The Development of a Protein Integrity Metric for Biorepository Specimens: Introduction of the "SPIN" Index

Timothy Geddes[1], Samreen Ahmed[1], Bryan Thibodeau[1], Barb Pruetz[1], Dianna Larson[1], Jan Akervall[1,2]. George Wilson[1,3]

[1]Beaumont BioBank, [2]Department of General Surgery, [3]Department of Radiation Oncology
Beaumont Health System, Royal Oak, Michigan, USA

Beaumont® | HEALTH SYSTEM

Background

A growing concern in the advancing biorepository field is specimen integrity. Although quality metrics exist for the assessment of nucleic acids (i.e. the universally accepted RNA Integrity Number, or RIN), there is not to date an accepted index of protein integrity for the myriad of biological samples that are collected, processed and stored in biorepositories around the world. Here we describe the creation of a program to discover proteins in individual sample types with the goal of establishing an index to assess the integrity of stored specimens for protein-based biomarker studies. The study mimics potential variation in pre-analytical conditions common in biorepositories that may result in proteolysis or other proteome-associated changes. Surface-enhanced laser desorption time-of-flight mass spectrometry (SELDI-TOF MS) is employed to assess changes in multiple proteins and peptides in a high-throughput manner. Candidate proteins from representative sample types were employed to develop a relative index which is derived from a ratio of non-stable vs. stable proteins for each sample type in the investigation. We refer to this metric as the Sample-specific Protein Integrity Number, or SPIN.

Materials and Methods

Serum, plasma, urine, thyroid papillary cancer tissue, and periprosthetic spinal vertebral tissue were collected according to institutional HIC policies and processed according to the Beaumont Research Institute's BioBank Standard Operating Procedures (SOPs) with various modifications designed to introduce pre-analytical variability which may be encountered in biorepository laboratories (Figure 1). Serum and urine specimens were divided into aliquots immediately after collection and separate aliquots were exposed to variation. Tissue specimens were snap frozen then minced to homogeneity, divided equally, and each portion exposed to processing variation. Extracts were then prepared from tissues and normalized to protein. Samples were diluted into binding buffer and spotted in duplicate onto ProteinChip CM10 (weak cation exchange) surfaces and subjected to SELDI analysis. The resulting spectra were normalized and peak clusters within one experimental condition were created. Following univariate statistical analysis and P value calculation, differences in spectra patterns and peak intensities were studied. Significant peaks clusters ($p<0.05$) were then considered candidates for future identification and protein integrity focus. Interim SPIN numbers (ratio of non-stable to stable SELDI peak intensities) were assigned.

To demonstrate the effect of sub-optimal sample handling (as indicated by the SPIN index) on the outcomes of laboratory analytical determinations, quantitative analysis (cytokine bead array or analyte-specific ELISA) was then performed according to standard protocols.

