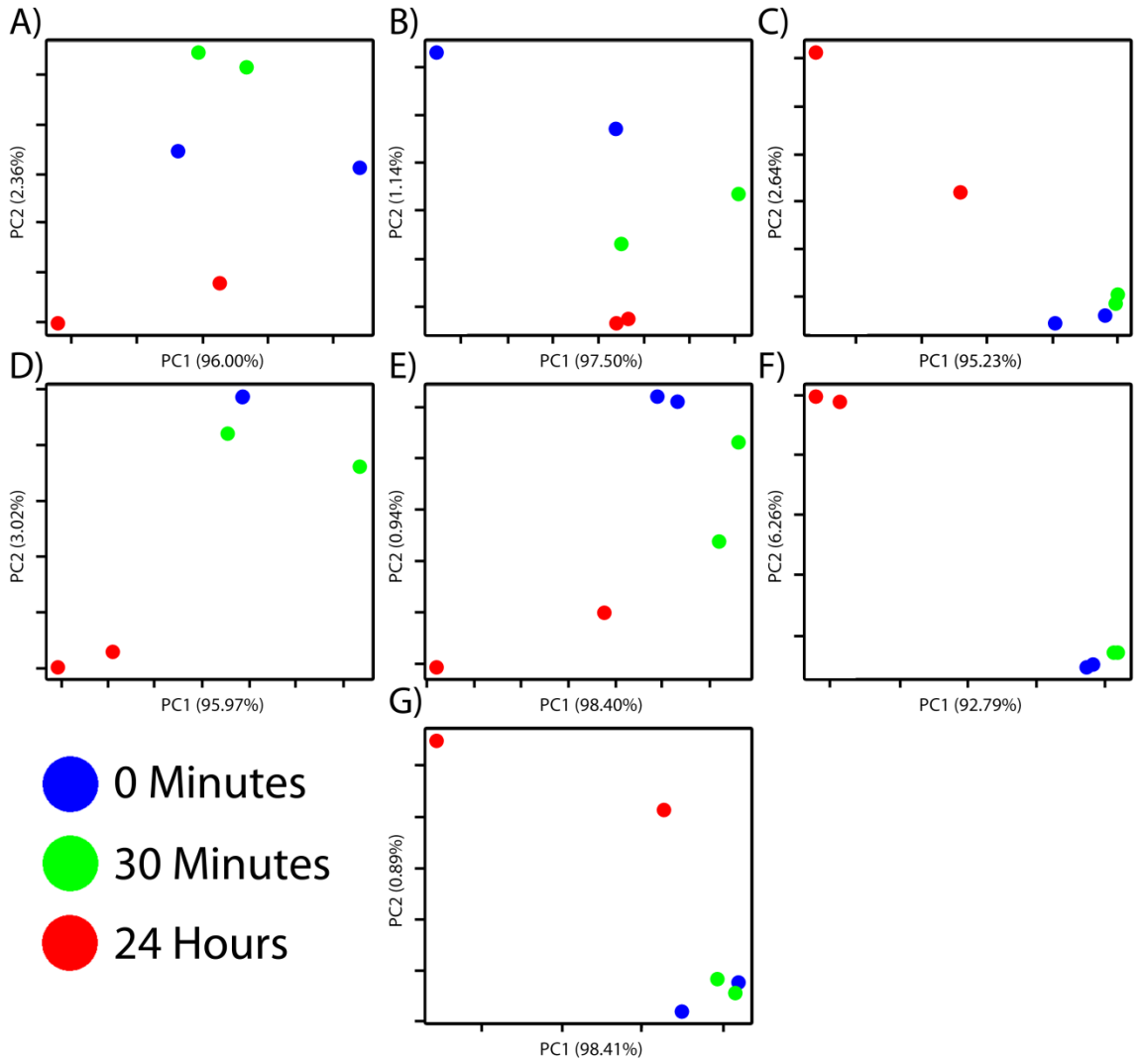

Figure 5.1. PCA plots of each bottle of wine following 24 hours of breathing. Samples of wine were taken at 0 minutes, 30 minutes, and 24 hours after uncorking, treated with the PSA before data acquisition by LC-MS. Data was then visualized for each bottle using PCA: A) malbec, B) merlot, C) blend 1, D) cabernet sauvignon, E) cabernet franc, F) petit verdot, G) blend 2. For each plot, blue dots represent 0 minutes, green dots represent 30 minutes, and red dots represent 24 hours post uncorking.

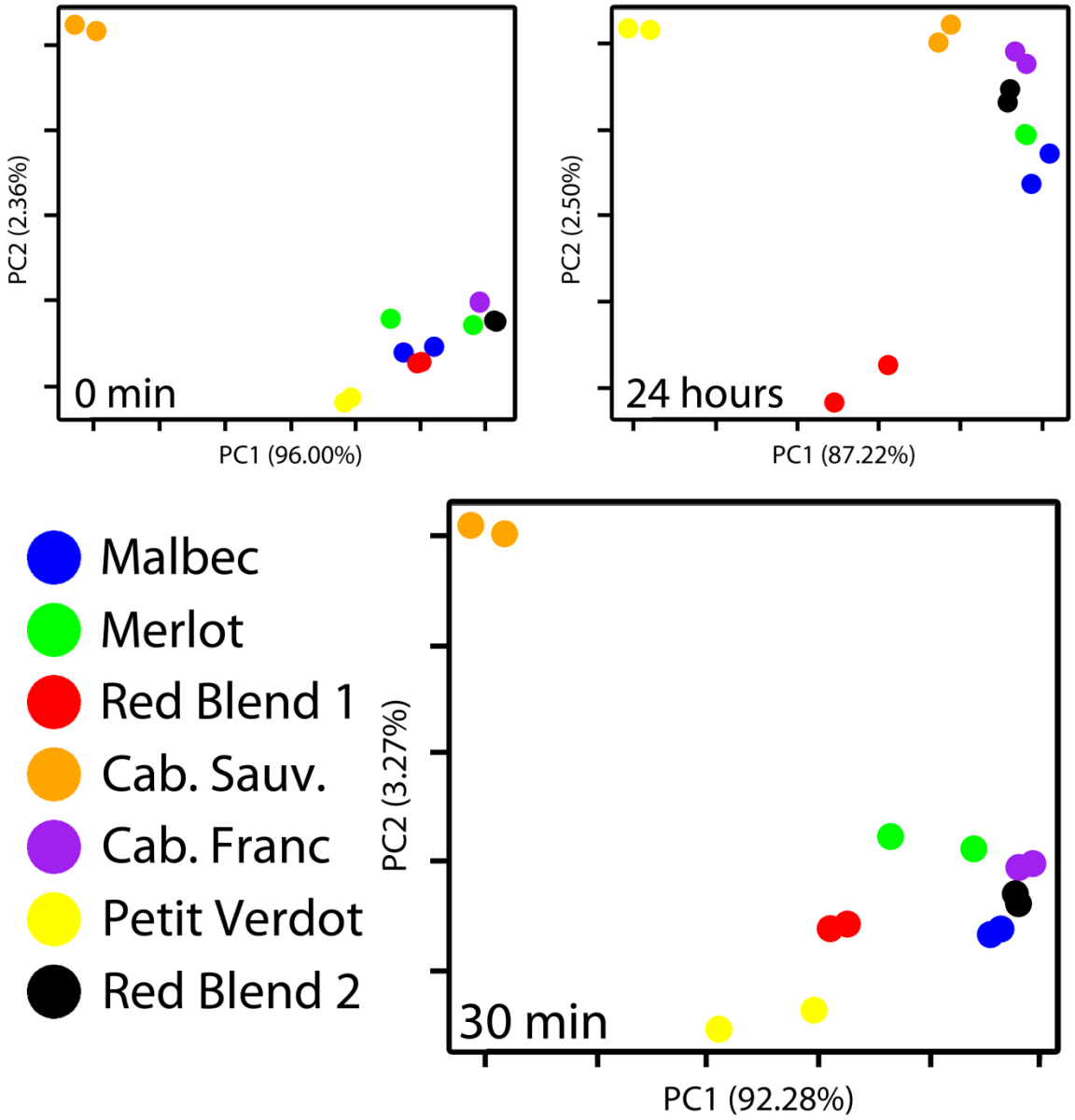

Figure 5.2. PCA plots of wine samples after 0 minutes, 30 minutes, and 24 hours post uncorking. Samples of wine were taken at 0 minutes, 30 minutes, and 24 hours after uncorking, treated with the PSA before data acquisition by LC-MS. Data was then visualized for each bottle using PCA for 0 minutes, 30 minutes, and 24 hours post uncorking. For each plot; blue dots represent malbec, green dots represent merlot, red dots represent blend 1, orange dots represent cabernet sauvignon, purple dots represent cabernet franc, yellow dots represent petit verdot, black dots represent blend 2.

30 minutes of breathing not only shows significant separation of each bottle, but also shows the power of the PSA with respect to blends of varietals. Each of the five pure varietals exists on the edges of the PCA plot (indicating more variation or difference among them), while the two blends are situated in the middle of the plot (indicating less variation from the other bottles). The blends fall in relatively close proximity to the pure varietals of which they are composed. For example, blend 1 contains merlot, cabernet sauvignon, and petit verdot is situated closest to the data points associated to the bottles of pure varietals. This is similar for the blend 2. Even more interesting is that both blends contain a moderate level of syrah grapes and as a result the two blends are positioned in close proximity to one another.

Taking data from all three time points for all seven bottles of wine and analyzing them by PCA provides an interesting picture that highlights a number of trends not previously observed in our analysis of wine. If the PCA plot is generated where the data points are colored base on the bottle (grape varietals), nearly every set of three time points is separated from one another (Figure 5.3A). The exception to this is the cabernet franc which has one time point intermixed with the merlot. Aside from this anomaly, the PSA can clearly differentiate not only the wine varietals themselves, but can also separate the time points from each bottle as well. Some bottles are tightly clustered among their respective time points, while other bottles have one time point which deviates greatly from the other two. This is in agreement with the analysis of each bottle individually, where some bottles showed almost no difference at 0 and 30 minutes post uncorking, but

had a large difference after 24 hours. The same trends can be seen when all wines and time points are plotted together.

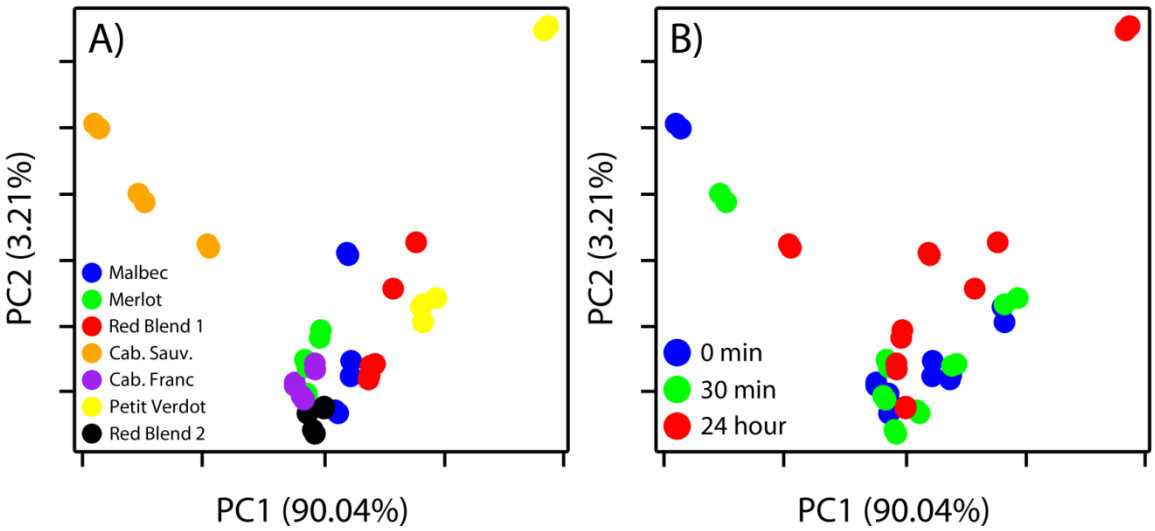

Figure 5.3. PCA plots of all wine samples together following breathing post uncorking. Samples of wine were taken at 0 minutes, 30 minutes, and 24 hours after uncorking, treated with the PSA before data acquisition by LC-MS. Data was then visualized for each bottle using PCA. Plots are colored by wine varietal (A) or time post uncorking (B). For plot A; blue dots represent malbec, green dots represent merlot, red dots represent blend 1, orange dots represent cabernet sauvignon, purple dots represent cabernet franc, yellow dots represent petit verdot, black dots represent blend 2. For plot B; blue dots represent 0 minutes, green dots represent 30 minutes, and red dots represent 24 hours post uncorking.

