Development of a Synthetic Protein Quality Control (QC) Standard for the Assessment of Sample Proteolysis Reproducibility

Bryan Krastins[1], Amol Prakash[1], Michael Athanas[2], David Sarracino[1], Mary Lopez[1]
[1]Thermo Fisher Scientific, Cambridge, MA, USA ; [2]Vast Scientific, Wayland, MA, USA

Overview

Purpose: Develop a synthetic protein (concatenated peptide) that can be used to optimize and monitor proteolytic digestion prior to selected reaction monitoring (SRM) or other mass spectrometry (MS) assays.

Methods: A long, concatenated peptide was created with multiple cleavage sites varying in accessibility as determined by single or multiple arginines and lysines.

Results: The data demonstrate that the accessibility of the digestion sites (as determined by single or multiple arginines or lysines) had a significant effect on the optimal incubation times for digestion. Therefore, the concatenated peptide can be useful as a quality control (QC) tool for monitoring digestion efficiency in experimental samples before quantitative SRM and other MS-based assays.

Introduction

In many proteomic applications, proteins are subjected to proteolytic digestion. Following digestion, peptides are often analyzed by a mass spectrometer resulting in qualitative or quantitative data. Although enzymatic protein digestion is broadly used, the results may be widely variable depending on sample, enzyme-to-protein ratio and digestion conditions. Because enzymatic digestion is the starting point for many proteomic experiments and assays, inconsistent sample preparation can compromise the downstream analysis of the resulting data, especially in SRM assays where accurate quantification of peptides is crucial. Therefore, a thorough understanding of proteolytic digestion efficiency during the sample preparation process is critical to the success of these methods.

Methods

Samples

Human serum samples were collected from a donor with full consent and approval.

Synthetic Peptides

A concatenated synthetic protein which could be used to assess the quality and efficiency of proteolytic digestion during sample preparation was developed. The 4707 Da protein (see Figure 1) contained 42 amino acids, eight of which were heavy isotope-labeled (either arginine or lysine). The protein was purified by HPLC to >97% purity and stored frozen in solution at a 5 pmol/µL concentration. Following complete proteolytic digestion, the synthetic QC protein produced exactly four proteolytic peptides. In addition, light and heavy isotope-labeled versions of the resulting proteolytic peptides were synthesized. The proteolytic QC protein was tested by adding a known concentration to a serum sample and then subjecting the mixture to enzymatic digestion followed by mass spectrometry analysis in a triple quadrupole SRM-based assay.

Proteolytic Digestion

Proteolytic digestion was carried out as previously described[1, 2] except three different buffer compositions (Table 1) were tested, and aliquots were taken at 1, 4, and 24 hours for SRM assays.

SRM Assay Development

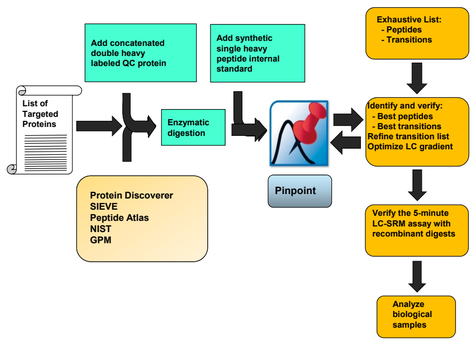
SRM assays were developed as previously described[1, 2] and as shown in Figure 1B. The initial step is input of a target protein(s) sequence. Prediction algorithms create an exhaustive list of peptides and transitions. In addition, previous spectral data from discovery experiments can be mined to facilitate the choice of optimum peptides and transitions. Iterative cycles using a recombinant protein digest or cocktail of synthetic peptides are used to optimize the assay conditions automatically. The efficiency of enzymatic digestion can be monitored by the initial addition of the concatenated peptide with post-digestion addition of isotopically-labeled single heavy peptides and the calculation of the ratios between the single and double heavy labeled peptides.

FIGURE 1A. Sequence and properties of the synthetic protein. 1B. SRM assay development using Thermo Scientific Pinpoint software and concatenated peptide for digestion QC.

A. TGISPLA*LIKSS*ALFYQ*KKKVTI*ASLP*RRRF*LVGPDGIPIM*R

- Long concatenated protein was synthesized
- Extra lysines and arginines make digestion more difficult
- Theoretical proteolytic digestion (to completion) produces four peptides
- Heavy versions of the single peptides were also synthesized
- Spiking concatenated peptide before digestion into sample allows quantification of the digest
- Ratio of light to heavy peptides provides precise quantification

B.



