Fractionation-Dependent Improvements in Proteome Resolution in the Mouse Hippocampus by Isoelectric Focusing

Joseph L. Bundy, Brian D. Inouye, Roger S. Mercer, and Richard S. Nowakowski

Abstract

Mass spectrometry is a tool for investigating the abundance of small molecules including peptides. Mass spectrometry-based proteomics can identify and quantify simultaneously hundreds of proteins in a single biological sample. By pre-fractionating protein extracts with isoelectric focusing (IEF), the number of proteins identified in a single experiment can be increased from hundreds to thousands. However, fractionation also increases analysis time and is cost-prohibitive. Therefore, it is advantageous to ascertain and understand the benefits and drawbacks of IEF when designing large IEF-enabled experiments. To understand the benefits of IEF in an investigation of hippocampal tissue, a systematic analysis of IEF-fractions and pooled fractions was conducted. This analysis focused on improvements in functionally relevant protein identifications, quantitative resolution, and statistical power.

Isoelectric Focusing (IEF)

Isoelectric focusing is a preparative technique by which peptides in a complex mixture are separated based on their respective isoelectric points. This is can be accomplished by running peptides through an polyacrylamide gel with an immobilized pH gradient. When electric current is applied to the gel, peptides will migrate. Peptides will cease migrating when the reach a pH where they no longer have an electric charge. Peptides can then be harvested from separate locations on the gel and subsets of peptides can undergo analysis individually.

Methods

Extract Proteins

- Hippocampus of C57BL/6J mice were dissected, homogenized, and digested with trypsin
- Non-protein contaminants were removed and protein concentrated

Fractionate Peptides

- Tryptic peptides were separated into 12 fractions by isoelectric focusing
- Fractionation of peptides reduces sample complexity and improves proteome coverage

Pool adjacent fractions

- Aliquots of fractionated lysate were pooled into samples comprised of fewer, more complex fractions.

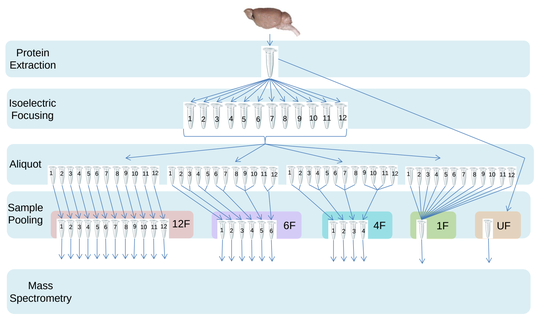
Mass Spectrometry

- Fractionated samples were injected into the Orbitrap Velos mass spectrometer which assesses the mass and charge of individual peptides

Data Analysis

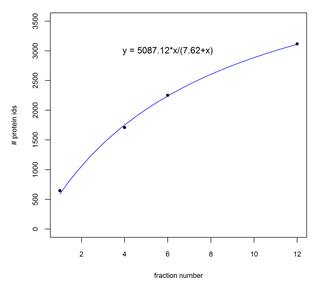
- .raw files generated by mass spectrometer were processed by commercial and open-source software packages

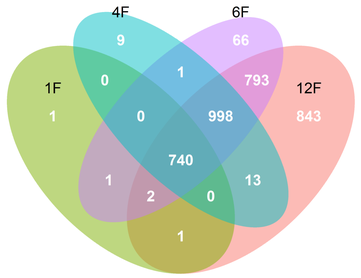
Experimental Design



Samples were derived from aliquots of protein extract that was either unfractionated (UF) or fractionated via peptide IEF from a single mouse hippocampus. One aliquot of all IEF fractions was individually analyzed via LC-MS/MS (12F). The 3 remaining aliquots from the 12 IEF fractions were combined into different sets of pooled samples (6F, 4F, and 1F), which were then analyzed via LC-MS/MS.

