



Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis

Authors: Timothy Hamerly & Brian Bothner

This is a postprint of an article that originally appeared in [Applied Biochemistry and Biotechnology](#) on September 2015. The final publication is available at Springer via <http://dx.doi.org/10.1007/s12010-015-1861-2>.

Hamerly, T., & Bothner, B. (2015). Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis. *Applied Biochemistry and Biotechnology*, 178(1), 101–113.
doi:10.1007/s12010-015-1861-2

Made available through Montana State University's [ScholarWorks](http://scholarworks.montana.edu)
scholarworks.montana.edu

Investigations into the Use of a Protein Sensor Assay for Metabolite Analysis

Timothy Hamerly¹ • Brian Bothner¹

¹ Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717, USA



Brian Bothner

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 the 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 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 the 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 [1, 2]. 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 have led to a gradual shift toward metabolomics for biomarker discovery [3]. Current enthusiasm for metabolomics may be justified because metabolites are a direct readout of genetic and proteomic potential [4].

In a previous report, we showed that small molecules bound to the protein bovine serum albumin (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 [5]. 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) [5]. 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 [5]. 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 [6–15]. SA is found in the blood of all mammals and assists in transporting nonpolar molecules to muscle tissue and organs [15–18]. 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 [7, 9, 13–16, 18, 19]. The working model is that SA binds a subset of all molecules present in solution based on each molecule's 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 archaeal 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 [20–22]. Over the past decade, it 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 [20, 23, 24]. 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 [25–27].

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 (NaCl) 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) and 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 [20]. Briefly, ATCC strain *S. solfataricus* P2 (SSP2) was cultured in DSMZ media 182 with the addition of 0.1 % glucose to OD₆₅₀ of 0.60. A control sample (0 min) was collected by taking 50 mL of culture into a Falcon tube and cell pellet collected by centrifugation. Following this sample, cells were stressed with hydrogen peroxide (H₂O₂) at a final concentration of 30 μM and cells were collected at 15, 30, 60, and 105 min 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 [20]. 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 h. The supernatant was divided into multiple vials and evaporated to dryness using a SpeedVac. 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 3000 Dalton (3-kD) 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 and 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 min and then washed over a 3-kD MWCO spin filter. Samples were centrifuged for 5 min and 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 supernatant was combined and dried using a SpeedVac. Samples were stored at -80 °C until 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 × 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 min, to 95 % B over 10 min, then held at 95 % B for 2 min, before finally returning to 2 % B for 1 min, for a total run time of 15 min. 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 3500 V and a fragmentor voltage of 120 V. 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 min giving a total run time of 30 min. 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 LC-MS data, for both PSA-treated and untreated, has been previously described [5]. Briefly, molecular features from PSA-treated and untreated samples were extracted using Masshunter Software (Agilent, Santa Clara, CA) and analyzed using standard XCMS protocols [28–30]. Following XCMS analysis, differential reports (ANOVA) from each time point were imported into metaXCMS for alignment [31]. The resulting aligned

molecular feature lists were analyzed by principal component analysis (PCA) using XLStat and visualized using the R package rgl [32, 33]. Data for Venn diagrams were annotated by hand and visualized using the R package VennDiagram [34].

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 0.1 mg mL^{-1} was mixed with a solution of metabolites from *S. solfataricus*. This mixture was washed over 3-kD 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 min (standard PSA time) or 3, 8, or 24 h, before centrifugation. Metabolites remaining bound to BSA were eluted with MeOH 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 min were analyzed. Batch cultures of cells, at 80°C , pH 3.2, were spiked with hydrogen peroxide (H_2O_2) to a final concentration of $30 \text{ }\mu\text{M}$. This low dose does not significantly slow growth but has been shown to induce a cellular response at the level of mRNA and protein [24]. Cells were collected at five time points to track the metabolic response to hydrogen peroxide ($T = 0, 15, 30, 60$, and 105 min). 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 min after the addition of hydrogen peroxide are shown in Fig. 1. The metabolite fractions analyzed directly by LC-MS were dominated by early-eluting (more polar) compounds (Fig. 1a). Small differences between biological replicates were visible (gray 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 (Fig. 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 PCA. The complex datasets resulting from LC-MS are well suited for PCA because it reduces dimensionality providing a straightforward, unsupervised visualization of sample variance [35]. A plot of the samples analyzed directly by LC-MS showed that the biological replicates fail to cluster in a consistent manner (Fig. 2a). However, if the same samples were PSA treated before LC-MS analysis, biological replicates clustered tightly (Fig. 2b). The data also segregated temporally 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 min) and 15-min post stress samples which were overlapped, presumably because the metabolome had yet to change significantly.

