ProDM - A Colorimetric Kit for Rapid Monitoring and Standardization of Tryptic Digestion for Improved and Reproducible Shotgun Proteomics

Kutralanathan Renganathan[1], Stephen Russell[1], Steven Wolfe[1], Florentina Mayko[1], Stella B. Somiari[2], Richard I. Somiari[1]

[1]ITSI Biosciences, 633 Napoleon Street, Johnstown, PA 15901, USA
[2]Windber Research Institute, 620 Seventh Street, Windber PA 15963, USA.

Introduction

Proteomics has advanced significantly over the past decade [1], and it is being applied in many laboratories for protein characterization, biomarker discovery/verification and complex disease research [2]. The central objective of shotgun proteomics is to identify as many individual proteins as possible in a single run. The more the number of proteins identified in a sample the higher the possibility of elucidating biologically-relevant mechanisms and pathways associated with the condition under study.

A critical pre-analytical step in sample preparation which can directly affect the success of shotgun proteomics is the digestion of proteins into peptides prior to the mass spectrometry step. Like all enzymes, the specific activity of trypsin towards a protein substrate and completion of digestion will depend on the type of protein substrate, the buffer system used, pH of the buffer, incubation temperature and the overall duration of incubation. Thus, the addition of trypsin to a protein mixture is not a guarantee that the trypsin will adequately digest every protein present and a useful MS sequence coverage will be obtained.

The main objective of the study was therefore to identify a simple, sensitive and practical method that could be used routinely to assess the quality of tryptic digestion. Although there are established methods that can be used to monitor protein digestion, the methods reported, including HPLC, circular dichroism, SDS-PAGE, MS and real-time monitoring with a fluorescent dye [3] are either time and sample consuming or expensive to perform.

Our report shows that ProDM can be used to precisely determine the degree of protein digestion and hence optimize the trypsin digestion step. Optimizing the digestion step will help save time and reduce sample wastage and ultimately increase the success rate of proteomic experiments in many laboratories.

Methods

Blood collection and sample preparation

Blood sample was collected using Proteomics Grade Plasma and Serum Tubes (PGPT/PGST, ITSI-Biosciences). Plasma w/Serum was isolated from the PGPT /PGST tube according to standard plasma/serum isolation protocol, aliquoted in 1 ml amounts to avoid repeated freeze/thaw of the same aliquot and stored at -80ºC until required. Prior to analysis, samples were thawed on ice and total protein content determined for all samples using the ToPA Kit (ITSI-Biosciences). All experiments were performed in duplicate and all results presented are averages of two separate readings.

Tryptic digestion of plasma proteins

Digestion of plasma proteins were performed in a buffer system containing Trifluoroethanol (TFE). Briefly, 50 µl of diluted plasma or serum was mixed with 50 µl of TFE. Sample was reduced with DTT (10 mM) and alkylated with IAA (20 mM). Incubation was for 8h and 24hrs at 37ºC. At the end of the incubation period, 5 % (v/v) Formic acid was added to stop the digestion and the extent of protein digestion was monitored.

Tryptic Digestion Monitoring

Undigested plasma sample and samples digested for 8h and 24h at 37ºC were analyzed with a) ProDM and b) Agilent Bioanalyzer.

ProDM

Samples were analyzed according to the manufacturers' protocol. The ProDM Kit contains *ready-to-use* reagents including a Standard Buffer (Urea-Tris Buffer pH 8.5), Reaction Buffer (Tris-Buffer pH 8.5), Reaction Quencher (Buffered Phosphoric acid) and Colorimetric Reagent. Samples without trypsin (time zero) and samples digested for 8h and 24h were processed. The colorimetric reagent was added and absorbance was read at 595nm within 1 min of adding the color reagent.

% Protein Digested (%PD) = $[(A_{1(T0,\ I)} - A_{2(Tx,\ I)})/ A_{1(T0,\ I)}]*100$,

where $A_{1(T0,\ I)}$ is the absorbance of the first aliquot at time zero and $A_{2(Tx,\ I)}$ is the absorbance of the second aliquot after 8h or 24h digestion.

Agilent BioAnalyzer

50µl of undigested and digested samples were processed using a Agilent Protein 80 chip according to manufactures recommendations.

Protein identification by LC/MS/MS

After digestion with trypsin, the peptides were acidified with 5 % Formic acid and dried down in a Speedvac. The dried sample was reconstituted in 2 % Acetonitrile / 0.1 % Formic acid and loaded onto a PicoFrit C18 nanospray column using a Thermo Scientific Surveyor Autosampler operated in no waste injection mode. Peptides were eluted from the column using a linear Acetonitrile gradient from 2 to 40% over 60 minutes into a LTQ XL mass spectrometer (Thermo Scientific) via a nanospray source with the spray voltage set to 1.8kV and the ion transfer capillary set at 180ºC. A data-dependent Top 5 method was used were a full MS scan from m/z 400-1500 was followed by MS/MS scans on the five most abundant ions [4].

