Implementation of MS1 based real–time retention time adjustment on a hybrid Orbitrap mass spectrometer for directed and targeted proteomics

Markus Kellmann, Yovany Cordero Hernandez, Hanno Resemann, Philip Remes, Jesse Canterbury, Will Barshop, Julia Kraegenbring, Heiner Koch

Thermo Fisher Scientific, Bremen, Germany

Abstract

Purpose: Implementation of real-time retention time alignment on a modified Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer without spiking standards.

Methods: MS1 spectra from a reference experiment are stored and compared with spectra during subsequent experiments to perform time alignment.

Results: Excellent agreement between retention time shifts measured and those estimated by the procedure

Introduction

Targeted proteomics studies with large cohorts often suffer from retention time instabilities caused by column aging, especially when low flow HPLC is utilized. To compensate for this the windows for scheduling the targets are set to larger windows, which in turn reduces the number of targets measurable in a meaningful way.

Real-time correction of retention times provides the possibility to get reproducible identification and accurate quantification of many peptide targets during a LC run regardless of column aging effects which usually result in instabilities of the retention times in large cohorts. Several solutions have already been developed and implemented to tackle this drawback. For example, Orbitrap Exploris mass spectrometers have implemented a landmark-based method that utilizes the Thermo Scientific™ Pierce™ Retention Time Calibration Mixture (PRTC). Although this method is simple and effective, it has the disadvantage that the retention time is corrected at relatively long intervals (distances between the eluting peptides), measuring time is "lost" due to the monitoring of the standards and the costs for the synthetic peptides. A robust method without the need of spiking synthetic standards has been presented by Remes et al.[1] and was recently commercialized in a different flavour on the Thermo Scientific™ Stellar™ mass spectrometer. This method utilizes cross-correlations at MS1 level between a reference run and current measurements for calculating an estimated RT shift and applies this correction in real-time to the scheduled target list (fig.1 & fig2). Here we show the implementation on a modified Orbitrap Exploris 480 mass spectrometer.

Materials and methods

Sample

Thermo Scientific™ Pierce™ HeLa Digest Standard 20 µg/vial was reconstituted by adding 200 µL of 0.1 % Formic acid A. The vial was subsequently sonicated for 5 min, followed by multiple sample aspiration and release cycles with a pipette to dissolve it completely. Pierce™ Retention Time Calibration Mixture was spiked to a final concentration of 50fmol/µL spiked into 100 ng/µL of Pierce HeLa protein digest standard.

MS

A modified Orbitrap Exploris 480 mass spectrometer was operated under the following conditions:

Table1. MS settings for experiments

		Reference generation	Alignment runs
Experiment	**Setting**	**Value**	**Value**
MS (RTalignment)	RES:	15k	15k
	Mass Range:	350-1200	350-1200
	AGC:	100% (1e6)	100% (1e6)
	maxIT:	auto	auto
	RealTimeAlign:	rtbin generation	RT align
MS	RES:	60k	60k
	Mass Range:	350-1200	350-1200
	AGC:	200% (2e6)	200% (2e6)
	maxIT:	auto	auto
ddA	RES:	15k	15k
	Mass	350-1200	350-1200
	AGC:	100% (1e5)	100% (1e5)
	maxIT	auto	auto
	Include List:	-	14 PRTC Peptides
	TopN:	cycleTime: 2s	cycleTime: 2s

To use RT alignment, a regular DDA experiment (MS1 + ddA) must be preceded by an additional MS experiment (*RT alignment*, see Table 1). This additional MS experiment is used only for inter-run RT alignment. However, it is sufficient to set a relatively low resolution that does not significantly affect the overall scanning speed.
A reference run serves as a reference point for all subsequent sample runs. The additional MS experiment used for the reference and the samples must contain identical settings for some method parameters, in particular for the mass range, since this parameter is essential for the RT correction.

Chromatography

Samples were separated on a Thermo Scientific EASY Spray™ column ES906 (particle size: 2µm, ID:150µm, length: 150mm) using a Thermo Scientific Vanquish Neo™ UHPLC system. The Vanquish Neo system was operated in a direct injection configuration. 1µL (100ng) was injected analysed using different 30 min gradients (fig.1) with 1µL/min flow. The column with internal emitter was inserted into an Thermo Scientific EASY-Spray™ Source.

