Automating the Retention Time Determination for Targeted Peptide Detection/Quantitation: Building Flexibility into Scheduled SRM Assay Development

Scott M. Peterman[1], Amol Prakash[1], Daniela Tomazela[2], Barbara Frewen[2], Brendan MacLean[2], and Michael J. MacCoss[2] • [1]BRIMS Center, Thermo Fisher Scientific, Cambridge, MA and [2]Department of Genome Sciences, University of Washington, Seattle, WA

Overview

Purpose: To demonstrate a flexible method of predicting retention times for targeted peptides in a timed SRM assay.

Methods: Identify a well characterized set of peptides classified as a trainer set used to characterize the linear relationship between hydrophobicity factors and retention time .

Results: The calculated linear response was used to predict retention time windows (±4 min) that identified 83% of the 262 targeted plasma peptides in a scheduled SRM assay.

Introduction

The role of targeted protein/peptide detection via SRM-based assays has increased over the past year from primarily quantification to discovery and verification of putative biomarkers. Advances in determining proteotypic peptides and SRM transition information based on bioinformatics have resulted in larger numbers of peptides being monitored.[1] The large SRM transition list requires scheduling SRM transitions to maintain analytical relevance. The difficulty associated with method development now lies in determining RT information. Previous attempts to build timed SRM experiments simply transfer the entire experimental method used from discovery experiments acquired on a different mass spectrometer. This approach requires precise duplication of the LC experimental conditions or the use of wide retention time windows to ensure that the targeted peptide is detected in the timed SRM experiment. Additionally, any deviation in experimental details must be eliminated without the need to acquire the samples on the second instrument using the desired conditions, which limits the flexibility to change the chromatographic gradient conditions to increase throughput. Our approach is to use a well-characterized trainer set of peptides to provide a means of determining the linear relationship between hydrophobicity factors and retention times for the peptides used in the trainer set.[2,3] The mathematical expression can then be applied to the list of targeted peptides to identify retention times and set retention time windows according to the gradient length of the experiment.

Methods

Samples

The initial set of trainer set peptides were obtained from Michrom standard digests samples commercially available (Michrom Bioresources, Auburn, CA). The stock solutions were diluted to 10 fmol/µL and injected on column. At least 10 peptides per digest were used in the initial trainer set resulting in 92 peptides. The enolase digest sample was used as a trainer set for the plasma protein digest study. Seven peptides were selected that had a range of MW and hydorphobicity factors that spanned those values included from the plasma protein digest. A human plasma protein digest was used to identify 261 peptides. Each of the peptides were previously identified and contained in an LTQ-CID spectral library.

Instrumentation

All experiments were performed using a Thermo Scientific TSQ Vantage™ triple quadrupole mass spectrometer operated in SRM mode. SRM transition information (e.g. precursor/product ion information and Q2 collision energies) was determined using prototype software. The calculated hydrophobicity factors used in the study were based on the values and methods proposed by Krokhin et al. Both samples were analyzed using a standard LC-MS/MS method with no scheduling as well as acquired using a timed SRM acquisition. For the acquisition of the 261 plasma peptides, three different injections were used to evaluate detection efficiency and experimentally determined RT. The errors reported in the poster were based on the difference between the predicted and the experimentally determined RT values.

HPLC Conditions

All injections were performed using a Eksigent nanoLC-Ultra™ system (Dublin, CA) and used with a binary solvent system comprised of A) 0.1% formic acid and B) MeCN (0.1% formic acid). All experiments were performed using a 150 x 0.1 mm Halo column (Michrom Bioresources, Auburn, CA). Initial studies on the standard protein digests used three different gradients 30, 60, and 90 minutes running from 5 → 40% B prior to holding. The plasma protein digest analysis used a 60 minute gradient.

Data Processing

All data was processed using the Thermo Scientific SRM Workflow software. The resulting data was processed with the software to determine peak shapes, RT, ion ratios, and reproducibility.

Results

The first step of the described approach is determining the elution time reproducibility for proteotypic peptides. Figure 1 shows the retention time reproducibility (±5 seconds) for the yeast ADH peptide IGDYAGIK for 5 technical replicates using a 60 minute gradient. Knowing the retention time shift enables more confidence to be placed on the retention time windows used for timed SRM assay construction for targeted peptides. The second step is to identify a set of peptides from a well characterized digest sample that can be used to establish the relationship between calculated hydrophobicity factors and retention time. For the experiments shown here, the 5 yeast equal molar digest sample was used. Figure 2 shows an example of 7 trainer set peptides selected from earlier experiments used to determine the linear relationship for a 30-minute gradient.