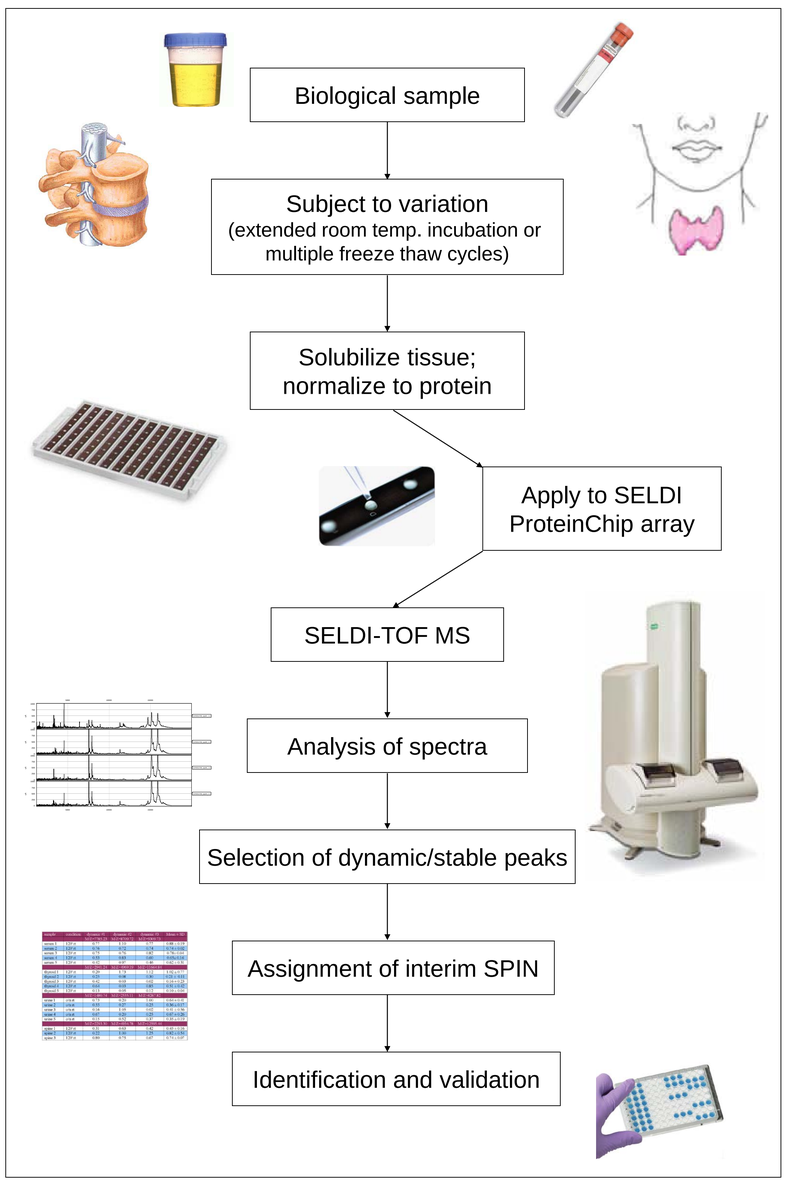
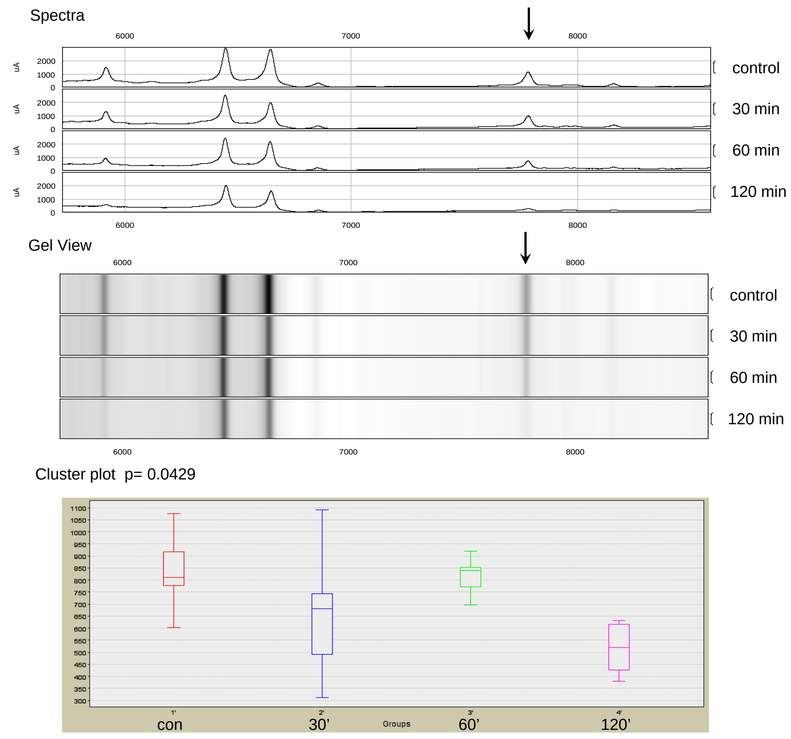


Figure 1: Process flow for SPIN candidate peptide or protein selection

Results

Selection of SPIN candidate peaks: For each sample type in the study, two categories of protein species (M/Z) were selected from SELDI proteomic profiles. Of these, one displayed dynamic changes in peak intensity when samples were exposed to extended incubations at room temperature or multiple freeze/thaw cycles, and the second was chosen as a stable internal reference (i.e. peak intensity remained the same in all conditions). Selection criteria for the dynamic protein peaks include: 1.) peak is present in all conditions studied, 2.) peak intensity changes from control to more rigorous handling conditions (> 2 fold change), 3.) easily identified as a clearly separated peak, 4.) intensity greater than 10, and 5.) p value <0.05 upon univariate statistical analysis of clustered peaks within one experimental condition. The selected stable control peak demonstrated a p value >0.95 (Figure 2 and Figure 3).

Dynamic protein M/Z = 7785.23 (serum at room temp.)



Stable protein M/Z = 3899.35 (serum at room temp.)