Figure 1. Different gradients to mimic column instabilities for RT alignment validation

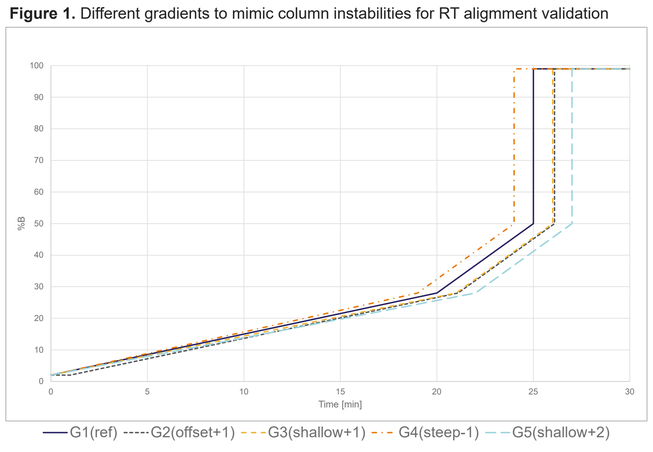


Table1. Gradients G1-G4

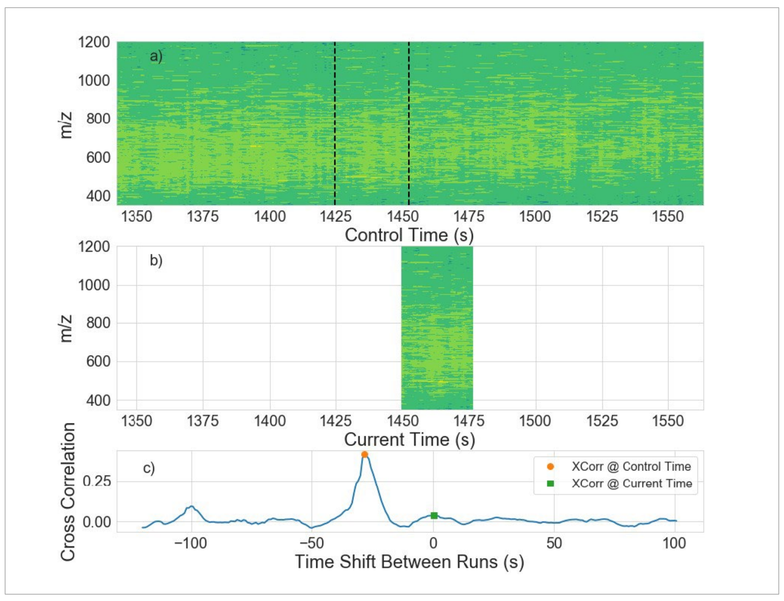
G1(ref)		G2(offset+1)		G3(shallow+1)		G4(steep-1)	
min	%B	min	%B	min	%B	min	%B
0	2	0	2	0	2	0	2
20	28	1	2	21	28	19	28
25	50	21.1	28	26	50	24	50
25	99	26.1	50	26	99	24	99
30	99	26.1	99	30	99	30	99
		30	99				

Implementation

The method is based on a periodic acquisition of MS1 scans in a reference LC-MS(MS) experiment followed by real-time alignment to those scans during subsequent runs. We implemented the method by adding an MS1 experiment to the experiment setup that is solely responsible for real-time alignment, either during reference generation or during the alignment task. During reference generation, a binned and compressed spectral matrix is generated which is used for fast cross correlation with the spectra that were generated during the alignment runs. For alignment, every incoming scan is binned and compressed the same way as the reference.

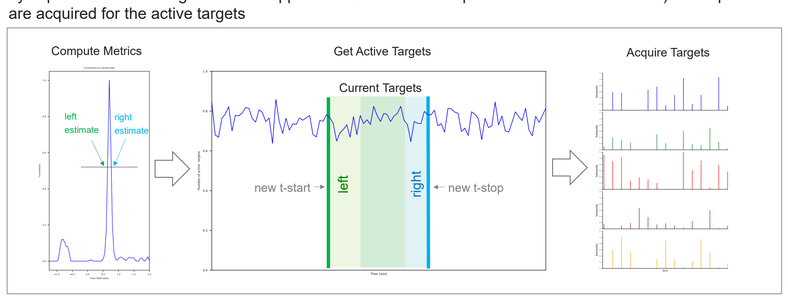
Cross-correlations are now calculated between the current spectrum and successive spectra of the reference over a defined period of time (e.g. +-10min). The time difference between the current time and the maximum of the cross-correlation defines the chromatographic shift (fig.2).