To test the approach three different gradients were used to monitor 92 peptides from the yeast protein digest standards. Three different gradients were used to test the relationship of hydrophobicity factor and retention times for the ten yeast peptides. Figure 3 shows the response for the 30, 60, and 90 minute gradients and the calculated R^2 value was ca. 0.98 for each. Following the trainer peptide detection, a method used to monitor the 92 peptides from the Michrom digest standards were analyzed using the same gradients

The difference between the experimentally determined RT and theoretical RT based on the linear response for the targeted peptides for the three gradients are shown in Figure 4. Percent difference was used to evaluate the RT errors for the three gradients. A 15% RT error was deemed acceptable for this study as it would represent a ±2.6, 3.6, and 5.3 minute window for the 30, 60, and 90 minute gradients, respectively. For the 92 peptides, 91% of them had RT errors less than 15% while the percentage of targeted peptides fell to 83 and 66% for the 60 and 90 minute gradient studies.

FIGURE 1. Chromatographic reproducibility of a targeted peptide for 5 technical replicates.

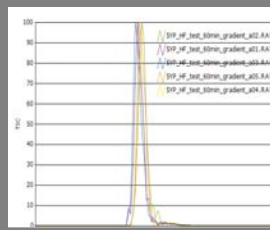
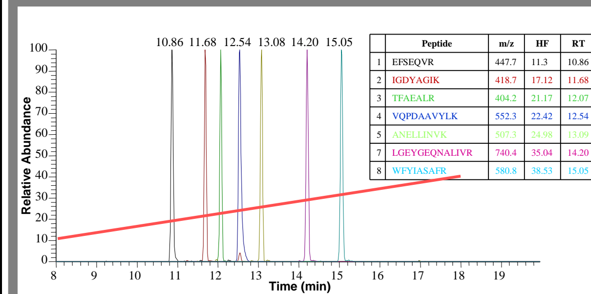


FIGURE 2. Chromatographic relationship for a set of known peptides from a yeast protein digest mixture standard. The inset table shows the peptide sequence, the hydrophobicity factors and the measured retention time for a 30 minute gradient.



	Peptide	m/z	HF	RT
1	EFSEQVR	447.7	11.3	10.86
2	IGDYAGIK	418.7	17.12	11.68
3	TFAEALR	404.2	21.17	12.07
4	VQPDAAVYLK	552.3	22.42	12.54
5	ANELLINVK	507.3	24.98	13.09
7	LGEYGEQNALIVR	740.4	35.04	14.20
8	WFYIASAFR	580.8	38.53	15.05

FIGURE 3. Plotting standard peptide retention times as a function of hydrophobicity factors for three different chromatographic gradients. The blue circles represents a 30 minute gradient, the red squares represents a 60 minute gradient, and the black triangles represents a 90 minute gradient. The three experiments were acquired using the same column and mobile phases.

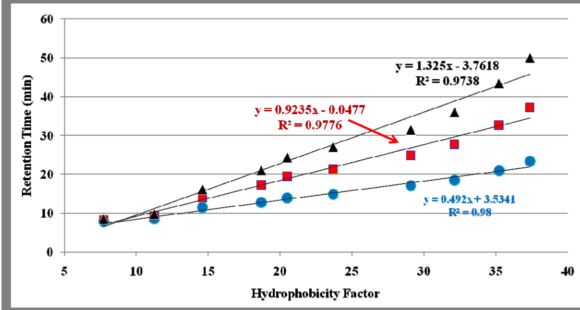
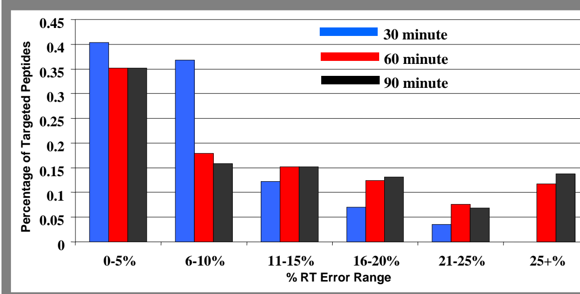


FIGURE 4. Comparison of retention time error for the standard proteins for the yeast protein digest mixture for the three chromatographic gradients.



The study was expanded to monitor a large peptide set from a human plasma digest. The peptides were selected from a spectral library previously acquired on a Thermo Scientific LTQ Orbitrap™ hybrid mass spectrometer. A 60-minute gradient was used to monitor the targeted peptides following the acquisition of the enolase trainer set. A small sample of the hydrophobicity distribution for 84 of the 262 targeted peptides is displayed in Figure 5. In addition to the plasma peptides, 7 enolase peptides used to determine the linear response is provided. A bulk of the targeted peptides fall within a narrow hydrophobicity factor range of 12 to 38. The enolase trainer set peptides and are shown to cover this HF range to provide a reasonable estimation of the RT for the targeted plasma peptides. To identify the experimentally determined retention time for the plasma peptides, the 786 SRM transitions (3 per peptide) were initially acquired using 3 different injections. Following this analysis, the enolase trainer set was analyzed to determine the linear equation applied to the peptides to estimate the RT used in building a timed SRM method. A ±4 minute window was used for acquisition.

FIGURE 5. Hydrophobicity factors for a set of targeted human plasma peptides acquired using a scheduled SRM assay. The black triangles represents a set of enolase peptides used as the trainer set to establish the relationship between hydrophobicity factors and retention time.