If the PCA plot is instead colored based on the time point post uncorking, another trend arises which was not previously seen. Although each bottle is separated from one another, the time points paint an interesting picture where as the wines age, they appear to become more similar to each other (Figure 5.3B). At 0 minutes and 30 minutes post uncorking, the wines are more closer to the edge of the PCA plot, where at 24 hours they have come more towards the middle, indicating they are becoming more similar to one another (with the exception being the petit verdot which has a time point of 24 hours at the outskirts of the plot). Additionally, this plot explains why the cabernet franc at 24

hours post uncorking appeared to have a time point intermixed with the merlot data. In essence, the 24 hour time point of the cabernet franc is more closely associated with the other 24 hour time points. This would appear to indicate that over time these particular wines begin to coalesce, with the variation among them decreasing. Further experimentation with other wines may shed more light on this idea.

The use of the PSA to differentiate wines based on their time post uncorking and varietal content shows the power of this analytical tool. In a previous report, a smaller, unrelated set of wines were proven to also be separated under these same conditions (varietal and time post uncorking). The current results provide more conclusive evidence of the true capabilities of the PSA in classifying wines. The classification of the blended varietals among the pure varietals based on the content of each was enlightening, and could be of great use in helping consumers to understand the flavor profile of wines or to find similar tasting wines. Another benefit of the assay is the ability to be useful in understanding the breathing of wines. Not only can the assay classify wines, the PSA could be used to tell which types of wines require breathing prior to drinking. These findings give new purpose to an already beneficial assay.

Pesticide Analysis in Wines

Another potential use of the assay with respect to wine is in the detection of pesticides or other contaminants which make it through manufacturing process into the bottle. Often vineyards will use pesticides to help ensure the successful growth of their wine grapes (Abhilash and Singh 2009; Dasgupta et al. 2010). These pesticides have

proven to be of great use in the wine industry but have the side effect of potentially being present in the wine, and eventually making their way into the digestive system of the consumer. Regulations are in place to help control for such things, limiting the amount that can be found in a given bottle (High and Coppin 1988; Sannino et al. 2004; Garbi et al. 2010). However, it is not common practice to present the levels of pesticide to the consumer on a wine label, nor is it possible to do so if they cannot be detected by traditional methods due to matrix interference or low levels relative to other components of the wine (Dasgupta et al. 2010; Garbi et al. 2010). To evaluate the potential use of the PSA in determining pesticide and contaminant levels in wine, a series of experiments were done to test for the presence of pesticides before and after treatment with the assay. The experiments included a selection of red and white wines and included organic wines as well.

As an initial test, wines were analyzed by LC-MS both with and without PSA treatment and the presence of pesticides was evaluated. A database of over 180 pesticides and other contaminants stemming from the manufacturing process taken from a previous publication were searched in wine samples with and without PSA treatment (Dasgupta et al. 2010). For this experiment, accurate mass was the only criteria used to match pesticides in samples from a database which was built from a two dimensional gas chromatography coupled to mass spectrometry (2D GC-MS) experiment (Dasgupta et al. 2010). A list of pesticides in both treated and untreated wine samples was generated from this analysis (Table 5.1). In total, 31 compounds were putatively identified in the PSA treated samples, while only 29 were found in the untreated samples. Comparison of the

compounds revealed that 22 were in common between the treated and untreated samples, leaving 9 unique compounds found in the PSA treated, and 7 in the untreated (Table 5.2).

Table 5.1. List of pesticides identified with and without PSA treatment.

PSA Treated	**Untreated**
Acephate	Atrazine
Anthracene	Bifenthrin
Bifenthrin	Bisphenol A
Buprofezin	Buprofezin
Butachlor	Butachlor
Cyprodinil	Cyhalothrin
Diafenthiuron	Diafenthiuron
Diflubenzuron	Edifenphos
Edifenphos	Enilconazole
Enilconazole	Etofenprox
Fluoranthene	Fenpropathrin
Fluorene	Fluoranthene
Fluvalinate	Fluorene
Isoproturon	Fluvalinate
Methabenzthiazuron	Isoprothiolane
Metoxuron	Isoproturon
Myclobutanil	Kresoxim-methyl
Oryzalin	Methabenzthiazuron
Paclobutrazol	Metoxuron
Pendimethalin	Myclobutanil
Permethrin	Paclobutrazol
Phenanthrene	Pendimethalin
Phenothrin	Permethrin
Prometryn	Prallethrin
Propiconazole	Propiconazole
Propoxur	Propoxur
Pyrene	Pyrene
Pyridafenthion	Pyridafenthion
Tetramethrin	Tetramethrin
Thiabendazole	Triphenyl phosphate
Triphenyl phosphate	

Without standards to validate a number of these compounds by matching accurate mass and retention times, followed by tandem mass spectrometry (MS-MS)

fragmentation pattern matching, a different approach for identifying these molecules was undertaken. A structural and functional component of many pesticides is that they contain chlorine and bromine atoms which play a key role in their efficacy (Abhilash and Singh 2009). The presence of these elements in particular are very clear in a mass spectrometer due to their distinct isotope patterns (Siuzdak 2006; Gross 2011). To this end, each compound putatively identified in PSA treated and untreated wines that contained a chlorine or bromine was further investigated for their distinct isotope patterns.

Table 5.2. List of unique pesticides identified with and without PSA treatment.

PSA Treated	**Untreated**
Acephate	Atrazine
Anthracene	Bisphenol A
Cyprodinil	Etofenprox
Diflubenzuron	Fenpropathrin
Oryzalin	Isoprothiolane
Phenanthrene	Kresoxim-methyl
Phenothrin	Prallethrin
Prometryn	
Thiabendazole	

A mass spectrum of one compound which had this distinct pattern, and therefore was a more confident identification, is shown in Figure 5.4. For this compound (myclobutanil) an intense peak at 289.12 m/z was observed for the molecular ion which contains one ^{35}Cl atom. A less intense peak at 290.12 m/z is seen which correlates to the substitution of one ^{13}C in the compound and another peak is seen at 291.11 m/z correlating to the ^{37}Cl isotopic peak. This shift of two mass units is characteristic of a chlorine molecule when analyzed by mass spectrometry (Siuzdak 2006; Gross 2011). Furthermore, the predicted isotopic pattern (red bars) matches the observed spectrum

(blue bars) closely providing strong evidence that myclobutanil is present in the wine samples.

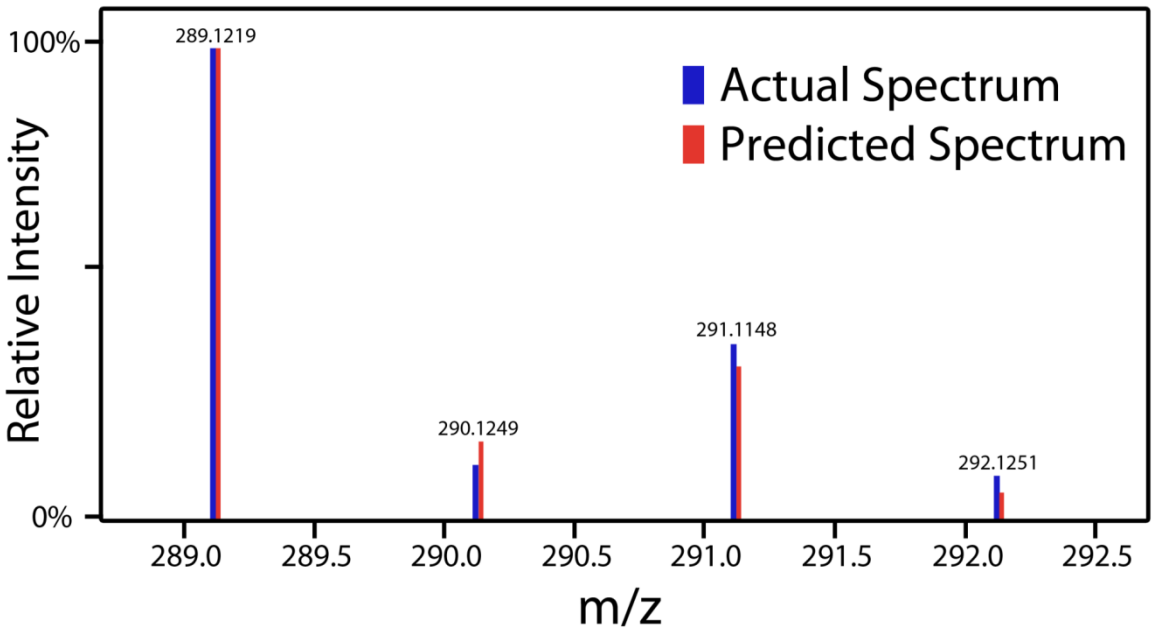

Figure 5.4. Comparison of myclobutanil actual vs predicted spectrum. Mass spectrum (blue bars) of the putatively identified pesticide myclobutanil and the predicted spectrum (red bars) generated from the chemical formula $C_{15}H_{18}ClN_4$. Intensity of peaks from the predicted spectrum match closely with that of the actual data, indicating a positive match for myclobutanil.

These results show the utility of the PSA to pull out pesticides from wine samples; these compounds were not found in untreated samples. To provide further confirmation of the previous results, a set of authentic pesticide standards was obtained to compare against the data from the wine samples. Authentic standards were obtained from the local agricultural and water testing facility at Montana State University, which regularly tests for trace contaminants in the ground water of rural areas in southwest Montana. A series of roughly 100 pesticides were analyzed by LC-MS using the same chromatographic methods for PSA treated samples and untreated samples. The goal of

this was to build a library of compounds which could be detected, establish their retention times, and the most common adduct form during ESI (M+H or M+Na where M = the monoisotopic mass, H = proton adduct, Na = sodium adduct).

Using this library of authentic standards, identification of pesticides from wine samples with and without PSA treatment could be accomplished with a higher confidence than the previous analysis. From this analysis, a total of four pesticides were confirmed to be present in the wine samples, three of which were unique to PSA treated samples, with the fourth compound present in both treated and untreated samples (Table 5.3). As with the previous analysis using only accurate mass more pesticides were identified in PSA treated samples than compared untreated samples, using shorter chromatographic runs. This means that not only can more pesticides be found using the PSA, but the time required to analyze each sample by LC-MS is less, a potential improvement on traditional approaches.

Table 5.1. List of pesticides identified with and without PSA treatment based on matches to authentic standards.

PSA Treated	**Untreated**
Aldicarb sulfoxide	Aldicarb sulfoxide
Dicrotophos	
Imazamethabenz acid	
Prometryn	

The result of fewer identified pesticides in this analysis compared with the initial study which had more putatively identified pesticides is not surprising. First, matching a compound based solely on accurate mass measurements within a set amount of error is more likely than matching accurate mass, retention time, and fragmentation pattern. This

is why authentic standards become so important in identifying a compound as they allow for comparison of three metrics. The second explanation as to why there are fewer identified compounds in the latter analysis using authentic standards relates to the origin of the library used. The library was built using authentic standards obtained from a local testing facility at Montana State University, where the primary clientele are residents of rural areas in Southwest Montana. This means that the library has pesticides which were biased towards pesticides used in Montana and the surrounding area. Not a region known for wine production. A deeper look at the compounds reveals pesticides such as propiconazole, a fungicide used on turf grass, and clothianidin, an insecticide used to treat perennial seeds before they are planted in the ground (Clark and Kenna 2010; Sheets 2010). Other examples of pesticides which would not be expected to be used by vineyards to grow grapes include hexaconazole, cyanazine, and siduron. In this respect, the library of pesticides was not ideal for the analysis of wine, but stilled proved to be of value in showing that more pesticides could be identified in PSA treated wine samples than compared to untreated samples. This also validates the results of the previous analysis which used pesticides actually known to be used in vineyards on grapes. Further experiments aimed at identifying pesticides in wine using the PSA will require a more tuned library of pesticides, such as those used in the initial study.

Conclusions

Building upon previously published results which showed the power of the PSA to differentiate wines based on varietal and time post uncorking, a selected series of

Bordeaux red wines were treated with the PSA, with the goal of classifying them in a logical and meaningful way to the consumer, with input from a wine sommelier. As was the case in the previous report, the PSA again classified each individual bottle based on the time post uncorking, as well as separated the bottles from one another. This experiment also shed light on the effect that breathing has on the wines as they became more similar to one another after 24 hours. These results could be useful in determining how long a given wine should be allowed to breathe before consuming, or conversely, how long until the wine spoils. A second use of the PSA in identifying organic pesticides from wine which was not previously reported was also explored in this work. Wine samples with and without treatment using the PSA were analyzed by LC-MS. PSA treated samples had more unique pesticides identified in them when compared to untreated samples. The PSA was previously shown to pull out low abundance molecules and this is likely the case in this experiment, the pesticides are at much lower concentrations compared to the bulk wine and are pulled out by the PSA. This work confirms the ability of the PSA to be used to classify wines based on varietal and time post uncorking and is extended to show the usefulness in identifying potential containments, such as pesticides, in wine when traditional analysis fails to detect them. Further experiments to identify the metabolites which bind to BSA that allow for the separation of the wine varietals is currently being undertaken with the goal being to generate unique flavor profiles for each. Additionally, confirmation of pesticides in the wine by authenticated standards using tandem mass spectrometry is also being done. The classification of wine varietals in a meaningful way or identification of pesticide

contamination in wine is of great interest to consumers and these experiments demonstrate the potential application of the PSA to do both.