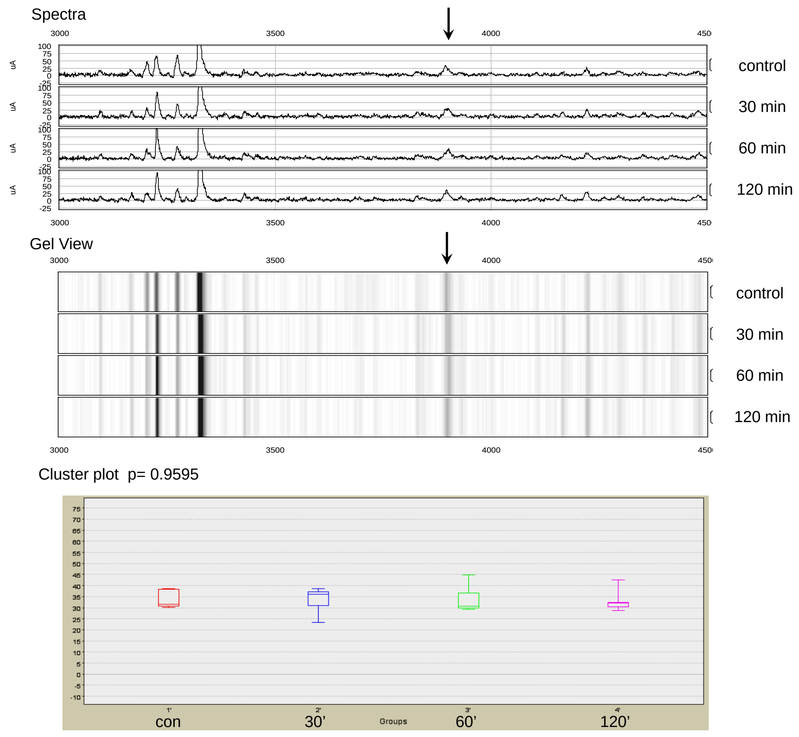


Figure 2: Representative spectra/gel view of dynamic and stable protein peak species as determined by univariate analysis of peak clusters. A serum was incubated at room temperature for 30, 60, or 120 minutes.

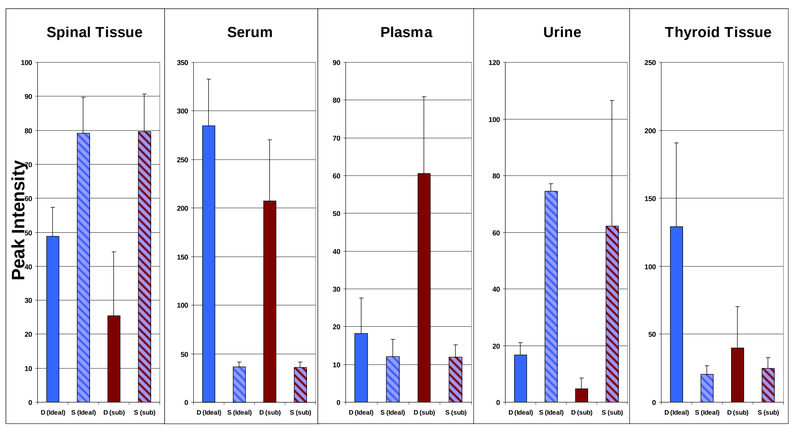


Figure 3: Dynamic protein peak intensity variation upon exposing samples to pre-analytical variability. Various sample types were subjected to ideal processing (Ideal) conditions or incubated at room temperature for two hours prior to freezing (sub). D = dynamic protein, S = stable protein

Interim SPIN assignment in each sample type: After selecting a stable protein peak and a dynamic protein peak using SELDI analysis and statistics software and p value analysis (Table 1), a SPIN index for each sample was calculated by dividing the reported peak intensities of dynamic protein peaks by the peak intensity of the stable protein. Indexes were averaged for each experimental condition and expressed as mean SPIN. Samples were then scored as sub-optimal based on the comparison of the dynamic peak intensity/stable peak intensity ratio (Table 2).