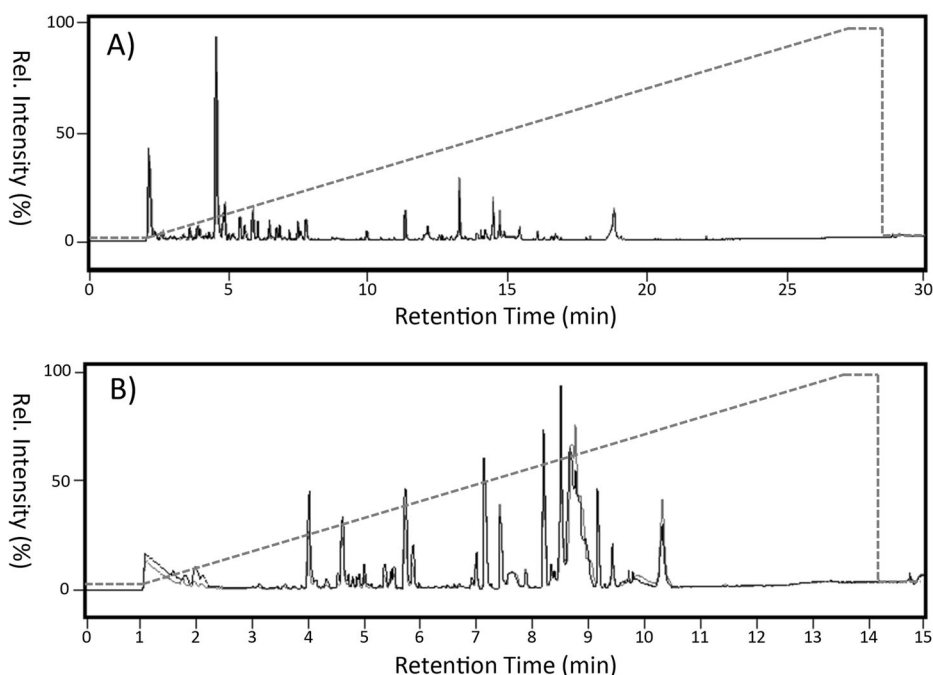


Fig. 1 LC-MS chromatograms of metabolite extracts of *S. solfataricus* cells. **a** TIC of reverse-phase LC-MS for two biological replicates of metabolite extracts 30 min after addition of 30 μ M hydrogen peroxide. **b** The same two biological samples treated with the PSA before LC-MS analysis. Biological replicates are similar in both analyses; however, the reverse-phase elution profile is different. Dashed line indicates the % acetonitrile across the chromatogram

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-enhanced differentiation. To begin, the LC-MS data was analyzed to find the specific markers that differentiate at later