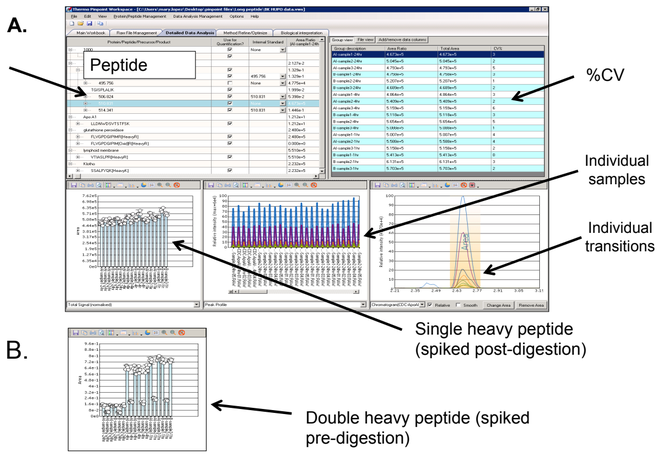
Results

Serum samples containing the synthetic protein (Figure 1A) were subjected to enzymatic digestion. In order to understand the efficiency of proteolysis, each sample underwent a time course experiment in three different buffers, as shown in Table 1. Aliquots of each sample were taken at the following time points during digestion: 1 hr, 4 hr and 24 hr. At each time point, an aliquot of sample was removed for triplicate analysis. Calibration curves were generated using light and heavy-labeled versions of the peptides from the proteolytic protein and the LOQ and LLOD in serum matrix were calculated . The LLOD of the synthetic peptides ranged from 500 attomoles to 1 femtomole (data not shown). The addition of the proteolytic QC protein to the SRM assay experiments provided increased confidence in the quantitative measurements for targeted proteins in the experimental sample.

Figures 2-5 illustrate that the accessibility of the enzymatic digestion sites (as determined by single or multiple arginines or lysines) had a significant effect on optimal digestion time. Peptides with easily accessible sites such as TGISPLALIK, SSALFYQK and FLVGPDGIPIMR had optimal digestion times of 1-4 hr with recoveries of 70-80%. Peptide VTIASLPR, which is bordered by 2 lysines and 2 arginines, and would therefore be more difficult to digest, had an optimal digestion time of 24 hr. However, a digestion time of 4 hr in sample buffers 1 or 3 was sufficient to deliver recoveries of 60-70% for all the peptides. Figure 6 demonstrates that buffer composition (Table 1) has a dramatic effect on digestion. Buffer containing SDS reduced digestion efficiency by more than 50% in all cases, even when SDS was completely removed prior to the addition of a proteolytic enzyme.

FIGURE 2. Digestion efficiency *versus* time for peptide TGISPLALIK. The single heavy peptide data (A) show quantitative consistency of the internal single heavy peptide standard added after digestion. Percent CV ranged from 0-6. Digestion efficiency for the double heavy peptide (B) was greatest, ~ 70-80%, at 1 hr in sample buffers 1 and 3 .

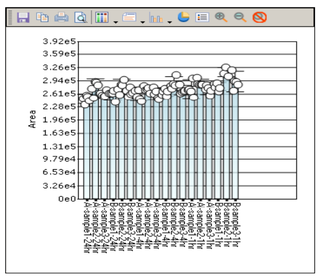
A.



B.

FIGURE 3. Digestion efficiency *versus* time for peptide VTIASLPR. The single heavy peptide data (A) show quantitative consistency of the internal single heavy peptide standard added after digestion. Digestion efficiency for the double heavy peptide (B) was greatest, ~ 60-70%, at 24 hr in sample buffers 1 and 3.

A. Single heavy peptide (spiked post-digestion)



B. Double heavy peptide (spiked pre-digestion) ratio to single

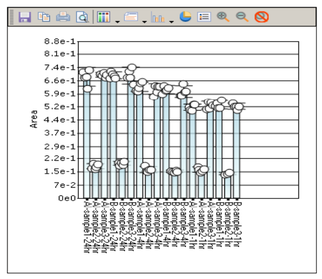
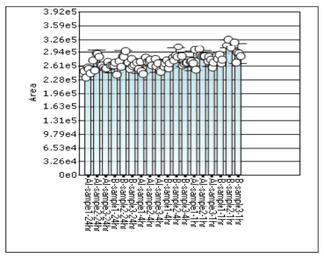


FIGURE 4. Digestion efficiency *versus* time for peptide FLVGPDGIPIMR. The single heavy peptide data (A) show quantitative consistency of the internal single heavy peptide standard added after digestion. Digestion efficiency for the double heavy peptide (B) was greatest, ~ 60-70%, at 4 hr in sample buffers 1 and 3.