Protein identification and the number of missed cleavages were determined with the Proteome Discoverer 1.3 software as previously described [4]. Trypsin was the selected enzyme and we allowed for up to three missed cleavages per peptide. Carbamidomethyl Cysteine was used as a fixed modification. Precursor and fragment ion peaks were searched with a mass tolerance of 5000 ppm and 2 Da, respectively. Proteins were identified when unique peptides had X-correlation scores greater than 1.5, 2.0, and 2.5 for respective charge states of +1, +2, and +3 [4].

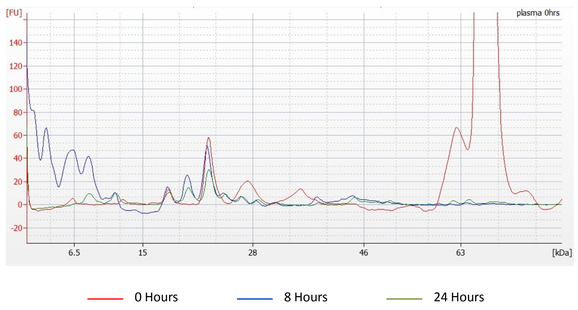

Figure 1A: Agilent Bioanalyzer Protein 80 traces showing the profile of undigested plasma and plasma digested in a buffer system containing ammonium bicarbonate (AMBIC) and Trifluoroethanol (TFE). The presence of visible peaks in Plasma suggests the presence of undigested proteins.

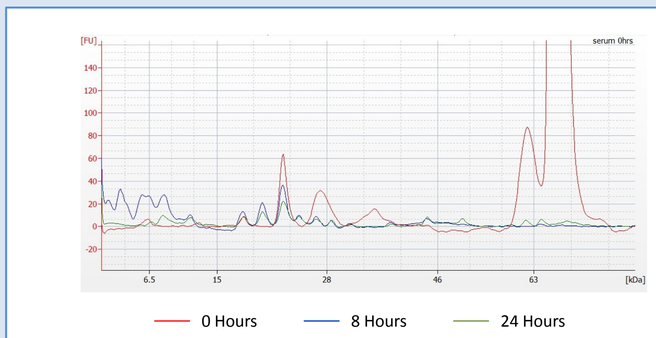

Figure 1B: Agilent Bioanalyzer Protein 80 traces showing the profile of undigested Serum and Serum digested in a buffer system containing ammonium bicarbonate (AMBIC) and Trifluoroethanol (TFE). The presence of visible peaks in Serum suggests the presence of undigested proteins.

Table 1: Percentage (%) of total proteins digested in 8h and 24h in Plasma and Serum measured with the ProDM kit.

Incubation Time (hr)	% Protein Digested	
	Plasma	Serum
0	0	0
8	46.5	46.2
24	56.1	50.2

Table 1 shows that about 46% of proteins were digested in plasma and serum after 8h incubation with trypsin, and this percentage increased to 56% in plasma and 50% in serum after 24h incubation.

Table 2: Total number of unique peptides, proteins identified and albumin sequence coverage in Plasma and Serum after tryptic digestion for 8h and 24h at 37ºC.

	Plasma			Serum		
	8h	24h	% Change	8h	24h	% Change
Number of Proteins	125	118	5.93	127	107	18.69
Number of Peptides	991	733	35.20	897	693	29.44
Serum Albumin Coverage	79	74	6.76	77	66	16.67

Remarkably, fewer numbers of peptides were identified after 24h compared to 8h in both plasma and serum. Similarly albumin sequence coverage is also reduced when 8h and 24h incubation is compared.

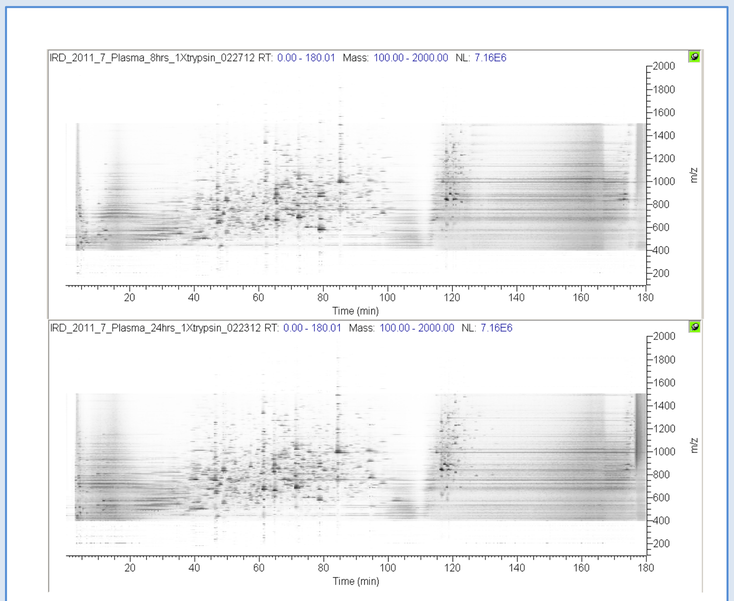

Figure 2 2D-density plot consisting of time (X-axis) vs. m/z (Y-axis) vs. relative abundance (Z-axis). Figure 2 A (upper figure) and B (lower figure) shows only subtle differences between 8h and 24h samples indicating good reproducibility between the mass spectrometry runs.