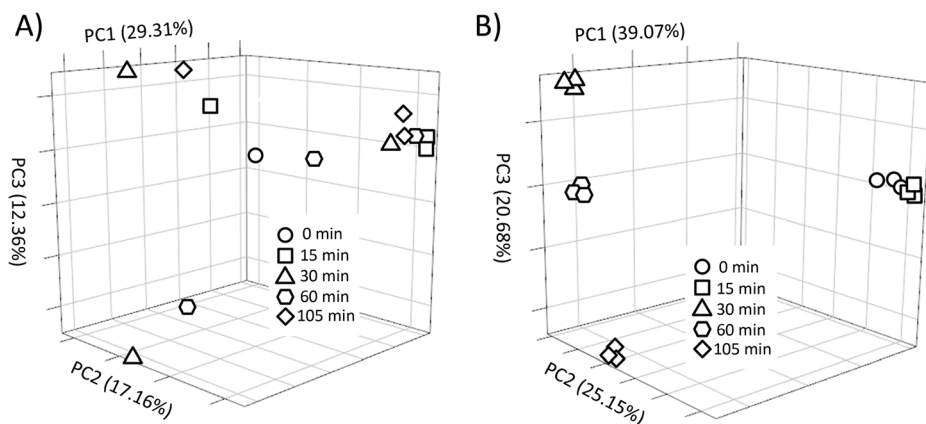


Fig. 2 Principal component analysis of *S. solfataricus* cells undergoing oxidative stress. **a** PCA plot of LC-MS data from metabolite samples directly after extraction from cells. **b** The same samples after treatment with the PSA, followed by LC-MS analysis. Three biological replicates were used at each time point. Key indicates time after induction of mild oxidative stress

time points from control (0 min). The time point 105 min 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-min samples were selected because the system was perturbed, but based on PC3 showed signs of recovery (see Fig. 2b and previously published work from our lab) [24]. To assess which molecular features were 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.

In total, nine features were selected for further analysis. Table 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 italics are those with significant change (fold change 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 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 [36–39].

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 (Fig. 3a). Histogram plots of features

Table 1 Molecular features of interest

Number	<i>m/z</i>	PSA treated		Untreated		METLIN
		Fold	<i>p</i> value	Fold	<i>p</i> value	
1	<i>229.14</i>	9.13	0.002	1.00	0.777	Fatty acid
2	<i>611.35</i>	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	<i>226.10</i>	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

Features in italics are those with significant change (fold change greater than 1.5 and *p* value less than 0.05) in the PSA groups