References

Abhilash PC, Singh N (2009) Pesticide use and application: An Indian scenario. J Hazard Mater 165:1–12. doi: 10.1016/j.jhazmat.2008.10.061

Bisson LF, Waterhouse AL, Ebeler SE, et al (2002) The present and future of the international wine industry. Nature 418:696–699. doi: 10.1038/nature01018

Cabras P, Angioni A (2000) Pesticide residues in grapes, wine, and their processing products. J Agric Food Chem Chem 48:967–973. doi: 10.1021/jf990727a

Clark JM, Kenna MP (2010) Hayes' Handbook of Pesticide Toxicology. Elsevier

Dasgupta S, Banerjee K, Patil SH, et al (2010) Optimization of two-dimensional gas chromatography time-of-flight mass spectrometry for separation and estimation of the residues of 160 pesticides and 25 persistent organic pollutants in grape and wine. J Chromatogr A 1217:3881–3889. doi: 10.1016/j.chroma.2010.04.003

Garbi A, Sakkas V, Fiamegos YC, et al (2010) Sensitive determination of pesticides residues in wine samples with the aid of single-drop microextraction and response surface methodology. Talanta 82:1286–1291. doi: 10.1016/j.talanta.2010.06.046

Goto T, Ito Y, Yamada S, et al (2006) The high throughput analysis of N-methyl carbamate pesticides in fruits and vegetables by liquid chromatography electrospray ionization tandem mass spectrometry using a short column. Anal Chim Acta 555:225–232. doi: 10.1016/j.aca.2005.09.055

Gross JH (2011) Mass Spectrometry - A Textbook, 2nd edn. Springer

Hamerly T, Bothner B (2015) Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis. Appl Biochem Biotechnol 1–13.

Hamerly T, Heinemann J, Tokmina-Lukaszewska M, et al (2014) Bovine serum albumin as a molecular sensor for the discrimination of complex metabolite samples. Anal Chim Acta 818:61–66. doi: 10.1016/j.aca.2014.01.058

High J, Coppin C a (1988) Wiley and the Whiskey Industry: Strategic Behavior in the Passage of the Pure Food Act. Bus Hist Rev 62:286–309.

Khan PK, Sinha SP (1996) Ameliorating effect of vitamin C on murine sperm toxicity induced by three pesticides (endosulfan, phosphamidon and mancozeb). Mutagenesis 11:33–36. doi: 10.1093/mutage/11.1.33

Sannino A, Bolzoni L, Bandini M (2004) Application of liquid chromatography with electrospray tandem mass spectrometry to the determination of a new generation of pesticides in processed fruits and vegetables. J Chromatogr A 1036:161–169. doi: 10.1016/j.chroma.2004.02.078

Sheets LP (2010) Hayes' Handbook of Pesticide Toxicology. Elsevier

Siuzdak G (2006) The Expanding Role of Mass Spectrometry in Biotechnology, 2nd edn. MCC Press, San Diego, California

CHAPTER 6

INVESTIGATIONS INTO THE USE OF A PROTEIN SENSOR ASSAY FOR METABOLITE ANALYSIS

Contribution of Authors and Co-Authors

Manuscript in Chapter 6

Author: Timothy Hamerly

Contributions: Grew cell cultures, extracted metabolites, treated samples with PSA, acquired LC-MS data, analyzed data, generated figures, wrote manuscript.

Co-Author: Brian Bothner

Contributions: Provided insight and interpretation of results, edited manuscript.

Manuscript Information Page

Authors: Timothy Hamerly, Brian Bothner

Journal: Applied Biochemistry and Biotechnology

Status of Manuscript:
___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
__X_ Published in a peer-reviewed journal

Published by Springer

Issue in which manuscript appears: Online September 1-13 (2015)

INVESTIGATIONS INTO THE USE OF A PROTEIN SENSOR ASSAY FOR METABOLITE ANALYSIS

Timothy Hamerly,† and Brian Bothner*†

† Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717

*To whom correspondence should be addressed:
Brian Bothner
Department of Chemistry and Biochemistry
Montana State University
Bozeman, MT 59717
Phone: 406-994-5270
FAX: 406-994-4807
E-mail: bbothner@chemistry.montana.edu

Abstract

Rapid and definitive classification of biological samples has application in industrial, agricultural, and clinical settings. Considerable effort has been given to analytical methods to address such applications over the past 50 years, with the majority of successful solutions focusing on a single molecular target. However, in many cases a single or even a few features are insufficient for accurate characterization or classification. Serum albumin (SA) proteins are a class of cargo carrying proteins in blood that have evolved to transport a wide variety of metabolites and peptides in mammals. These proteins have up to seven binding sites which, communicate allosterically to, orchestrate a complex pick-up and delivery system involving a large number of different molecules at any time. The ability of SA proteins to bind multiple molecular species in a sophisticated manner, inspired development of assays to

differentiate complex biological solutions. The combination of SA and high resolution liquid chromatography mass spectrometry (LC-MS) is showing exciting promise as a protein sensor assay (PSA) for classification of complex biological samples. In this study, the PSA has been applied to cells undergoing and recovering from mild oxidative stress. Analysis using traditional LC-MS based metabolomics analysis failed to differentiate samples into treatment or temporal groups, whereas samples first treated with the PSA were cleanly classified into both correct treatment and temporal groups. The success of the PSA could be attributed to selective binding of metabolites, leading to a reduction in sample complexity and a general reduction in chemical noise. Metabolites important to successful sample classification were often enriched by 100 fold or more, yet displayed a wide-range of affinities for SA. The end result of PSA treatment is better classification of samples with a reduction in the number of features seen overall. Together, these results demonstrate how the use of a protein based assay before LC-MS analysis can greatly improve separation and lead to more accurate and successful tracking of metabolic state in an organism, suggesting potential application in a wide range of fields.

Introduction

A major challenge for biologists investigating states of disease or stress is that the variation between experimental replicates, termed biological replicates, is often similar in magnitude to the dependent variable. Omics experiments are particularly prone to this problem because data is collected for thousands of features (e.g. mRNA, proteins, metabolites), the vast majority of which are assumed to not be impacted by the

independent variable (Broadhurst and Kell 2006; Lay et al. 2006). A further complication for biologists is that it is often not possible to synchronize organisms. With omics experiments, the goal is often the discovery of a biomarker that can be used for sample differentiation, disease staging, or prognosis. The maturation of metabolomics as a discipline and the limited number of clinically useful biomarkers that have arisen from proteomics, has led to a gradual shift toward metabolomics for biomarker discovery (Kondo 2014). Current enthusiasm for metabolomics may be justified because metabolites are a direct readout of genetic and proteomic potential (Siuzdak 2014).

In a previous report, we showed that small molecules bound to the protein bovine serum album (BSA), could be used to differentiate wine varietals and correctly categorize animals into control and treatment groups in a clinically relevant model of hemorrhagic shock (Hamerly et al. 2014). These results were significant because biological replicates clustered more tightly and experimental groups were better separated when samples were treated using the BSA-based protein sensor assay (PSA) compared to direct analysis by liquid chromatography mass spectrometry (LC-MS) (Hamerly et al. 2014). The subset of molecules which bind to serum albumin (SA) generates a less complex profile that can be used to discriminate samples such as varietals of wine and urine (Hamerly et al. 2014). The choice of SA for this assay is based on the knowledge that these proteins bind a wide array of molecules with a range of affinities (Reynolds et al. 1968; Richieri et al. 1993; Peters Jr. 1996; Curry et al. 1998; Curry et al. 1999; Petitpas et al. 2001a; Zunszain et al. 2003; Fielding et al. 2005; Varshney et al. 2010; Nafisi et al. 2011). SA is found in the blood of all mammals and assists in transporting nonpolar molecules to muscle tissue and

organs (He and Carter 1992; Dockal et al. 2000; Varshney et al. 2010; Fanali et al. 2012). Structural models of albumins from several mammals reveal a set of distinct binding pockets that form the basis for selectivity; each one has a slightly different specificity while being allosterically connected (Reynolds et al. 1968; He and Carter 1992; Peters Jr. 1996; Curry et al. 1999; Petitpas et al. 2001b; Petitpas et al. 2001a; Varshney et al. 2010; Fanali et al. 2012). The working model is that SA binds a subset of all molecules present in solution based on each molecules affinity for a particular pocket. While the PSA was clearly effective, numerous questions arose about the usefulness of the assay for tracking subtle biological changes.

In this study, we expand upon the utility of the PSA by investigating the ability to differentiate microbial samples from the arachaeal organism *Sulfolobus solfataricus* during mild oxidative stress, gaining key insights to the mechanism of the PSA. The differentiation will be based on changes in the profile of intracellular metabolites. *S. solfataricus* is a crenarchaeal hyperthermal acidophile (Brock et al. 1972; Esser et al. 2011; Heinemann et al. 2014). Over the past decade, is has become a model for investigating the biology of extremophiles. Understanding how *S. solfataricus* copes with extreme conditions, particularly oxidative stress, has made important contributions to the understanding of intracellular reduction potential and the evolutionary relationship of antioxidant systems across the three domains of life (Woese 1987; Maaty et al. 2009; Heinemann et al. 2014). Oxidative stress has been well studied in this organism and the level of stress is easily manipulated, making it a good test for assessing sensitivity of the PSA (Buttke and Sandstrom 1994; Finkel and Holbrook 2000; Apel and Hirt 2004).

In addition to testing if the PSA can differentiate samples exposed to mild oxidative stress, we looked at the general properties of the molecules binding to SA. To start, a more in-depth analysis of the molecules binding to SA was undertaken. This revealed a novel set of features which show no classification strength in the raw samples. In other words, features which classify samples treated with the PSA, do not show significant change between time points of recovery. Binding studies showed that the off-rates for SA-ligand interaction were variable, but in general molecules having the most power for classification, increased in relative abundance after PSA treatment. Together, these results support a mechanism in which allosteric modulation of SA leads to selective binding and a powerful assay for differentiating biologically important samples.

Materials and Methods

Materials

Bovine serum albumin (greater than 98% agarose gel electrophoresis pure) was purchased from Sigma-Aldrich (St. Louis, MO). Molecular weight cutoff spin filters and syringe filters were purchased from Pall (Port Washington, NY). The buffer agent 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES) and Sodium Chloride were purchased from VWR (Radnor, PA) at a purity of 99% or greater. All solvents were purchased as HPLC grade; water from Avantor (Center Valley, PA), methanol and acetonitrile from EMD Chemicals Inc. (Gibbstown, NJ). Formic acid (98% GR ACS) for use as an LC ion pairing agent was purchased from EMD Chemicals Inc. (Gibbstown, NJ). Metabolite standards were purchased from Sigma-Aldrich (St. Louis, MO).

Cell Culturing and Oxidative Stress of *S. solfataricus*

S. solfataricus was grown in batch cultures at 80 C and pH 3.2, as previously described (Heinemann et al. 2014). Briefly, ATCC strain *S. solfataricus* P2 (SSP2) was cultured in DSMZ media 182 with the addition of 0.1% glucose to OD_{650} of 0.60. A control sample (0 minute) was collect by taking 50 mL of culture into a falcon tube and cell pellet collected by centrifugation. Following this sample, cells were stressed with H_2O_2 at a final concentration of 30 μM and cells were collected at 15, 30, 60, and 105 post stress. All cell pellets were stored at -80 °C until processing.

Metabolite Extraction from S. solfataricus Cells

Metabolites were extracted from cells as previously described (Heinemann et al. 2014). Briefly, metabolites were extracted from cells using sonication of cell pellets in 50% aqueous methanol (v/v). This was repeated three times and the extracts were pooled together. Proteins were precipitated using a 5:1 (v/v) of acetone by incubation at -80 °C for 12 hrs. Supernatant was divided into multiple vials and evaporated to dryness using a speed-vac. Samples were stored at -80 °C until analysis.

Preparation of BSA

A solution of BSA was prepared at 30 mg mL^{-1} and then washed with 10 mM HEPES buffer (pH 7.0) using 3kD molecular weight cutoff (MWCO) spin filters. This process was repeated three times to remove small molecules and contaminants remaining that could interfere with the assay and subsequent analysis by mass spectrometry. Following the wash step, BSA was re-suspended in 250 mM NaCl, 10 mM HEPES buffer

(pH 7.0) at a concentration of 30 mg mL^{-1} and frozen at -20 °C until analysis. Prior to analysis, the BSA stock was diluted using the same buffer to 0.1 mg mL^{-1}. MWCO spin filters were washed with deionized water prior to use to remove any preservatives as recommended by the manufacturer.

Treatment of Cell Extracts with PSA

Metabolite extracts from *S. solfataricus* cells were re-suspended in 50% aqueous (v/v) methanol (MeOH) (100 µL) and mixed with an equal volume of 0.1 mg mL^{-1} BSA. The sample was allowed to equilibrate for 5 mintues and then washed over a 3kD MWCO spin filter. Samples were centrifuged for 5 minutes, then washed with a solution of 10 mM HEPES buffer (pH 7.0, 250 mM NaCl) to remove nonspecific material. Samples were then washed and centrifuged three times with 70% MeOH to elute compounds retained by BSA. The supernant was combined and dried using a speed vac. Samples were stored at -80 °C until liquid chromatography mass spectrometry (LC-MS) analysis. Experimental blanks were conducted at the same time to control for protein, solvent, and filter contamination throughout the process.