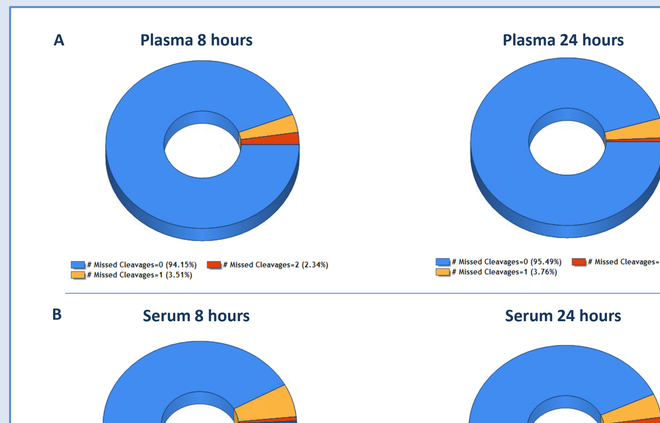

Figure 3: Number of misses cleavages in plasma and serum proteins digested for 8h and 24h with trypsin.
The number of peptides with zero missed cleavages in Plasma **(A)** or Serum **(B)** increased after 24 h compared to 8h indicating the progression of digestion. However, the number of proteins identified decreased in 24 h (**Table 2**).

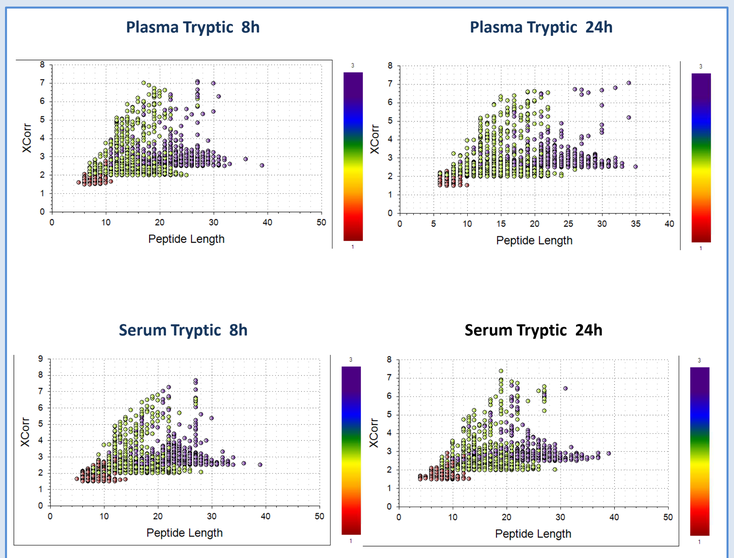

Figure 4 : Distribution of peptide length vs. Xcorr. +3 and +1 charged peptides had relatively lower Xcorr scores compared to +2 charged peptides. 24 h trypsin incubation produces more +1 and +3 charged peptides compared to +8h digestion time. Peptides with 8 – 20 amino acids seems to offer high Xcorr scores and 24 h digestions seems to produce peptides longer than 20 AA compared to 8h leading to less confident peptide identifications [5].

Table 3: Charge State Distribution at 8h and 24 h digestion

Plasma			Serum		
Charge State	% in 8h Incubation	% 24 h Incubation	Charge State	% in 8h Incubation	% 24 h Incubation
1	0.93	1.99	1	0.76	3.2
2	72.43	68.16	2	85.50	82.4
3	26.64	29.85	3	13.74	14.4

24h shows more +3 charged peptides compared to 8h and +3 charged peptides show relatively lower Xcorr compared to +2 charged peptides on a LTQ-XL machine (Figure 4).

Conclusions

- Our results demonstrates that ProDM can be used to precisely monitor protein digestion prior to mass spectrometry.
- The result is important because it shows that the duration of trypsin digestion apparently affects a) the total number of proteins identified and b) protein sequence coverage.
- This is particularly critical for applications like quantitation and biomarker identification, where reproducibility is critical.
- Optimizing the trypsin digestion step will ensure that sample and equipment time are not wasted because of the need to repeat experiments or the inability to interpret results.

References

1. Klammer, A., MacCoss, M. *J Proteome Res 2005,* 5, 695–700.
2. Pastwa, E., Somiari, S.B., Czyz, M., Somiari, R.I. *Proteomics Clin Appl* 2007, 1, 4-17.
3. Karuso, P., Crawford, A.S., Veal, D.A., Scott, G.B., Choi, H.Y. *J Proteome Res* 2008, 7, 361–366.
4. El-Bayoumy, K., Das, A., Russell, S., Wolfe, S., *et al. J Proteomics* 2012, 75(3), 1018-1030.
5. Rudnick P A, Clauser, K R, Kilpatrick L E, Tchekhovskoi D V, Neta P, Blonder N, and Stein S E. *Molecular & Cellular Proteomics*, 2010, 9(2), 225-241.