PE DAG phosphatidyl ethanolamine diacylglycerol

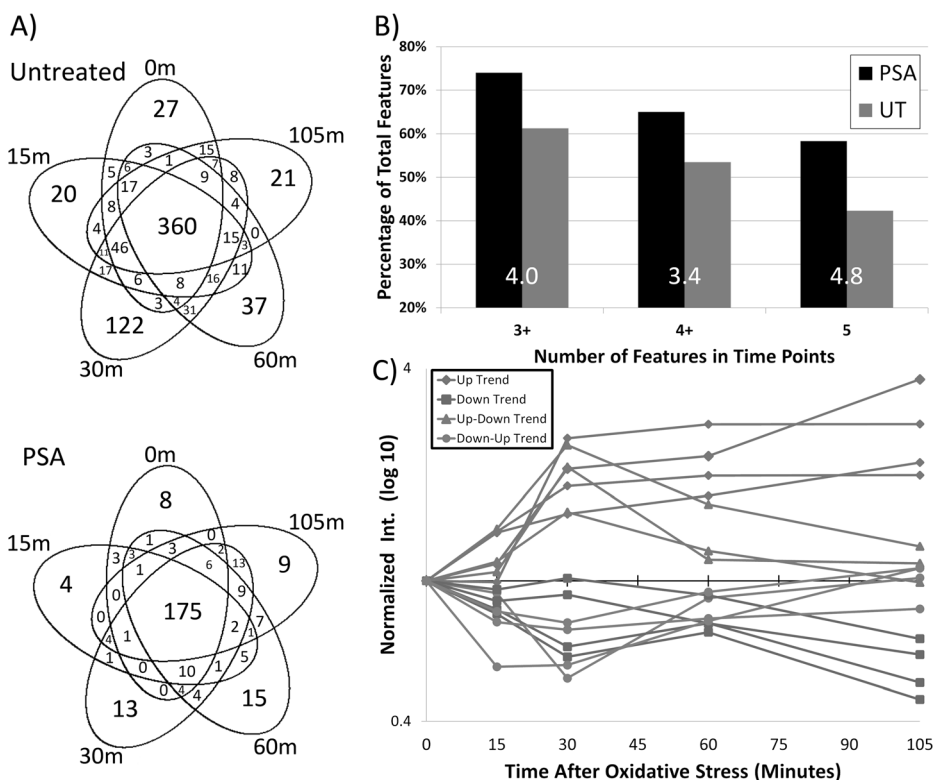


Fig. 3 Analysis showing the consistency of features throughout a time course after PSA treatment. **a** Venn diagram showing the distribution of features when samples were analyzed directly by LC-MS (*top*) and when treated with the PSA prior to LC-MS (*bottom*). **b** Histogram showing the percentage of metabolites in three or more, four or more, and all five 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 0.001). **c** Four basic abundance trends were present in features detected in all five time points after PSA treatment. Metabolites that trend up (*diamonds*), metabolites that trend down (*triangles*), metabolites which trend up at first and then come back down (*squares*), and metabolites which trend down at first and then come back up (*circles*). Error bars for these data points were smaller than the markers and were left off the plot for clarity

present in three or more, four or more, and all five time points showed that a greater percentage of features were common between multiple time points when the PSA was used (Fig. 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 five 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 (Fig. 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).

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 [9, 10, 14, 19]. 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 h 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 (Fig. 4). The fraction bound was calculated by dividing the ion intensity at each equilibration period by the intensity after 5 min (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) and 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 h of equilibration were compared. While most of the features in Table 2 are more intense after being 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 2). Thus, the

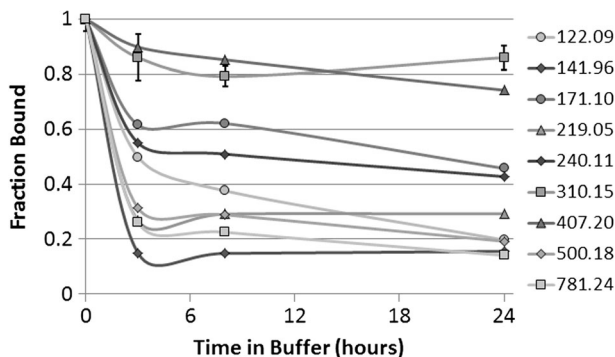


Fig. 4 Disassociation curves of BSA-bound metabolites. Curves showing fraction bound to BSA of nine different metabolites over a 24-h period of time. Metabolite extracts from *S. solfataricus* were treated using the PSA and left to equilibrate before standard elution with MeOH to remove bound metabolites. Fraction bound was calculated based on LC-MS ion intensity after equilibration compared to the standard 5-min period used in the PSA. Lines connecting data points were added to improve visualization. Error bars have been added for the feature 310.15 m/z to show the reproducibility of the binding, while not cluttering the plot

Table 2 Intensity and rank for selected molecular features before and after PSA treatment

<i>m/z</i>	Ion intensity		Rank ^a		Percentile ^a	
	Untreated	PSA	Untreated	PSA	Untreated (%)	PSA (%)
122.09	57,980	66,314	771/845	213/240	4.56	11.61
141.96	58,299	618,234	769/845	46/240	4.97	80.91
171.10	442,902	156,162	227/845	136/240	61.41	43.56
219.05	204,666	3,383,842	416/845	14/240	32.78	94.19
240.11	49,510	861,739	794/845	38/240	3.31	84.23
310.15	125,195	541,025	548/845	53/240	20.22	78.00
407.20	184,743	283,405	443/845	91/240	29.44	62.24
500.18	40,287	315,432	810/845	87/240	2.25	63.90
781.24	64,705	52,160	755/845	227/240	5.37	5.80

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

larger the rank value, the smaller the percentile value, indicates that a feature is not very abundant. By using a percentile rank, it is possible to compare specific ions between untreated and PSA-treated 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 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 are 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 [20]. As a comparison, a standard nontargeted 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 (Fig. 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 [15]. 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- μ M H₂O₂ treatment is referred to as “mild” because it does not inhibit growth of *S. solfataricus* [24]. 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 (Fig. 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 min were clearly separated, and a logical trajectory appeared (Fig. 2b). PC1 accounted for the movement of these three time points away from the control and 15-min groups. Metabolic recovery is also evident because at the final time point (105 min), 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 min) and 105 min 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 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 (Fig. 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 (Fig. 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 (Fig. 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 (Fig. 4). This agrees with previous studies showing that SAs have a number of binding pockets with a range of affinities [16]. 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 m/z increased by as much as 100-fold in the PSA LC-MS data (Table 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 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 albumin 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.