Mass Spectrometry Analysis of PSA Treated Samples

For LC-MS analysis, samples treated with the PSA were re-suspended in 50% aqueous MeOH. For LC separation, a Kinetex 1.7 µm C18 150 mm x 2.1 mm column (Phenomenex, Torrance, CA) kept at 50 °C with a flow rate of 600 µL min^{-1} was used. The solvent system consisted of 0.1% formic acid in water and 0.1% formic acid in acetonitrile for solvent A and B respectively. An elution gradient of 2% B for 2 minutes,

to 95% B over 10 minutes, and then held at 95% B for 2 minutes, before finally returning to 2% B for 1 minute, for a total run time of 15 minutes. An Agilent 1290 UPLC (Agilent, Santa Clara, CA) system was used for LC separation, and was connected to an Agilent 6538 Q-TOF Mass Spectrometer (Agilent, Santa Clara, CA) for MS acquisition. The mass spectrometer operated in positive ion mode, with a cone voltage of 3500V and fragmentor voltage of 120V. Drying gas temperature was 350 °C with a flow of 12 L min^{-1} and the nebulizer was set to 60 psig. Spectra were collected at a rate of 2.5 Hz with a mass range of 50 to 1000 m/z.

Mass Spectrometry Analysis of Untreated Samples

Samples were handled as above for the PSA except that after drying, material was re-suspended in 50% aqueous MeOH. LC-MS analysis was as described above, except that the gradient for solvent B was extended; 2% to 95% B over 24 minutes giving a total run time of 30 minutes. This change was required to prevent major ion suppression due to the complexity of the untreated samples and allowed an unbiased comparison between the methods to be made. No changes were made to the mass spectrometer settings.

LC-MS Data Analysis

A pipeline for analysis of LCMS data, for both PSA treated and untreated, has been previously described (Hamerly et al. 2014). Briefly, molecular features from BSA treated and untreated samples were extracted using Masshunter Software (Agilent, Santa Clara, CA), and analyzed using standard XCMS protocols (Smith et al. 2006; Tautenhahn et al. 2008; R Core Team 2012). Following XCMS analysis, differential reports

(ANOVA) from each time point were imported into metaXCMS for alignment (Tautenhahn et al. 2011). The resulting aligned molecular feature lists were analyzed using principal component analysis (PCA) using XLStat and visualized using rgl, an R packages (Adler and Murdoch 2012; Addinsoft 2013). Data for Venn diagrams were annotated by hand, and visualized using the R package VennDiagram (Chen 2013).

Binding Studies

To compare the off-rates of molecules bound to SA, the PSA method was modified to include a variable length equilibration period during the wash step. Protein at a concentration of 0.01 mg mL^{-1} was mixed with a solution of metabolites from *S. solfataricus*. This mixture was washed over 3kD MWCO spin filters and the filters were centrifuged as described above. Following this, HEPES buffer was added to the filters and the solution was allowed to equilibrate for 5 minutes (standard PSA time), 3, 8, or 24 hours, before centrifugation. Metabolites remaining bound to BSA were eluted with methanol and analyzed by LC-MS (as above). Experiments were repeated a minimum of three times. Data analysis to determine percent bound used ion intensity by the same method as outlined for analysis of LC-MS data above.

Results

Analysis of Oxidative Stress

To determine if the PSA was sensitive to subtle differences in cellular state, *S. solfataricus* cells recovering from mild oxidative stress over the course of 105 minutes were analyzed. Batch cultures of cells, at 80 ^{o}C, pH 3.2, were spiked with hydrogen

peroxide (H_2O_2) to a final concentration of 30 μM. This low dose does not significantly slow growth, but has been shown to induce a cellular response at the level of mRNA and protein (Maaty et al. 2009). Cells were collected at five time points to track the metabolic response to hydrogen peroxide (T=0, 15, 30, 60, and 105 minutes). Intracellular metabolites were analyzed using two different methods. Samples were split into aliquots with one analyzed directly using standard reverse-phase LC-MS. A second aliquot was treated using the PSA before LC-MS analysis. Representative total ion chromatograms (TICs) from two biological replicates 30 minutes after the addition of hydrogen peroxide are shown in Figure 6.1. The metabolite fractions analyzed directly by LC-MS were dominated by early eluting (more polar) compounds (Figure 6.1A). Small differences between biological replicates were visible (grey and black TIC traces). LC-MS analysis of the same two samples after PSA treatment revealed a shift in the elution position of the small molecules (Figure 6.1B). The treated samples were significantly enriched in less polar molecules. The average number of molecular features was reduced, going from 850 to approximately 300 after PSA treatment.

Data from each time series were analyzed using principal component analysis (PCA). The complex data sets resulting from LC-MS are well suited for PCA because it reduces dimensionality providing a straight-forward, unsupervised visualization of sample variance.(Worley and Powers 2013) A plot of the samples analyzed directly by LC-MS showed that the biological replicates fail to cluster in a consistent manner (Figure 6.2A). However, if the same samples were PSA treated before LC-MS analysis, biological replicates clustered tightly (Figure 6.2B). The data also segregated temporally

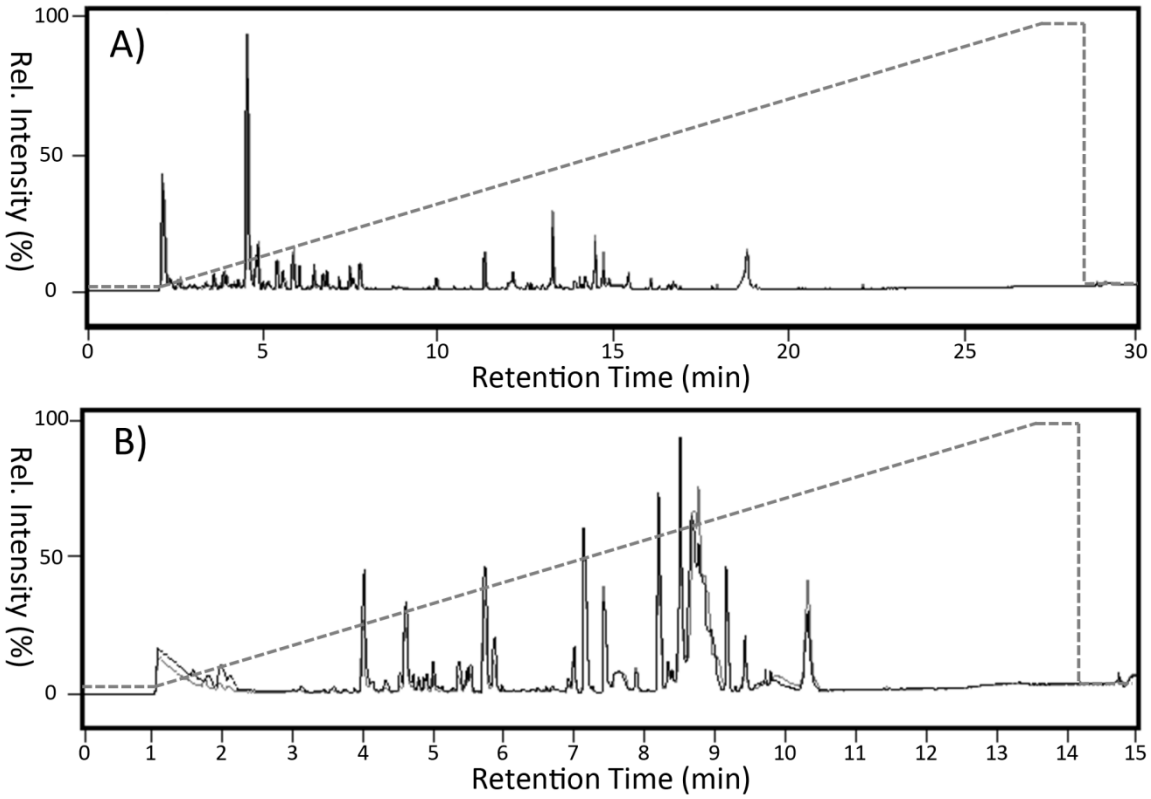

Figure 6.1. LC-MS chromatograms of metabolite extracts of *S. solfataricus* cells. A) TIC of reverse phase LC-MS for two biological replicates of metabolite extracts 30 mins after addition of 30 μM hydrogen peroxide. B) The same two biological samples treated with the PSA before LC-MS analysis. Biological replicates are similar in both analyses, however, the reverse-phase elution profile is different. Dashed line indicates the % acetonitrile across the chromatogram.

for PSA treated samples. Individual time points throughout the time course can be discriminated from one another when treated with the PSA, whereas this was not possible for the direct analysis using only LC-MS. The one exception was the control (0 minutes) and 15 minutes post stress samples which were overlapped, presumably because the metabolome had yet to change significantly.

The ability to correctly categorize samples using LC-MS after PSA treatment when LC-MS alone did not, prompted us to investigate the mechanism behind PSA

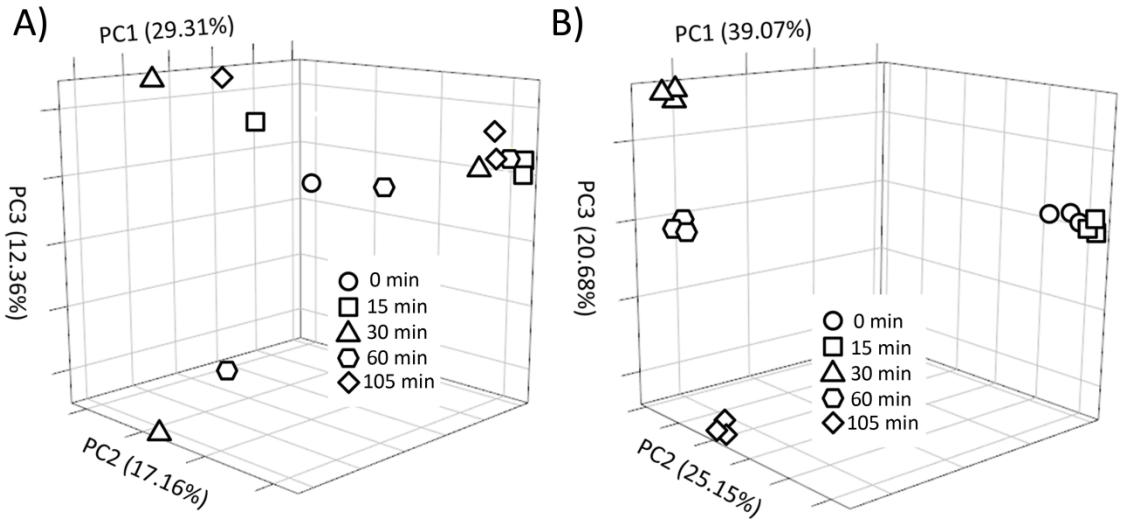

Figure 6.2. Principal component analysis of *S. solfataricus* cells undergoing oxidative stress. A) PCA plot of LC-MS data from metabolite samples directly after extraction from cells. B) The same samples after treatment with the PSA, followed by LC-MS analysis. Three biological replicates were used at each time point. Color key indicates time after induction of mild oxidative stress.

enhanced differentiation. To begin, the LC-MS data was analyzed to find the specific markers that differentiate later time points from control (0 minutes). The time point 105 minutes after induction of stress was compared with the control sample, to look for markers of discrimination and to compare these markers with samples not treated with the PSA. The 105 minute samples were selected because the system was perturbed, but based on PC3 showed signs of recovery (see Figure 6.2B and previously published work from our lab) (Maaty et al. 2009). To assess which molecular features where responsible for differentiating the sample groups, an ANOVA was used to identify features which had a fold change of greater than 1.5 and a p-value less than 0.05. Only three features met the criteria, even though the samples showed a clear separation by PCA. Examination of the loading plot showed that a number of features were distributed on the perimeter. This

indicated that the distinction between sample groups relied on a group of features that, individually, were only subtly changed. A number of features in close proximity to those with significant difference were also selected for further investigation because, while not significant in the ANOVA analysis, they contributed to the clustering by PCA.