Figure 2. a) Binned and compressed reference run utilizing MS1 spectra b) Real-time region at current time for cross correlation c) correlation curve of real time spectra at current time and retention time span in the reference. The difference between retention time with max. Xcorr and current time describes the RT-shift



To increase robustness and account for small shifts between runs the implementation provides ‚uncertainty bound'. The ‚uncertainty bounds' correspond to the right and left sides of the correlation peak at FWHM. These right and left estimates are applied to the start and stop retention times in an inclusion list and the superset is used for determining the active set of targets in a scheduled targeted or directed experiment (fig3).

Figure 3. a) Cross correlation curve with left and right estimates at FWHM. b) active targets are determined by superset of left and right estimate applied to Rtstart and Rtstop values in an inclusion list c) MS2 spectra are acquired for the active targets



Validation experiments

Reference generation

The reference generation run was used to generate the binned and compressed reference file required for the alignment and to determine the retention times of the PRTC peptides that serve as targets during the alignment runs. For this purpose, the full MS1 data from experiment 1 was used for automatic generation of the reference file (*.rtbin) for later use in real-time correction experiments. Apex retention time of the PRTC peptide were determined manually in Thermo Scientific Freestyle software.

Real-Time-alignment of PRTC standards

Initial evaluation of the procedure was done by monitoring the shift of the PRTC peptides in different gradients. For this, a logfile providing detailed information about the retention time alignment was used to capture the left and right estimates along the run. For determining shifts of the PRTC peptides, apexes were again manually determined and shifts relative to the reference were calculated. Data processing and visualization were done using Jupyter Notebooks showing excellent alignment of actual peak shifts vs. retention time shift estimations of the procedure (fig X).

Real-Time-alignment of Hela digest peptides

After achieving the first proof of principle with PRTC peptides, we wanted to prove that the alignment also works out for a larger set of peptides within a real proteomics sample. For that purpose, we used the same raw data file, processed all different gradients including the reference run with Proteome Discoverer and used the intersection of the identified unique peptides between reference and alternative gradient as test set. The corresponding m/z values were used together with the automatically determined apexes for the scheduled inclusion list.

Again, the logfile generated in real-time was used to extract the right and left estimates provided by the method, but this time these were applied to different RT windows on the inclusion list to simulate the new (aligned) retention time windows for the alternating gradients.

Now, we can directly assess whether the new suggested window would cover the peptide at its shifted position. Fore that, we calculated the effective retention time window after alignment, since the introduction of ‘uncertainty bounds' slightly broadens the scheduling window intrinsically. To evaluate the performance of the new implementation, we calculated the percentage of coverage of the test set's peptide signal apexes within the new scheduling window.

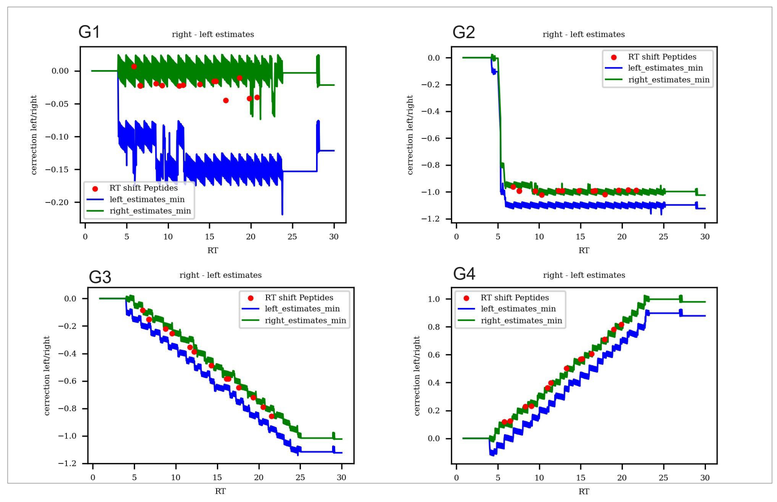
Since processing speed is critical for real-time applications, we monitored the processing speed of each alignment closely.