Acknowledgments 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

1. Lay, J. O., Liyanage, R., Borgmann, S., & Wilkins, C. L. (2006). Problems with the “omics.”. *Trends in Analytical Chemistry*, 25(11), 1046–1056.
2. Broadhurst, D. I., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, 2(4), 171–196.
3. Kondo, T. (2014). Inconvenient truth: cancer biomarker development by using proteomics. *Biochimica et Biophysica Acta*, 1844(5), 861–865.
4. Siuzdak, G. (2014). Mass spectrometry-based metabolomics as a unique biochemical approach for understanding disease pathogenesis. *The FASEB Journal*, 28, 350.3.
5. Hamerly, T., Heinemann, J., Tokmina-Lukaszewska, M., Luszczyk, E. R., Mulier, K. E., Beilman, G. J., & Bothner, B. (2014). Bovine serum albumin as a molecular sensor for the discrimination of complex metabolite samples. *Analytica Chimica Acta*, 818, 61–66.
6. 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(29), 7574–7580.
7. Peters Jr., T. (1996). *All about albumin*. San Diego: Academic Press.
8. 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. *Nature Structural Biology*, 5(9), 827–835.
9. Curry, S., Brick, P., & Franks, N. P. (1999). Fatty acid binding to human serum albumin: new insights from crystallographic studies. *Biochimica et Biophysica Acta*, 1441(2–3), 131–140.
10. Zunszain, P. A., Ghuman, J., Komatsu, T., Tsuchida, E., & Curry, S. (2003). Crystal structural analysis of human serum albumin complexed with hemin and fatty acid. *BMC Structural Biology*, 3, 1–9.
11. Fielding, L., Rutherford, S., & Fletcher, D. (2005). Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems. *Magnetic resonance in chemistry: MRC*, 43(6), 463–470.
12. Nafisi, S., Bagheri Sadeghi, G., & PanahYab, A. (2011). Interaction of aspirin and vitamin C with bovine serum albumin. *Journal of Photochemistry and Photobiology B: Biology*, 105(3), 198–202.
13. Reynolds, J., Herbert, S., & Steinhart, J. (1968). The binding of some long-chain fatty acid anions and alcohols by bovine serum albumin. *Biochemistry*, 7(4), 1357–1361.
14. Petitpas, I., Bhattacharya, A. A., Twine, S., East, M., & Curry, S. (2001). Crystal structure analysis of warfarin binding to human serum albumin: anatomy of drug site I. *The Journal of Biological Chemistry*, 276(25), 22804–22809.
15. Varshney, A., Sen, P., Ahmad, E., Rehan, M., Subbarao, N., & Khan, R. H. (2010). Review article ligand binding strategies of human serum albumin : how can the cargo be utilized? *Chirality*, 22(January 2009), 77–87.
16. He, X. M., & Carter, D. C. (1992). Atomic structure and chemistry of human serum albumin. *Nature*, 358, 209–215.
17. Dockal, M. (2000). Conformational transitions of the three recombinant domains of human serum albumin depending on pH. *Journal of Biological Chemistry*, 275(5), 3042–3050.
18. Fanali, G., di Masi, E., Trezza, V., Marino, M., Fasano, M., & Ascenzi, P. (2012). Human serum albumin: from bench to bedside. *Molecular Aspects of Medicine*, 33(3), 209–290.