Table 6.1. Molecular features of interest

		PSA Treated		**Untreated**		
Number	**m/z**	**Fold**	**p-value**	**Fold**	**p-value**	**METLIN**
1	**229.14**	**9.13**	**0.002**	**1.00**	**0.777**	Fatty Acid
2	**611.35**	**2.38**	**0.015**	**1.01**	**0.752**	Tetra-Peptide
3	186.13	1.06	0.318	1.09	0.685	
4	620.43	1.37	0.149	1.09	0.251	PE DAG
5	407.20	1.15	0.142	1.48	0.027	Tetra-Peptide
6	145.10	1.11	0.138	1.13	0.659	
7	**226.10**	**1.52**	**0.021**	**1.20**	**0.260**	
8	477.30	1.41	0.240	1.14	0.523	Tetra-Peptide
9	324.15	1.26	0.439	1.03	0.950	Tri-Peptide

PE DAG = phosphatidyl ethanolamine diacylglycerol

In total, nine features were selected for further analysis. Table 6.1 shows the fold change and p-value of these molecules. Also shown in Table 1 are the fold changes and p-values of these features from the untreated samples. Features in bold type are those with significant change (fold greater than 1.5 and p-value less than 0.05) in the PSA groups. As can be seen, the fold changes for each feature were greater when first treated with the PSA, and the p-values were more significant than for untreated samples. The one exception was feature number 5 (407.20 m/z) which had a higher fold change and lower p-value without PSA treatment. A number of the features in Table 6.1 were putatively identified using the METLIN database as long chain fatty acids or peptides. Both of these

molecular classes are among those expected to be selected by BSA (Spector et al. 1969; Lowenthal et al. 2005; Jupin et al. 2013; Jupin et al. 2014).

To further investigate the mechanism behind the PSA improved differentiation, the data was analyzed to show the consistency of features across the five time points. A Venn diagram analysis emphasized that fewer features were present after PSA. Also, there was a lower relative number of unique features and a higher relative number shared at each time point in the treated samples (Figure 6.3A). Histograms plots of features present in 3 or more, 4 or more, and all 5 time points showed that a greater percentage of features were common between multiple time points when the PSA was used (Figure 6.3B). In each case, the enrichment of consistently detected features was statistically significant in PSA treated samples, as ascertained by Z-test scores of greater than 3.3 (p-value of less than 0.001). We reasoned that features present in all 5 time points would have a large influence on temporal separation of the samples. To see if there was a specific trend in abundance of these features, a plot of relative abundance over time was constructed. Four distinct patterns were observed. Representative features which demonstrate the abundance trends were plotted (Figure 6.3C). This included metabolites which trended steadily up (diamond shapes), metabolites which trended steadily down (triangle shapes), metabolites which first increased in intensity, before coming back down (square shapes), and metabolites which initially decreased in intensity, before coming back up at the end of the time course (circle shapes).

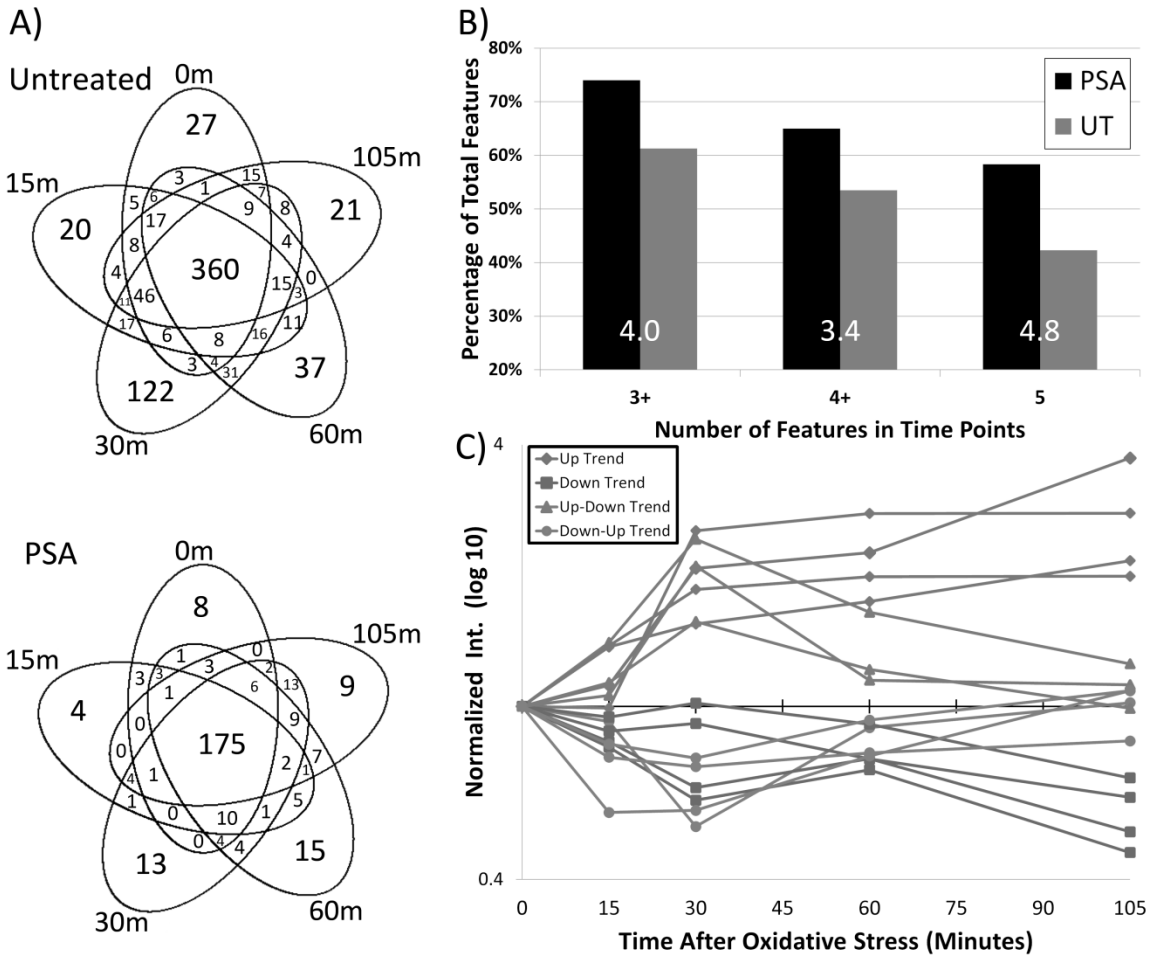

Figure 6.3. Analysis showing the consistency of features throughout a time course after PSA treatment. A) Venn diagram showing the distribution of features when samples were analyzed directly by LC-MS (top) and when treated with the PSA prior to LC-MS (bottom). B) Histogram showing the percentage of metabolites in 3 or more, 4 or more, and all 5 time points before and after PSA treatment. The z-score, shown for each pair, indicates a significant difference between untreated and treated metabolites (p-value less than 0.01). C) Four basic abundance trends were present in features detected in all five time points after PSA treatment. Metabolites that trend up (blue), metabolites that trend down (red), metabolites which trend up at first, then come back down (green), and metabolites which trend down at first, and then come back up (purple).

Binding Studies

Based on our results showing an enrichment (increased intensity) of metabolites after treatment with the PSA, a binding study was conducted to investigate the off-rates

of metabolites bound to BSA. Also of interest was the relative contribution to sample composition with and without PSA treatment. It is well known that allostery plays a role in molecular affinity of SA (Curry et al. 1999; Petitpas et al. 2001b; Petitpas et al. 2001a; Zunszain et al. 2003). This means that the affinity of BSA for a specific molecule can be influenced by other molecules in the solution. In this experiment, metabolites from *S. solfataricus* were mixed with BSA following the protocol described above for the PSA, except that after binding and washing, the solution of BSA and metabolites was left to equilibrate at room temperature for 3, 8, or 24 hours in a HEPES buffer solution. Samples were then analyzed by LC-MS at each equilibration time to determine the fraction bound for each metabolite. Data for nine molecules were selected to show the range of observed trends (Figure 6.4). The fraction bound was calculated by dividing the ion intensity at each equilibration period by the intensity after five minutes (the equilibration period used in the PSA). The curves show that the bound molecules had a range of off-rates from BSA, with some showing very little loss in intensity (slow off-rate) with others showing almost complete disassociation at the first time point (fast off-rate). In other words, the time to equilibrium varied for the highlighted features with some showing faster off rates than others.

The large difference in observed off-rates suggested that molecules with a range of affinities for BSA were being selected in the PSA. This raised a question about the concentration of a particular species in the sample. To address this, the intensities of each metabolite without PSA treatment and with PSA treatment after 8 hours of equilibration were compared. While most of the features in Table 6.2 are more intense after being

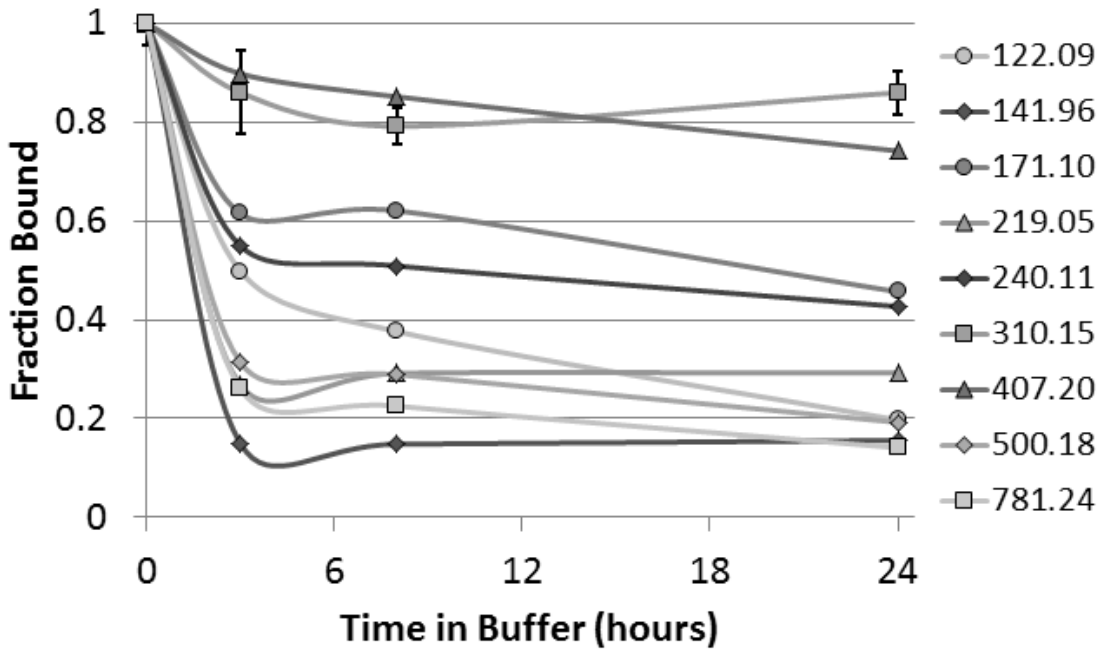

Figure 6.4. Disassociation curves of BSA bound metabolites. Curves showing fraction bound to BSA of nine different metabolites over a 24 hour period of time. Metabolite extracts from *S. solfataricus* were treated using the PSA and left to equilibrate before standard elution with methanol to remove bound metabolites. Fraction bound was calculated based on LCMS ion intensity after equilibration divided compared to the standard 5 minute period used in the PSA. Lines connecting data points were added to improve visualization.

treated with the PSA, there were multiple features (122.09, 141.96, 240.11, 500.18, and 781.24) nearly absent without treatment using the PSA. These features increase in intensity by as much as 100 fold when treated with the PSA, showing the enhancement of features based on affinity to BSA. The intensity of these features with and without PSA treatment was compared by ranking them on percentile, where the most intense ion was given a rank of 1 and a percentile value of 99% (Table 6.2). Thus, the larger the rank value, smaller percentile value, indicates a feature which is not very abundant. By using a percentile rank, it is possible to compare specific ions between untreated and PSA treated

Table 6.2. Intensity and rank for select molecular features before and after PSA treatment.

m/z	Ion Intensity		Rank #		Percentile #	
	Untreated	PSA	Untreated	PSA	Untreated	PSA
122.09	57980	66314	771/845	213/240	4.56%	11.61%
141.96	58299	618234	769/845	46/240	4.97%	80.91%
171.10	442902	156162	227/845	136/240	61.41%	43.56%
219.05	204666	3383842	416/845	14/240	32.78%	94.19%
240.11	49510	861739	794/845	38/240	3.31%	84.23%
310.15	125195	541025	548/845	53/240	20.22%	78.00%
407.20	184743	283405	443/845	91/240	29.44%	62.24%
500.18	40287	315432	810/845	87/240	2.25%	63.90%
781.24	64705	52160	755/845	227/240	5.37%	5.80%

most intense ion was given a rank of 1 and a percentile value of 99%

samples, where simply looking at the rank within a dataset may be misleading due to the difference in total ions detected. While some ions do not change in percentile or intensity before and after PSA treatment, many ions in Table 6.2 drastically increase in intensity, and in percentile rank, where these ions may be important for biological differentiation.

Discussion

The complexity and variability of biological samples is a constant challenge for analytical biochemists. To test the usefulness of a new analytical method, we examined the metabolic differences caused by mild oxidative stress. We chose to investigate *S. solfataricus* because, as an extremophile, it is markedly different than the other systems tested using the PSA and the metabolic changes associated with mild oxidative stress had not been previously investigated (Heinemann et al. 2014). As a comparison, a standard non-targeted LC-MS based metabolomics approach was conducted.