Validation results

Real-Time-alignment of PRTC standards

Figure 4 shows the overlay between shift estimations and actual peak shifts of 14 PRTC peptides. For all tested gradient shifts the slope of the PRTC signals and the estimated correction is very similar, which indicates that the alignment works very nicely for those examples. Additionally, the apexes are covered nicely between the uncertainty borders suggesting small inclusion windows can be applied.

Figure 4. Overlayed retention time shift estimations for gradients G1 to G4 exemplifying scenarios applying the reference gradient (G1), RT offset (G2), shallow gradient (G3) and steep gradient (G4).



Real-Time-alignment of Hela digest peptides

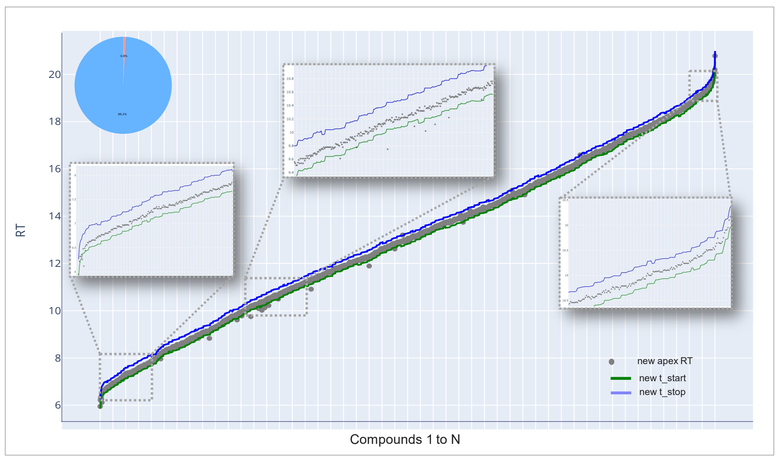
Table XY shows the results of ~2500 unique peptide targets for different gradients (G1-G4) and different starting RT windows (6sec, 12sec, 18sec & 30sec). In all cases we see a very good coverage of over 95% of the shifted peptide signals within the new estimated RT windows after alignment. As expected, the windows were broadened at maximum by 6-9 seconds in every case, caused by the addition of uncertainty bounds.

Figure 5 shows a visualization for the gradient G4 with an initial RT window of 18sec. The vast majority of Peptides signals lie within the RT windows indicating a very good chance to be triggered even under these conditions with shifts from 6 seconds to 55 seconds.

Table2. Results of retention time alignment for different gradients and different starting windows

Gradient	#peptides	Coverage% (30s)	$Rtwin_{max}$ (30s)	Coverage% (18s)	$Rtwin_{max}$ (18s)	Coverage% (12s)	$Rtwin_{max}$ (12s)	Coverage% (6s)	$Rtwin_{max}$ (6s)
G1 (ref)	2726	100	39	100	24	100	18	98.9	12
G2(offs +1)	2591	99.6	39	99.6	27	99.6	21	99.6	15
G3 (shal1)	2722	99.4	39	98.9	27	98.8	21	98.6	15
G4(steep-1)	2449	99.6	39	99.3	27	99.1	21	97	15

Figure 5. Overlay of 2449 peptide signals from a HeLa digest with new RT windows after real-time chromatographic alignment. The insertions show how well the apexes fit into the recalculated windows. Overall, 99.3% of all signals lie within the new boundaries.



Conclusions and outlook

We could demonstrate that real-time retention time correction based on MS1 spectra works very well even when LC conditions vary greatly. Over 97% of the tested peak apexes are kept within the boundaries recalculated on-the-fly during runtime , even with scheduled retention time windows as short as 15 seconds. The algorithmic broadening of the initial retention time windows can be viewed from different perspectives. In fully targeted experiments, it helps to cover the entire elution profile. However, it may be too conservative for data-directed approaches such as TMT experiments. In future experiments, the new algorithm will be applied to experiments at a larger scale to overcome chromatographic limitations caused by column aging without the need to spike isotope-labeled standards.

References

1. Philip M. Remes, Ping Yip, and Michael J. MacCoss *Highly Multiplex Targeted Proteomics Enabled by Real-Time Chromatographic Alignment, Analytical Chemistry* **2020** *92* (17), 11809-11817, DOI: 10.1021/acs.analchem.0c02075

Trademarks/licensing