A. Single heavy peptide (spiked post-digestion)



B. Double heavy peptide (spiked pre-digestion) ratio to single

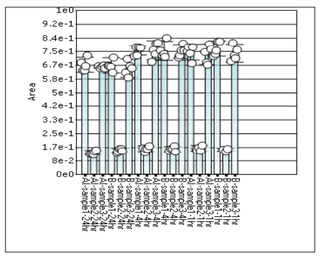
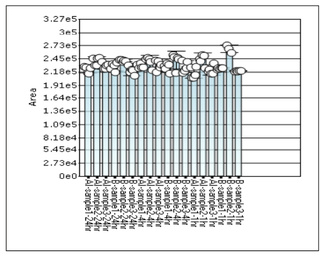


FIGURE 5. Digestion efficiency *versus* time for peptide SSALFYQK. The single heavy peptide data (A) show quantitative consistency of the internal single heavy peptide standard added after digestion. Digestion efficiency for the double heavy peptide (B) was greatest, ~ 60-70%, at 1 and 4 hr in sample buffers 1 and 3.

A. Single heavy peptide (spiked post-digestion)



B. Double heavy peptide (spiked pre-digestion) ratio to single

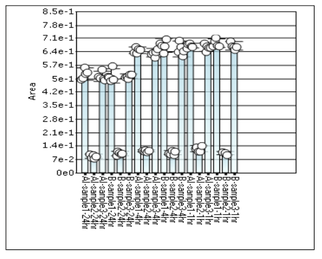
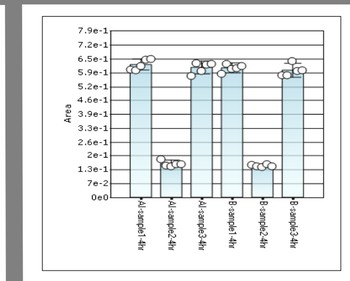


TABLE 1. Buffer compositions for enzymatic digestion. Three buffers (see Methods) were compared for digestion efficiency.

Sample Name	Sample 1 (SOP)	Sample 2	Sample 3
Buffer Composition	8M Urea 2.5% n-propanol 200mM Tris 10mM DTT pH 8.5	8M Urea 0.5% SDS 200mM Tris 10mM DTT pH 8.5	Remove lipids Mix 12.5ul of delipidated serum sample with 50ul of SOP

FIGURE 6. Effect of buffer composition on digestion efficiency. Three buffers (see Methods) were compared for digestion efficiency.



Conclusion

- The concatenated synthetic protein provided a simple, straightforward way to measure the efficiency of proteolytic digestion prior to SRM assays.
- Longer incubation times were required to produce cleavage of “difficult” sites such as those with multiple arginines or lysines.
- However, longer incubation times reduced recovery of peptides bounded by “easy” sites (single arginines or lysines).
- Four-hour incubation times are sufficient to provide cleavage at all sites with an overall efficiency of 60-80% and minimize peptide losses.
- Buffer composition is important to ensure optimum digestion efficiency. Buffer containing SDS with subsequent SDS removal reduced efficiency dramatically.

References

1. Lopez, M.F.; Rezai, T.; Sarracino, D.A,; Prakash, A.; Krastins, B.; Athanas, M.; Singh, R.J.; Barnidge, D.R.; Oran, P.; Borges, C.; Nelson, R.W. Selected Reaction Monitoring-Mass Spectrometric Immunoassay Responsive to Parathyroid Hormone and Related Variants. *Clin Chem.* **2010** 56:281-90.
2. Lopez, M.F.; Kuppusamy, R.; Sarracino, D.A.; Prakash, A.; Athanas, M.; Krastins, B.; Rezai, T.; Sutton, J.N.; Peterman, S.; Nicolaides, K. Novel MS discovery-to-targeted SRM workflow incorporating ROC curve analysis of putative biomarker candidates: Application to first-trimester maternal blood from Trisomy 21 pregnancies. *J Proteome Res.* **2010** Jun 4 E pub ahead of print.