The TICs from LC-MS analysis show that biological replicates were visually similar, both with and without PSA treatment (Figure 6.1). However, after treatment with the PSA, the chromatographic profile changed significantly. In general, metabolites in the PSA sample eluted at a higher concentration of acetonitrile, indicating that they were less polar in nature. This shift was expected due to the hydrophobic nature of the pockets on SA and the known binding preferences (Varshney et al. 2010). Additionally, the number of molecular features was reduced by nearly two-thirds, from 850 to 300 after PSA treatment. A result that is also consistent with selective binding by SA proteins.

The 30 μM hydrogen peroxide treatment is referred to as "mild" because it does not inhibit growth of *S. solfataricus* (Maaty et al. 2009). The mildness of this treatment was borne out when PCA of samples analyzed using standard LC-MS based metabolomics methods showed that replicates and time points are intermixed and cannot be differentiated (Figure 6.2A). This is in sharp contrast to samples which are first treated with the PSA. After PSA treatment, not only were replicates well grouped, but time points 30, 60, and 105 minutes were clearly separated, and a logical trajectory appeared (Figure 6.2B). PC1 accounted for the movement of these 3 time points away from control and 15 minute groups. Metabolic recovery is also evident because at the final time point (105 minutes) PC3 no longer contributes to separation from control. Another encouraging sign for PSA performance is the total contribution to variation from the first three principal components. When using the PSA, PC1-3 account for 85% of the separation, compared to 58% in the untreated data.

Based on the results showing PSA treated samples are more readily classified than the samples without, two experiments to investigate the mechanism responsible for PSA enhanced resolution were conducted. The first focused on the difference between control (0 minutes) and 105 minutes post stress. Before PSA treatment, no features had a statistically significant difference between these time points based on ANOVA. Post PSA, features in general have increased fold changes and improved significance values, as shown by the representative features in Table 6.1. PSA functionality arises from the ability to select and concentrate features, resulting in an amplification of weak trends that leads to separation of control and stress groups when otherwise it was not possible.

A second experiment to investigate the mechanism behind the PSA assay tracked the consistency of features across multiple time points. Nearly 60% of features are seen in all five time points in PSA treated samples, compared with roughly 40% in the untreated (Figure 6.3B). The increased consistency of metabolites means that trends exist for a greater percentage of metabolites and potential markers for stress can be elucidated more effectively (Figure 6.3C). Conversely, fewer metabolites are present in only one or two time points after PSA treatment, resulting in a higher frequency of metabolites which can be tracked throughout the entire time course (Figure 6.3A). It should be noted that trends in observed ion signals after PSA do not necessarily report accurately on the concentration of a metabolite in the original sample. This is because on top of binding to SA being competitive, the process is allosterically regulated.

The dynamics of ligand-SA protein affinity and ultimate selection of metabolites shows up in our results. Off-rates for metabolites selected by BSA can be very different

(Figure 6.4). This agrees with previous studies showing that SAs have a number of binding pockets with a range of affinities (He and Carter 1992). Metabolites which are low in concentration, but have an affinity for BSA are greatly enriched. For example, the intensity of features 141.96, 219.05, 240.11, 310.15, 407.20, and 500.18 amu increased by as much as 100 fold in the PSA LC-MS data (Table 6.2). These ions are among the most intense after PSA treatment, while in the untreated samples, they tend to have a much lower percentile ranking (Table 6.2). Many of these ions would be easily missed in a standard LC-MS based differential metabolomics analysis. The ion enrichment also improves the likelihood of identifying these potential biomarkers, because after PSA there is sufficient intensity for fragmentation based analysis using MS/MS. The observed off-rates indicate that metabolites with markedly different affinities for SA are important for discriminating samples. This is very different from how standard affinity chromatography works and may be a key factor in the success of the PSA.

Conclusion

The assay used herein, based on the natural affinity of serum album proteins for small molecules, shows great promise as a technique for fingerprinting complex biological fluids and biomarker discovery. Investigations into the nature of the molecules and their binding to BSA that facilitated the enhanced separation revealed two key details; the molecules had a range of affinities for BSA and were often present in low abundance. The PSA works because BSA is able to select molecules of low abundance despite the presence of other molecules that change abundance between samples. Put

another way, the PSA works by filtering chemical noise and integrating molecular concentrations into useful information.

Acknowledgements

The authors would like to thank Dr. Walid Maaty for advice on cell culture and stressing of *S. solfataricus* with hydrogen peroxide. The authors would also like to thank Dr. Jonathan Hilmer, for his tireless efforts and mass spectrometry expertise. Tim Hamerly would like to recognize support from the Kopriva Graduate Fellowship at MSU. This work was supported by the National Science Foundation, MCB102248 through an award to Brian Bothner. The Proteomics, Metabolomics, and Mass Spectrometry facility at MSU received support from the Murdock Charitable Trust and NIH 5P20RR02437 of the CoBRE program.

References

Addinsoft (2013) XLSTAT.

Adler D, Murdoch D (2012) rgl: 3D visualization device system (OpenGL).

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399. doi: 10.1146/annurev.arplant.55.031903.141701

Broadhurst DI, Kell DB (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics 2:171–196. doi: 10.1007/s11306-006-0037-z

Brock TD, Brock KM, Belly RT, Weiss RL (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84:54–68.

Buttke TM, Sandstrom PA (1994) Oxidative stress as a mediator of apoptosis. Immunol Today 15:7–10.

Chen H (2013) VennDiagram: Generate high-resolution Venn and Euler plots.

Curry S, Brick P, Franks NP (1999) Fatty acid binding to human serum albumin: new insights from crystallographic studies. Biochim Biophys Acta 1441:131–40.

Curry S, Mandelkow H, Brick P, Franks N (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat Struct Biol 5:827–35. doi: 10.1038/1869

Dockal M, Carter DC, Rüker F (2000) Conformational Transitions of the Three Recombinant Domains of Human Serum Albumin Depending on pH. J Biol Chem 275:3042–3050. doi: 10.1074/jbc.275.5.3042

Esser D, Kouril T, Zaparty M, et al (2011) Functional curation of the Sulfolobus solfataricus P2 and S. acidocaldarius 98-3 complete genome sequences. Extremophiles 15:711–2. doi: 10.1007/s00792-011-0392-1

Fanali G, di Masi A, Trezza V, et al (2012) Human serum albumin: from bench to bedside. Mol Aspects Med 33:209–90. doi: 10.1016/j.mam.2011.12.002

Fielding L, Rutherford S, Fletcher D (2005) Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems. Magn Reson Chem 43:463–70. doi: 10.1002/mrc.1574

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. Nature 408:239–247.

Hamerly T, Heinemann J, Tokmina-Lukaszewska M, et al (2014) Bovine serum albumin as a molecular sensor for the discrimination of complex metabolite samples. Anal Chim Acta 818:61–66. doi: 10.1016/j.aca.2014.01.058

He XM, Carter DC (1992) Atomic Structure and Chemistry of Human Serum Albumin. Nature 358:209–215.

Heinemann J, Hamerly T, Maaty WS, et al (2014) Expanding the paradigm of thiol redox in the thermophilic root of life. Biochim Biophys Acta 1840:80–85. doi: 10.1016/j.bbagen.2013.08.009

Jupin M, Michiels PJ, Girard FC, et al (2014) NMR metabolomics profiling of blood plasma mimics shows that medium- and long-chain fatty acids differently release metabolites from human serum albumin. J Magn Reson 239:34–43. doi: 10.1016/j.jmr.2013.11.019

Jupin M, Michiels PJ, Girard FC, et al (2013) NMR identification of endogenous metabolites interacting with fatted and non-fatted human serum albumin in blood plasma: Fatty acids influence the HSA-metabolite interaction. J Magn Reson 228:81–94. doi: 10.1016/j.jmr.2012.12.010

Kondo T (2014) Inconvenient truth: Cancer biomarker development by using proteomics. Biochim Biophys Acta 1844:861–865. doi: 10.1016/j.bbapap.2013.07.009

Lay JO, Liyanage R, Borgmann S, Wilkins CL (2006) Problems with the "omics." Trends Anal Chem 25:1046–1056. doi: 10.1016/j.trac.2006.10.007

Lowenthal MS, Mehta AI, Frogale K, et al (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem 51:1933–45. doi: 10.1373/clinchem.2005.052944

Maaty WS, Wiedenheft B, Tarlykov P, et al (2009) Something old, something new, something borrowed; how the thermoacidophilic archaeon Sulfolobus solfataricus responds to oxidative stress. PLoS One 4:e6964. doi: 10.1371/journal.pone.0006964

Nafisi S, Bagheri Sadeghi G, PanahYab A (2011) Interaction of aspirin and vitamin C with bovine serum albumin. J Photochem Photobiol B Biol 105:198–202. doi: 10.1016/j.jphotobiol.2011.09.002

Peters Jr. T (1996) All About Albumin. Academic Press, San Diego, California

Petitpas I, Bhattacharya AA, Twine S, et al (2001a) Crystal structure analysis of warfarin binding to human serum albumin: anatomy of drug site I. J Biol Chem 276:22804–9. doi: 10.1074/jbc.M100575200

Petitpas I, Gru T, Bhattacharya AA, Curry S (2001b) Crystal Structures of Human Serum Albumin Complexed with Monounsaturated and Polyunsaturated Fatty Acids. J Mol Biol 314:955–960. doi: 10.1006/jmbi.2001.5208

R Core Team (2012) R: A language and environment for statistical computing.

Reynolds J, Herbert S, Steinhardt J (1968) The Binding of Some Long-Chain Fatty Acid Anions and Alcohols by Bovine Serum Albumin. Biochemistry 7:1357–1361.

Richieri G V, Anel a, Kleinfeld a M (1993) Interactions of long-chain fatty acids and albumin: determination of free fatty acid levels using the fluorescent probe ADIFAB. Biochemistry 32:7574–80.

Siuzdak G (2014) Mass spectrometry-based metabolomics as a unique biochemical approach for understanding disease pathogenesis. FASEB J 28:350.3.

Smith CA, Want EJ, O'Maille G, et al (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Anal Chem 78:779–787.

Spector AA, John K, Fletcher JE (1969) Binding of long-chain fatty acids to bovine serum albumin. J Lipid Res 10:56–67.

Tautenhahn R, Boettcher C, Neumann S (2008) Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504.

Tautenhahn R, Patti GJ, Kalisiak E, et al (2011) metaXCMS: Second-Order Analysis of Untargeted Metabolomics Data. Anal Chem 83:696–700.

Varshney A, Sen P, Ahmad E, et al (2010) Review Article Ligand Binding Strategies of Human Serum Albumin : How Can the Cargo be Utilized ? Chirality 22:77–87. doi: 10.1002/chir

Woese CR (1987) Bacterial Evolution Background. Microbiol Rev 51:221–271.

Worley B, Powers R (2013) Multivariate Analysis in Metabolomics. Curr Metabolomics 1:92–107.

Zunszain P a, Ghuman J, Komatsu T, et al (2003) Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. BMC Struct Biol 3:1–9. doi: 10.1186/1472-6807-3-6

CHAPTER 7

CONCLUDING REMARKS

The results presented throughout this dissertation highlight the potential applications of a protein based sensor assay capable of classifying numerous complex biological samples. The wealth of knowledge which can be gleaned from the metabolome of biological systems has become of great interest in the field of systems biology, however, the complexity of these systems often times makes it difficult to make meaningful progress towards understanding them. Techniques such as LC-MS have made it possible to detect 1000's of metabolites from a cell, but the markers capable of differentiating a healthy and stressed state are often times lost as most of these metabolites do not change. A method which is capable of reducing the complexity while maintaining biological meaning therefore offers a great advantage when studying the metabolome. Using the protein serum albumin an assay has been successfully developed which is able to rapidly and easily classify complex biological samples.

The protein serum albumin was chosen for this assay due to its ability to bind a wide range of small molecules with some specificity. Serum albumin proteins *in vivo* bind fatty acids, drugs, and metal ions and transport them throughout the body. Binding to the protein is combination of affinity to a given pocket and allostery meaning that when the protein is at low concentrations compared to metabolites in solution, a competition to bind exists. The properties that allow for this binding make the protein an excellent candidate to build an assay upon. Additionally, serum albumin can be easily

extracted from the plasma of mammalians in high purity meaning the price when bought from vendors is relatively inexpensive and is easily accessible, both of which are key goals in developing the assay.

Early success of the PSA for the pretreatment of complex biological samples which reduced complexity while giving better classification than traditional methods has resulted in the assay being implemented in numerous other applications. Initially the PSA was applied to a series of wines which could be differentiated from one other regardless of PSA treatment but could only be differentiated based on time post uncorking after PSA treatment. This result was repeated in other samples of wine and whiskey where again the PSA excelled at differentiating subtle changes in the metabolome of these alcoholic beverages. Following this success the PSA was tested on a number samples which originated from medically relevant backgrounds. Cells recovering from oxidative stress over time were shown to be indiscernible with traditional methods. After PSA treatment the cells could be readably differentiated at each time point. From this data it was found that the PSA worked by pulling out both high and low abundance molecules from the metabolome in such a way as to allow the differentiation. Furthermore, these metabolites had a range of off rates from the protein likely due to a combination of allosteric regulation of the binding pockets and the chemical nature of the pockets themselves. This important discovery has resulted in a better understanding of how the PSA functions. The result of pulling out low abundance molecules was also observed when investigating pesticide contamination in wines, where a number of pesticides could only be seen in PSA treated samples. In a large animal model for hemorrhagic shock the PSA again

performed better than traditional methods, where only after PSA treatment could animals recovering from shock be classified separately from control animals. In this experiment a number of features were putatively found to be significantly different between control and shock animals. The PSA has already shown to be applicable with a diverse number of samples and is likely suited for more sample types than were explored in this dissertation.