Sample	Dynamic protein peak	Stable protein peak
spinal tissue	m/z 2293, p = 0.006	m/z 4028 p = 1.000
serum	m/z 7787, p = 0.028	m/z 3899 p = 0.917
plasma	m/z 1899, p = 0.000	m/z 17380 p = 0.993
urine	m/z 6268, p = 0.047	m/z 8875 p = 0.917
thyroid tissue	m/z 2941, p = 0.009	m/z 11430 p = 0.917

Table 1. Selection of dynamic (D) and stable (S) peaks in each sample type

Sample	dynamic peak, ideal	stable peak, ideal	Ideal **SPIN**	dynamic peak, 2hr RT	stable peak, 2hr RT	2hr RT **SPIN**	score
spinal tissue	48.89 ± 8.46	79.10 ± 10.69	≥0.62	25.47 ± 18.80	79.63 ± 11.11	0.32	sub
serum	284.49 ± 48.51	36.69 ± 4.61	≥7.75	207.46 ± 63.00	35.94 ± 5.79	5.77	sub
plasma	18.22 ± 9.31	12.15 ± 4.42	≤1.50	60.51 ± 20.33	11.90 ± 3.25	5.09	sub
urine	16.66 ± 4.32	74.46 ± 2.70	≥0.22	4.83 ± 3.69	62.15 ± 44.48	0.08	sub
thyroid tissue	129.06 ± 61.71	20.57 ± 6.43	≥6.27	39.97 ± 30.32	24.70 ± 7.93	1.62	sub

Table 2. Assignment of interim SPIN based on average SELDI clustered peak intensities

Correlation between SPIN and analytical results: Select assays performed on samples subjected to handling variation revealed a relationship between a sub-optimal SPIN index and inaccurate analytical results as compared to ideal handling conditions, suggesting that SPIN could be utilized to predict sample quality (Figure 4).

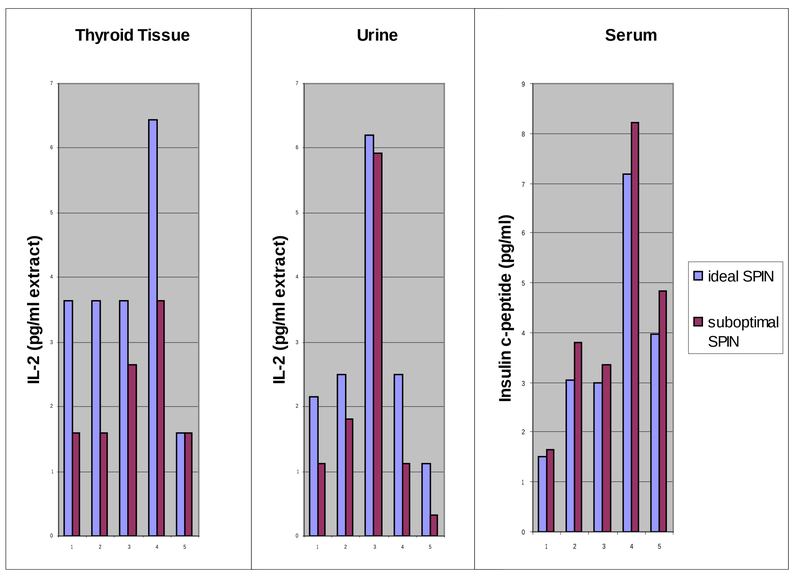


Figure 4: Comparison of analysis of common biomarkers in samples with an "ideal" and "suboptimal" SPIN. Sister aliquots of the same sample were either processed under standard BioBank SOPs (ideal) or subjected to conditions designed to introduce pre-analytical variability (suboptimal).

Discussion

It is widely accepted that variable biorepository specimen handling conditions can significantly alter outcomes of clinical research studies, establishing the need for metrics to determine sample quality. In line with the NIH and the OBBR Best Practices, it is vital that the integrity of specimens used for biomarker studies are of the highest standard to ensure validity of the data they generate and confidence in the application of new findings to clinical management. SPIN can be employed to determine the usability of a sample regardless of the processing procedure. This concept represents an innovative method to utilize specimens currently banked with varying consistency.

The BioBank at William Beaumont Hospital obtains, processes, stores, and analyzes biological specimens of virtually every type of biological tissue or bodily fluid. The process demonstrated will be further developed in order to discover and develop protein integrity indices for each. Once positively identified and quantified by standard laboratory techniques (i.e. MS/MS, ELISA, western blot, or multiplex bead array) absolute ratios will be used for the SPIN index. These indices could be applied universally throughout the biorepository community to standardize quality measurements for subsequent proteomic analyses.

Acknowledgment

We would like to thank the BioBank clinical staff for their assistance in consenting patients and acquiring samples.

This work is supported by the Erb Family and the Beaumont Research Institute.