19. Petitpas, I., Gru, T., Bhattacharya, A. A., & Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *Journal of Molecular Biology*, 314, 955–960.
20. Heinemann, J., Hamerly, T., Maaty, W. S., Movahed, N., Steffens, J. D., Reeves, B. D., Hilmer, J. K., Therien, J., Grieco, P. A., Peters, P. W., & Bothner, B. (2014). Expanding the paradigm of thiol redox in the thermophilic root of life. *Biochimica et Biophysica Acta*, 1840(1), 80–85.
21. Esser, D., Kouril, T., Zaparty, M., Sierocinski, P., Chan, P. P., Lowe, T., Van der Oost, J., Albers, S., Schomburg, D., Makarova, K. S., & Siebers, B. (2011). Functional curation of the *Sulfolobus solfataricus* P2 and *S. acidocaldarius* 98-3 complete genome sequences. *Extremophiles*, 15(6), 711–712.
22. Brock, T. D., Brock, K. M., Belly, R. T., & Weiss, R. L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Archiv für Mikrobiologie*, 84(1), 54–68.
23. Woese, C. R. (1987). Bacterial evolution background. *Microbiological Reviews*, 51(2), 221–271.
24. Maaty, W. S., Wiedenheft, B., Tarylkov, P., Schaff, N., Heinemann, J., Robison-Cox, J., Valenzuela, J., Dougherty, A., Blum, P., Lawrence, C. M., Douglas, T., Young, M. J., & Bothner, B. (2009). Something old, something new, something borrowed; how the thermoacidophilic archaeon *Sulfolobus solfataricus* responds to oxidative stress. *PLoS One*, 4(9), e6964.
25. Finkel, T., & Holbrook, N. J. (2000). Oxidants, oxidative stress and the biology of ageing. *Nature*, 408, 239–247.
26. Buttke, T. M., & Sandstrom, P. A. (1994). Oxidative stress as a mediator of apoptosis. *Immunology Today*, 15(1), 7–10.
27. Apel, K., & Hirt, H. (2004). Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology*, 55, 373–399.
28. Core Team, R. (2012). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
29. Tautenhahn, R., Boettcher, C., & Neumann, S. (2008). Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics*, 9, 504.
30. Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., & Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. *Analytical Chemistry*, 78, 779–787.
31. Tautenhahn, R., Patti, G. J., Kalisiak, E., Miyamoto, T., Schmidt, M., Lo, F. Y., McBee, J., Baliga, N. S., & Siuzdak, G. (2011). MetaXCMS: second-order analysis of untargeted metabolomics data. *Analytical Chemistry*, 83(3), 696–700.
32. Adler, D., & Murdoch, D. (2012). rg1: 3D visualization device system (OpenGL).
33. Addinsoft. (2013). XLSTAT. Addinsoft.
34. Chen, H. (2013). VennDiagram: generate high-resolution Venn and Euler plots.
35. Worley, B., & Powers, R. (2013). Multivariate analysis in metabolomics. *Current Metabolomics*, 1(1), 92–107.
36. Spector, A. A., John, K., & Fletcher, J. E. (1969). Binding of long-chain fatty acids to bovine serum albumin. *Journal of Lipid Research*, 10(1), 56–67.
37. Lowenthal, M. S., Mehta, A. I., Frogale, K., Bandle, R. W., Araujo, R. P., Hood, B. L., Veenstra, T. D., Conrads, T. P., Goldsmith, P., Fishman, D., Petricoin III, E. F., & Liotta, L. A. (2005). Analysis of albumin-associated peptides and proteins from ovarian cancer patients. *Clinical Chemistry*, 51(10), 1933–1945.
38. Jupin, M., Michiels, P. J., Girard, F. C., Spraul, M., & Wijmenga, S. S. (2013). NMR identification of endogenous metabolites interacting with fatty and non-fatty human serum albumin in blood plasma: fatty acids influence the HSA-metabolite interaction. *Journal of Magnetic Resonance*, 228, 81–94.
39. Jupin, M., Michiels, P. J., Girard, F. C., Spraul, M., & Wijmenga, S. S. (2014). NMR metabolomics profiling of blood plasma mimics shows that medium- and long-chain fatty acids differently release metabolites from human serum albumin. *Journal of Magnetic Resonance*, 239, 34–43.