The experiments and results described in this dissertation show that a protein based sensor assay has a myriad of applications; from the classification of wines to the use in biomedical research. The PSA has been shown to be of use in two agriculturally relevant applications and raises the question of where else it might be applied in the food industry. The use of the PSA to track milk spoilage has been explored and shown to be a promising new utilization in this area. In medically relevant fields the PSA could be of potential use for classifying disease states in a rapid manor by coupling it to an enzyme linked immunosorbent assay (ELISA). The PSA could be used to pull out markers of a given disease state which would then cause a color change in the ELISA. The PSA could also be modified to use the endogenously bound metabolites of an organisms own SA to look for disease states. Early experiments have already shown that this is more than just a possibility and could mark an entirely new field of exploration for the PSA. The PSA has been shown to be a powerful analytical tool for the simplification of complex biological samples.

REFERENCES CITED

Abhilash PC, Singh N (2009) Pesticide use and application: An Indian scenario. J Hazard Mater 165:1–12. doi: 10.1016/j.jhazmat.2008.10.061

Adams MM, Anslyn E V (2009) Differential sensing using proteins: exploiting the cross-reactivity of serum albumin to pattern individual terpenes and terpenes in perfume. J Am Chem Soc 131:17068–17069. doi: 10.1021/ja908319m

Addinsoft (2013) XLSTAT.

Adler D, Murdoch D (2012) rgl: 3D visualization device system (OpenGL).

Anslyn E V (2007) Supramolecular analytical chemistry. J Org Chem 72:687–99. doi: 10.1021/jo0617971

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol 55:373–399. doi: 10.1146/annurev.arplant.55.031903.141701

Bhattacharya B, Nakka S, Guruprasad L, Samanta A (2009) Interaction of bovine serum albumin with dipolar molecules: fluorescence and molecular docking studies. J Phys Chem B 113:2143–2150. doi: 10.1021/jp808611b

Bier M, Bakker BM, Westerhoff H V (2000) How yeast cells synchronize their glycolytic oscillations: a perturbation analytic treatment. Biophys J 78:1087–93. doi: 10.1016/S0006-3495(00)76667-7

Bisson LF, Waterhouse AL, Ebeler SE, et al (2002) The present and future of the international wine industry. Nature 418:696–699. doi: 10.1038/nature01018

Boulanger L, Joshi A V., Tortella BJ, et al (2007) Excess Mortality, Length of Stay, and Costs Associated with Serious Hemorrhage among Trauma Patients: Findings from the National Trauma Data Bank. Am Surg 73:1269–1274.

Braha O, Walker B, Cheley S, et al (1997) Designed protein pores as components for biosensors. Chem Biol 4:497–505. doi: 10.1016/S1074-5521(97)90321-5

Broadhurst DI, Kell DB (2006) Statistical strategies for avoiding false discoveries in metabolomics and related experiments. Metabolomics 2:171–196. doi: 10.1007/s11306-006-0037-z

Brock TD, Brock KM, Belly RT, Weiss RL (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84:54–68.

Bujacz A (2012) Structures of bovine, equine and leporine serum albumin. Acta Crystallogr Sect D Biol Crystallogr 68:1278–1289. doi: 10.1107/S0907444912027047

Burns E (2012) Bad Whisky. In: 3rd ed. Neil Wilson Publishing, pp 33–34

Buttke TM, Sandstrom PA (1994) Oxidative stress as a mediator of apoptosis. Immunol Today 15:7–10.

Cabras P, Angioni A (2000) Pesticide residues in grapes, wine, and their processing products. J Agric Food Chem Chem 48:967–973. doi: 10.1021/jf990727a

Chen H (2013) VennDiagram: Generate high-resolution Venn and Euler plots.

Cherkas D (2011) Traumatic hemorrhagic shock: advances in fluid management. Emerg Med Pract 13:1–19.

Clark JM, Kenna MP (2010) Hayes' Handbook of Pesticide Toxicology. Elsevier

Cuadros-Inostroza A, Giavalisco P, Hummel J, et al (2010) Discrimination of wine attributes by metabolome analysis. Anal Chem 82:3573–3580. doi: 10.1021/ac902678t

Cuperlovic-culf M, Culf AS, Morin PJ, Touaibia M (2013) Application of Metabolomics in Drug Discovery , Development and Theranostics. Curr Metabolomics 1:41–57.

Curry S, Brick P, Franks NP (1999) Fatty acid binding to human serum albumin: new insights from crystallographic studies. Biochim Biophys Acta 1441:131–40.

Curry S, Mandelkow H, Brick P, Franks N (1998) Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. Nat Struct Biol 5:827–35. doi: 10.1038/1869

Curvale R (2009) Buffer capacity of Bovine serum albumin.(BSA). J Argent Chem Soc 97:174–180.

Dasgupta S, Banerjee K, Patil SH, et al (2010) Optimization of two-dimensional gas chromatography time-of-flight mass spectrometry for separation and estimation of the residues of 160 pesticides and 25 persistent organic pollutants in grape and wine. J Chromatogr A 1217:3881–3889. doi: 10.1016/j.chroma.2010.04.003

Davidovic L, Navratil V, Bonaccorso CM, et al (2011) A metabolomic and systems biology perspective on the brain of the Fragile X syndrome mouse model. Genome Res 21:2190–2202. doi: 10.1101/gr.116764.110

Day WHE, Edelsbrunner H (1984) Efficient algorithms for agglomerative hierarchical clustering methods. J Classif 1:7–24. doi: 10.1007/bf01890115

Dockal M, Carter DC, Rüker F (2000) Conformational Transitions of the Three Recombinant Domains of Human Serum Albumin Depending on pH. J Biol Chem 275:3042–3050. doi: 10.1074/jbc.275.5.3042

Esser D, Kouril T, Zaparty M, et al (2011) Functional curation of the Sulfolobus solfataricus P2 and S. acidocaldarius 98-3 complete genome sequences. Extremophiles 15:711–2. doi: 10.1007/s00792-011-0392-1

Fanali G, di Masi A, Trezza V, et al (2012) Human serum albumin: from bench to bedside. Mol Aspects Med 33:209–90. doi: 10.1016/j.mam.2011.12.002

Fielding L, Rutherford S, Fletcher D (2005) Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems. Magn Reson Chem 43:463–70. doi: 10.1002/mrc.1574

Finkel T, Holbrook NJ (2000) Oxidants, oxidative stress and the biology of ageing. Nature 408:239–247.

Fitos I, Simonyi M (1988) Selective effect of clonazepam and (S)-uxepam on the binding of warfarin enantiomers to human serum albumin. J Chromatogr 450:217–220.

Fitos I, Tegyey Z, Simonyi M, et al (1986) Stereoselective binding of 3-acetoxy-, and 3-hydroxy-1, 4-benzodiaze- pine-2-ones to human serum albumin. Selective allosteric interaction with warfarin enantiomers. Biochem Pharmacol 35:263–269.

Gallagher LT, Heo JS, Lopez MA, et al (2012) Pattern-based discrimination of organic acids and red wine varietals by arrays of synthetic receptors. Supramol Chem 24:2:37–41.

Garbi A, Sakkas V, Fiamegos YC, et al (2010) Sensitive determination of pesticides residues in wine samples with the aid of single-drop microextraction and response surface methodology. Talanta 82:1286–1291. doi: 10.1016/j.talanta.2010.06.046

Ghosh S, Petrin M, Maki AH, Sousa LR (1987) Dependence of the triplet state properties on the orientation of metal ion perturbers in naphthalene-crown ether metal ion complexes. I. External heavy atom effect. J Chem Phys 87:4315–4323. doi: 10.1063/1.452890

Goodacre R (2005) Metabolomics – the way forward. Metabolomics 1:1–2. doi: 10.1007/s11306-005-1111-7

Goto T, Ito Y, Yamada S, et al (2006) The high throughput analysis of N-methyl carbamate pesticides in fruits and vegetables by liquid chromatography electrospray ionization tandem mass spectrometry using a short column. Anal Chim Acta 555:225–232. doi: 10.1016/j.aca.2005.09.055

Gritti F, Guiochon G (2012) Facts and Legends on Columns Packed with Sub-3-µm Core-Shell Particles. LCGC North Am 30:586–595.

Gross JH (2011) Mass Spectrometry - A Textbook, 2nd edn. Springer

Gutierrez G, Reines HD, Wulf-Gutierrez ME (2004) Clinical review: hemorrhagic shock. Crit Care 8:373–381. doi: 10.1186/cc2851

Haidekker M a, Brady TP, Lichlyter D, Theodorakis E a (2006) A ratiometric fluorescent viscosity sensor. J Am Chem Soc 128:398–9. doi: 10.1021/ja056370a

Halim AAA, Kadir HA, Tayyab S (2008) Bromophenol blue binding as a probe to study urea and guanidine hydrochloride denaturation of bovine serum albumin. J Biochem 144:33–8. doi: 10.1093/jb/mvn036

Hamerly T, Bothner B (2015) Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis. Appl Biochem Biotechnol 1–13.

Hamerly T, Heinemann J, Tokmina-Lukaszewska M, et al (2014a) Bovine serum albumin as a molecular sensor for the discrimination of complex metabolite samples. Anal Chim Acta 818:61–66. doi: 10.1016/j.aca.2014.01.058

Hamerly T, Tripet BP, Tigges M, et al (2014b) Untargeted metabolomics studies employing NMR and LC–MS reveal metabolic coupling between Nanoarcheum equitans and its archaeal host Ignicoccus hospitalis. Metabolomics 11:895–907. doi: 10.1007/s11306-014-0747-6

He XM, Carter DC (1992) Atomic Structure and Chemistry of Human Serum Albumin. Nature 358:209–215.

Headley LM, Hardy JK (1992) Adulterants In Whiskey - Detection And Identification By Principal Component Analysis. J Food Sci 57:980–984.

Heinemann J, Hamerly T, Maaty WS, et al (2014) Expanding the paradigm of thiol redox in the thermophilic root of life. Biochim Biophys Acta 1840:80–85. doi: 10.1016/j.bbagen.2013.08.009

Heinonen M, Shen H, Zamboni N, Rousu J (2012) Metabolite identification and molecular fingerprint prediction through machine learning. Bioinformatics 28:2333–41. doi: 10.1093/bioinformatics/bts437

Herskovits TT, Gadegbeku B, Jaillet H (1970) On the Structural Stability Denaturation of Proteins. J Biol Chem 245:2588–2598.

High J, Coppin C a (1988) Wiley and the Whiskey Industry: Strategic Behavior in the Passage of the Pure Food Act. Bus Hist Rev 62:286–309.

Ho JX, Holowachuk EW, Norton EJ, et al (1993) X-ray and Primary Structure of Horse Serum Albumin (Equus Caballus) at 0.27-nm resolution. Eur J Biochem 212:205–212.