Results – Protein Identifications

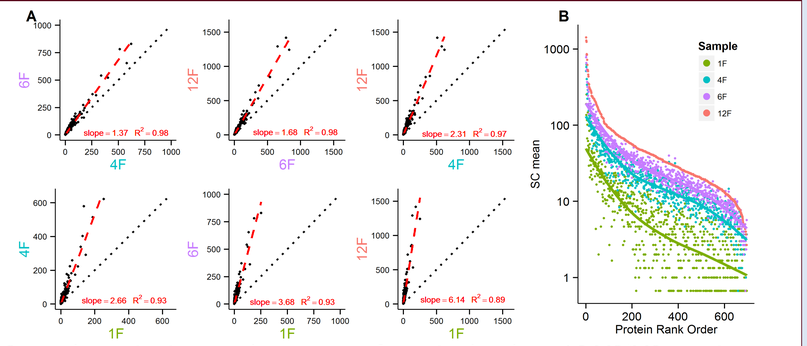




Variable	UF	1F	4F	6F	12F
# LC-MS/MS samples	1	1	4	6	12
# of proteins identified	967	745	1761	2601	3390
# of proteins identified in all replicates	673	420	1240	1907	2504
# of spectral counts	29,034	19,844	56,095	81,743	133,333
Mean protein sequence coverage*	5.61%	4.38%	10.64%	16.72%	21.91%
Cost of LC-MS/MS	$45	$45	$180	$270	$540
# of proteins with GO term "synapse"	78	64	134	176	198
# of proteins with GO term "synaptic vesicle"	38	35	47	57	63
# of proteins with GO term "neurogenesis"	94	84	157	214	251
# of proteins with GO term "neurotransmitter secretion"	14	12	22	25	30
# of proteins with GO term "voltage-gated channel activity"	8	7	17	28	37

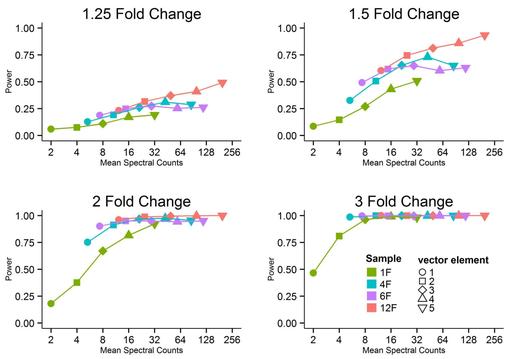
The sample consisting of the greatest number of fractions (12F) identified more proteins than those which consisted of fewer fractions. To better understand the relationship between fraction number and protein identifications, a regression analysis was performed using a saturating model (upper left). Additionally, a set-based analysis of protein identifications reveals that proteins identified in samples consisting of few fractions were generally a subset of those proteins identified in more extensively fractionated samples (upper right). Extensively fractionated samples also identified a greater number of proteins relevant to neuronal function as determined by association of gene ontology terms. Three-fold more neurogenesis and synapse - associated proteins were detected in the fully fractionated lysate (12F) than the most extensively pooled lysate (1F), demonstrating the benefit of IEF pre-fractionation for proteomic investigations of neuronal function and proliferation. However, considerable improvements are also evident in the 4F and 6F samples.

Results – Quantitative Resolution



Spectral counting is a popular measure of protein abundance LC-MS/MS experiments. Low-abundance proteins with few spectral counts are not robustly quantitated. We performed regression analysis on mean number of spectral counts associated with reproducibly identified proteins (A). Extensively fractionated samples identified proteins with more spectral counts than samples with few fractions, reflecting improved quantitation of those proteins. To compare simultaneously spectral counts from all samples, proteins were ranked from greatest to least abundance in 12F, and ranks were plotted against the spectral count mean in each sample (B).

Results – Statistical Power



Proteins quantified with many spectral counts have a lower signal to noise ratio than proteins quantified with few counts. Because extensively fractionated samples identified proteins with more spectral counts, we suspected that pre-fractionation with IEF may improve statistical power in differential expression analyses. Therefore, we performed a power analysis on synthetic data interpolated from regression based relationships in the dataset. Specifically, we investigated how statistical power is affected by spectral count level at a variety of fold-changes. This analysis revealed that differences in protein abundance between two biological conditions are more likely to be achieve significance in extensively fractionated samples than samples comprised of few fractions.

Conclusions

1. IEF pre-fractionation improves the breadth and depth of proteome resolution by increasing the number of protein identifications and their associated spectral counts respectively.
2. Improvements in quantitative depth translate into improved sensitivity to detect differential protein abundance.
3. Our future proteomics experiments on the mouse hippocampus will use the 6F pooling design. Fractionation improves data quality, but the cost to analyze all 12 fractions individually is not commensurate with the improvement in data quality.