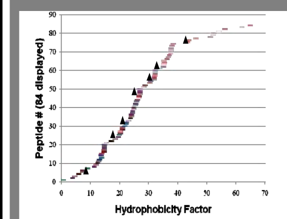
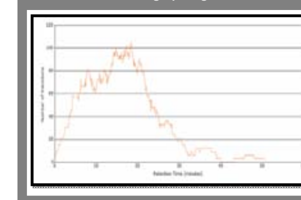


Figure 6 shows the estimated breakdown for monitoring the number of SRM transitions per unit time based on the number of total SRM transitions, the specified RT, RT window, and cycle time contained in the experimental method (2 sec). The shortest dwell time would occur ca. 19 minutes at be set at 9 msec.

FIGURE 6. Screen capture from SRM Workflow showing the distribution of SRM transitions across the chromatographic gradient.



The results from the plasma peptide analysis are shown in Figure 7. The colored peaks represent the enolase peptides that were also spiked in and monitored for RT overlap and are shown to bracket the majority of the plasma peptides. As predicted in Figure 6, the most congested elution time is between 15 and 30 minutes.

FIGURE 7. Chromatographic trace for the targeted plasma peptides acquired using predicted retention time windows for a scheduled SRM experiment. The colored peaks represents the overlaid enolase peptide trainer set used to determine the slope and intercept for determining the ±4 minute windows around the calculated retention time for each targeted peptide.

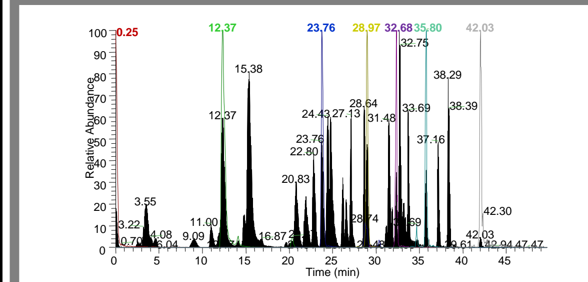
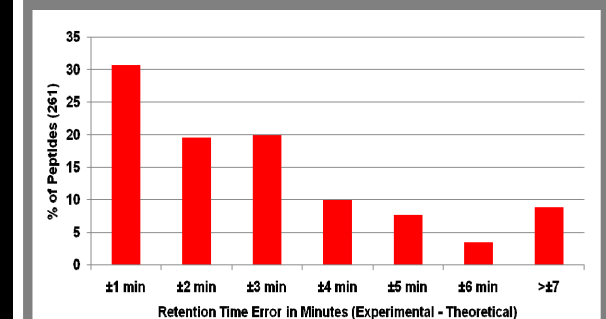


Figure 8 provides a view of prediction accuracy in the study. A total of 82% of the targeted peptides fell within the predicted ±4 minute time window. Within this group, 32% of the peptides showed better than ±1 min. Of the 18% that fell outside of the predicted window, 37% of the peptides had earlier RT than predicted and 62% of the peptide had sequence lengths in excess of 10 residues.

FIGURE 8. Retention time error distribution for the 261 targeted plasma peptides using a 60 minute gradient.



Conclusions

To help automate large-scale targeted peptide discovery/quantification experiments, a method to estimate the retention time is presented. A set of well characterized peptides from a known protein digest was used to establish the linear relationship between experimentally determined retention times and hydrophobicity factors. Once the slope, intercept, and retention time window were determined, a timed SRM assay can be built to accommodate hundreds of targeted peptides contained in a spectral library or database without having to first experimentally determine the retention times. This approach was shown to identify 82% of 261 human plasma peptides with a ±4 minute window around the predicted RT on a 60-minute gradient.

Future direction will be to model the failed and successful peptides to build up more robust predictive models. Also, library data will contain more experimental details showing the measured retention times and gradient conditions that can be used to help build better models and identify well behaved peptides to target for each selected protein.

References

1. Prakash, A., Tomazela, D. , Frewen, B., Merrihew, G., MacLean, B., Peterman, S., and MacCoss, M. Expediting the Development of Targeted SRM Assays: Using Data from Shotgun Proteomics to Automate Method Development, *submitted to J. Proteome Research*

2. Krokhin, V., Craig, R., Spicer, V., Ens, W., Standing, K., Beavis, R., Wilkins J., An improved model for prediction of retention times of tryptic peptides in ion-pair reverse-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Molecular and Cellular Proteomics* **2004** *3,* 908-19

3. Klammer A., Yi, X., MacCoss, M., Noble, W.. Improving tandem mass spectrum identification using peptide retention time prediction across diverse chromatographic conditions, *Anal. Chem.*, **2007**, *79(16)*, 6111-8

nanoLC-Ultra is a trademark of Eksigent Technologies LLC Corporation. All other trademarks are the property of Thermo Fisher Scientific, Inc. and its subsidiaries.