Hollywood K, Brison DR, Goodacre R (2006) Metabolomics: Current technologies and future trends. Proteomics 6:4716–4723. doi: 10.1002/pmic.200600106

Hopkins FG, Pinkus SN (1898) Observations on the Crystallization of Animal Proteids. J Physiol 23:130–136. doi: 10.1016/S0008-6215(00)00094-X

Jaitz L, Siegl K, Eder R, et al (2010) LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem 122:366–372. doi: 10.1016/j.foodchem.2010.02.053

Jansen JJ, Szymańska E, Hoefsloot HCJ, et al (2012) Between Metabolite Relationships: an essential aspect of metabolic change. Metabolomics 8:422–432. doi: 10.1007/s11306-011-0316-1

Johnson CH, Ivanisevic J, Benton HP, Siuzdak G (2015) Bioinformatics: The Next Frontier of Metabolomics. Anal Chem 87:147–156. doi: 10.1021/ac5040693

Jolliffe JT (2002) Principle Component Analysis, 2nd ed. Springer

Jupin M, Michiels PJ, Girard FC, et al (2014) NMR metabolomics profiling of blood plasma mimics shows that medium- and long-chain fatty acids differently release metabolites from human serum albumin. J Magn Reson 239:34–43. doi: 10.1016/j.jmr.2013.11.019

Jupin M, Michiels PJ, Girard FC, et al (2013) NMR identification of endogenous metabolites interacting with fatted and non-fatted human serum albumin in blood plasma: Fatty acids influence the HSA-metabolite interaction. J Magn Reson 228:81–94. doi: 10.1016/j.jmr.2012.12.010

Karp PD, Ouzounis C a, Moore-Kochlacs C, et al (2005) Expansion of the BioCyc Collection of Pathway/Genome Databases to 160 Genomes. Nucleic Acids Res 33:6083–6089. doi: 10.1093/nar/gki892

Kauvar DS, Wade CE (2005) The epidemiology and modern management of traumatic hemorrhage: US and international perspectives. Crit Care 9:S1–9. doi: 10.1186/cc3779

Khan PK, Sinha SP (1996) Ameliorating effect of vitamin C on murine sperm toxicity induced by three pesticides (endosulfan, phosphamidon and mancozeb). Mutagenesis 11:33–36. doi: 10.1093/mutage/11.1.33

Kim K, Aronov P, Zakharkin SO, et al (2009) Urine metabolomics analysis for kidney cancer detection and biomarker discovery. Mol Cell Proteomics 8:558–570. doi: 10.1074/mcp.M800165-MCP200

Kondo T (2014) Inconvenient truth: Cancer biomarker development by using proteomics. Biochim Biophys Acta 1844:861–865. doi: 10.1016/j.bbapap.2013.07.009

Krishnakumar SS, Panda D (2002) Spatial relationship between the prodan site, Trp-214, and Cys-34 residues in human serum albumin and loss of structure through incremental unfolding. Biochemistry 41:7443–7452. doi: 10.1021/bi025699v

Kubarych CJ, Adams MM, Anslyn E V (2010) Serum albumins as differential receptors for the discrimination of fatty acids and oils. Org Lett 12:4780–4783. doi: 10.1021/ol101906g

Kumar A, Remmel RP, Mann HJ, Beilman GJ (2011) Drug metabolism in hemorrhagic shock: pharmacokinetics of selective markers of cytochrome-P450 2C9, 2D6, and 3A4 enzyme activities in a porcine model. J Surg Res 167:e231–243. doi: 10.1016/j.jss.2010.06.040

Larroque C, Pelegrin A, Van Lier JE (1996) Serum albumin as a vehicle for zinc phthalocyanine: photodynamic activities in solid tumour models. Br J Cancer 74:1886–90.

Lay JO, Liyanage R, Borgmann S, Wilkins CL (2006) Problems with the "omics." Trends Anal Chem 25:1046–1056. doi: 10.1016/j.trac.2006.10.007

Lin L, Huang Z, Gao Y, et al (2011) LC-MS based serum metabonomic analysis for renal cell carcinoma diagnosis, staging, and biomarker discovery. J Proteome Res 10:1396–405. doi: 10.1021/pr101161u

Lowenthal MS, Mehta AI, Frogale K, et al (2005) Analysis of albumin-associated peptides and proteins from ovarian cancer patients. Clin Chem 51:1933–45. doi: 10.1373/clinchem.2005.052944

Maaty WS, Wiedenheft B, Tarlykov P, et al (2009) Something old, something new, something borrowed; how the thermoacidophilic archaeon Sulfolobus solfataricus responds to oxidative stress. PLoS One 4:e6964. doi: 10.1371/journal.pone.0006964

Majorek KA, Porebski PJ, Dayal A, et al (2012) Structural and immunologic characterization of bovine, horse, and rabbit serum albumins. Mol Immunol 52:174–182. doi: 10.1016/j.molimm.2012.05.011

Martí MP, Busto O, Guasch J (2004) Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing. J Chromatogr A 1057:211–217. doi: 10.1016/j.chroma.2004.08.143

Mcgovern P (2000) Origins and Ancient History of Wine. Psychology Press

Metzger A, Anslyn E V. (1998) A Chemosensor for Citrate in Beverages. Angew Chemie Int Ed 37:649–652. doi: 10.1002/(SICI)1521-3773(19980316)37:5<649::AID-ANIE649>3.0.CO;2-H

Mitzner SR, Stange J, Klammt S, et al (2001) Extracorporeal detoxification using the molecular adsorbent recirculating system for critically ill patients with liver failure. J Am Soc Nephrol 12:S75–S82.

Motta A, Paris D, Melck D (2010) Monitoring real-time metabolism of living cells by fast two-dimensional NMR spectroscopy. Anal Chem 82:2405–11. doi: 10.1021/ac9026934

Nafisi S, Bagheri Sadeghi G, PanahYab A (2011) Interaction of aspirin and vitamin C with bovine serum albumin. J Photochem Photobiol B Biol 105:198–202. doi: 10.1016/j.jphotobiol.2011.09.002

Patti G, Yanes O, Shriver L (2012a) Metabolomics implicates altered sphingolipids in chronic pain of neuropathic origin. Nat Chem … 8:232–234. doi: 10.1038/NCHEMBIO.767

Patti GJ, Tautenhahn R, Siuzdak G (2012b) Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc 7:508–16. doi: 10.1038/nprot.2011.454

Peters Jr. T (1996) All About Albumin. Academic Press, San Diego, California

Petitpas I, Bhattacharya AA, Twine S, et al (2001a) Crystal structure analysis of warfarin binding to human serum albumin: anatomy of drug site I. J Biol Chem 276:22804–9. doi: 10.1074/jbc.M100575200

Petitpas I, Gru T, Bhattacharya AA, Curry S (2001b) Crystal Structures of Human Serum Albumin Complexed with Monounsaturated and Polyunsaturated Fatty Acids. J Mol Biol 314:955–960. doi: 10.1006/jmbi.2001.5208

Pettersen EF, Goddard TD, Huang CC, et al (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1602–12.

Pluskal T, Castillo S, Villar-Briones A, Orešič M (2010) MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinformatics 11:395. doi: 10.1186/1471-2105-11-395

Pond SJK, Tsutsumi O, Rumi M, et al (2004) Metal-ion sensing fluorophores with large two-photon absorption cross sections: Aza-crown ether substituted donor-acceptor-donor distyrylbenzenes. J Am Chem Soc 126:9291–9306. doi: 10.1021/ja049013t

Pontes MJC, Santos SRB, Araújo MCU, et al (2006) Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. Food Res Int 39:182–189. doi: 10.1016/j.foodres.2005.07.005

R Core Team (2012) R: A language and environment for statistical computing.

R. Lusczek E, Nelson T, Lexcen D, et al (2011) Urine Metabolomics in Hemorrhagic Shock: Normalization of Urine in the Face of Changing Intravascular Fluid Volume and Perturbations in Metabolism. J Bioanal Biomed 03:38–48. doi: 10.4172/1948-593X.1000041

Reynolds J, Herbert S, Steinhardt J (1968) The Binding of Some Long-Chain Fatty Acid Anions and Alcohols by Bovine Serum Albumin. Biochemistry 7:1357–1361.

Richieri G V, Anel a, Kleinfeld a M (1993) Interactions of long-chain fatty acids and albumin: determination of free fatty acid levels using the fluorescent probe ADIFAB. Biochemistry 32:7574–80.

Rogers CW, Wolf MO (2002) Luminescent molecular sensors based on analyte coordination to transition-metal complexes. Coord Chem Rev 233-234:341–350. doi: 10.1016/S0010-8545(02)00023-1

Russell I, Stewart G (2014) Whiskey: Technology, Production and Marketing, 2nd edn. Elsevier Ltd

Sannino A, Bolzoni L, Bandini M (2004) Application of liquid chromatography with electrospray tandem mass spectrometry to the determination of a new generation of pesticides in processed fruits and vegetables. J Chromatogr A 1036:161–169. doi: 10.1016/j.chroma.2004.02.078

Scribner DM, Witowski NE, Mulier KE, et al (2010) Liver metabolomic changes identify biochemical pathways in hemorrhagic shock. J Surg Res 164:e131–139. doi: 10.1016/j.jss.2010.07.046

Shao J, Shen H, Havsteen B (1993) Purification, characterization and binding interactions of the Chinese-cobra (Naja naja atra) serum antitoxic protein CSAP. Biochem J 293:559–66.

Sheets LP (2010) Hayes' Handbook of Pesticide Toxicology. Elsevier

Singleton VL (1962) Aging of wines and other spiritous products, acceleration by physical treatments. Hilgardia 32:319–392. doi: 10.3733/hilg.v32n07p319

Siuzdak G (2014) Mass spectrometry-based metabolomics as a unique biochemical approach for understanding disease pathogenesis. FASEB J 28:350.3.

Siuzdak G (2006) The Expanding Role of Mass Spectrometry in Biotechnology, 2nd edn. MCC Press, San Diego, California

Sjöholm I, Ekman B, Kober A, et al (1979) Binding of drugs to human serum albumin:XI. The specificity of three binding sites as studied with albumin immobilized in microparticles. Mol Pharmacol 16:767–777.

Skarda DE, Putt KS, Hergenrother PJ, et al (2007) Increased poly(ADP-ribose) polymerase activity during porcine hemorrhagic shock is transient and predictive of mortality. Resuscitation 75:135–44. doi: 10.1016/j.resuscitation.2007.02.020

Smith C a, O'Maille G, Want EJ, et al (2005) METLIN: a metabolite mass spectral database. Ther Drug Monit 27:747–751.

Smith CA, Want EJ, O'Maille G, et al (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Anal Chem 78:779–787.

Soleas GJ, Diamandis EP, Goldberg DM (1997) Wine as a biological fluid: history, production, and role in disease prevention. J Clin Lab Anal 11:287–313.

Spector a a, Fletcher JE, Ashbrook JD (1971) Analysis of long-chain free fatty acid binding to bovine serum albumin by determination of stepwise equilibrium constants. Biochemistry 10:3229–32.

Spector AA, John K, Fletcher JE (1969) Binding of long-chain fatty acids to bovine serum albumin. J Lipid Res 10:56–67.

Sundale Research (2013) State of the Industry: Craft Distilleries and Distilled Spirits in the U.S. Bay Shore, NY

Sutharsan J, Lichlyter D, Wright NE, et al (2010) Molecular rotors: Synthesis and evaluation as viscosity sensors. Tetrahedron 66:2582–2588. doi: 10.1016/j.tet.2010.01.093

Szymanski J, Jozefczuk S, Nikoloski Z, et al (2009) Stability of metabolic correlations under changing environmental conditions in Escherichia coli--a systems approach. PLoS One 4:e7441. doi: 10.1371/journal.pone.0007441

Tautenhahn R, Boettcher C, Neumann S (2008) Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504.

Tautenhahn R, Cho K, Uritboonthai W, et al (2012) An accelerated workflow for untargeted metabolomics using the METLIN database. Nat Biotechnol 30:826–828. doi: 10.1038/nbt.2348

Tautenhahn R, Patti GJ, Kalisiak E, et al (2011) metaXCMS: Second-Order Analysis of Untargeted Metabolomics Data. Anal Chem 83:696–700.

Taylor JH, Beilman GJ, Conroy MJ, et al (2004) Tissue energetics as measured by nuclear magnetic resonance spectroscopy during hemorrhagic shock. Shock 21:58–64. doi: 10.1097/01.shk.0000101674.49265.93

Thompson JJ (1913) Rays of Positive Electricity and Their Applications to Chemical Analysis. Longmans, Green and Co., London

Thomson JJ (1907) On Rays of Positive Electricity. Philos Mag 13:561–575. doi: 10.1080/14786440709463633

Thomson JJ (1899) On the Masses of the Ions in Gases at Low Pressures. Philos Mag 48:547–567. doi: 10.1080/14786449908621447

Umali AP, Anslyn E V (2010) A general approach to differential sensing using synthetic molecular receptors. Curr Opin Chem Biol 14:685–692. doi: 10.1016/j.cbpa.2010.07.022

Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, et al (2015) Metabolomics for Biomarkers of Type 2 Diabetes Mellitus: Advances and Nutritional Intervention Trends. Curr Cardiovasc Risk Rep 9:12. doi: 10.1007/s12170-015-0440-y

Varshney A, Sen P, Ahmad E, et al (2010) Review Article Ligand Binding Strategies of Human Serum Albumin : How Can the Cargo be Utilized ? Chirality 22:77–87. doi: 10.1002/chir

Villiers A De, Vanhoenacker G, Majek P, Sandra P (2004) Determination of anthocyanins in wine by direct injection liquid chromatography–diode array detection–mass spectrometry and classification of wines using discriminant analysis. J Chromatogr A 1054:195–204. doi: 10.1016/j.chroma.2004.07.087

Wiskur SL, Ait-Haddou H, Lavigne JJ, Anslyn E V (2001) Teaching old indicators new tricks. Acc Chem Res 34:963–72.

Woese CR (1987) Bacterial Evolution Background. Microbiol Rev 51:221–271.

Worley B, Powers R (2013) Multivariate Analysis in Metabolomics. Curr Metabolomics 1:92–107.

Wulffen B, Buff MCR, Pofahl M, et al (2012) Caged glucosamine-6-phosphate for the light-control of riboswitch activity. Photochem Photobiol Sci 11:489–92. doi: 10.1039/c1pp05242j

Yamashita M, Fenn JB (1984) Electrospray ion source. Another variation on the free-jet theme. J PhysChem 88:4451–4459. doi: 10.1021/j150664a002

Zunszain P a, Ghuman J, Komatsu T, et al (2003) Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. BMC Struct Biol 3:1–9. doi: 10.1186/1472-6807-3-6