

UC Riverside

UC Riverside Previously Published Works

Title

Normalized retention time for scheduled liquid chromatography-multistage mass spectrometry analysis of epitranscriptomic modifications.

Permalink

<https://escholarship.org/uc/item/58t0q2n9>

Authors

Gonzalez, Gwendolyn
Cui, Yuxiang
Wang, Pengcheng
et al.

Publication Date

2020-07-19

DOI

10.1016/j.chroma.2020.461181

Peer reviewed



Published in final edited form as:

J Chromatogr A. 2020 July 19; 1623: 461181. doi:10.1016/j.chroma.2020.461181.

Normalized Retention Time for Scheduled Liquid Chromatography-Multistage Mass Spectrometry Analysis of Epitranscriptomic Modifications

Gwendolyn Gonzalez¹, Yuxiang Cui¹, Pengcheng Wang¹, Yinsheng Wang^{1,2,*}

¹Environmental Toxicology Graduate Program, University of California, Riverside, California 92521, United States

²Department of Chemistry, University of California, Riverside, California 92521, United States

Abstract

Investigations into post-transcriptional modifications of RNA and their regulatory proteins have revealed pivotal roles of these modifications in cellular functions. A robust method for the quantitative analysis of modified nucleosides in RNA may facilitate the assessment about their functions in RNA biology and disease etiology. Here, we developed a sensitive nano-liquid chromatography-multistage mass spectrometry (nLC-MS³) method for profiling simultaneously 27 modified ribonucleosides. We employed normalized retention time (iRT) and scheduled selected-reaction monitoring (SRM) to achieve high-throughput analysis, where we assigned iRT values for modified ribonucleosides based on their relative elution times with respect to the four canonical ribonucleosides. The iRT scores allowed for reliable predictions of retention times for modified ribonucleosides with the use of two types of stationary phase materials and various mobile phase gradients. The method enabled the identification of 20 modified ribonucleosides with the use of the enzymatic digestion mixture of 2.5 ng total RNA and facilitated robust quantification of modified cytidine derivatives in total RNA. Together, we established a scheduled SRM-based method for high-throughput analysis of modified ribonucleosides with the use of a few nanograms of RNA.

Introduction

There are over 100 types of modified nucleosides in RNA, primarily in transfer RNA (tRNA) and ribosomal RNA (rRNA).¹⁻³ These modifications in tRNA and rRNA can influence their stability and maturation, respectively, thereby impacting translation efficiency and cellular response to stress.^{1,4} Recently, several types of modified nucleosides have also been identified in other types of RNA, including messenger RNA (mRNA) and microRNA (miRNA).⁵⁻⁹ The biological functions of these RNA modifications, their regulatory enzymes, and their contributions to disease pathology, however, remain incompletely understood and are under intense investigation.⁸

*To whom correspondence should be addressed: Tel.: (951) 827-2700, Fax: (951) 827-4713; Yinsheng.Wang@ucr.edu.

Conventional methods for monitoring RNA modifications are often suitable for a single or a few types of modifications with relatively high modification stoichiometry, and are semi-quantitative.^{1,9-11} For instance, bisulfite sequencing is commonly employed for mapping 5-methylcytosine (m⁵C) in DNA and RNA.¹⁰ False positives can arise from imperfect deamination of cytosine to uracil during bisulfite treatment, and the method cannot distinguish m⁵C from its oxidation products of 5-hydroxymethylcytosine (5-hmrC), 5-formylcytosine (5-forC) and 5-carboxylcytosine (5-carC).^{10,12}

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) constitutes a powerful tool for accurately identifying and quantifying modified nucleosides, and most measurements are based on the characteristic neutral loss of a ribose or 2-deoxyribose from protonated nucleosides during collisional activation.^{1,8,9,13,14} Nevertheless, previously published LC-MS/MS methods for ribonucleoside analysis require relatively large amount of RNA (100-1000 ng) per injection.^{1,7,14} For RNAs of relatively low abundance, including mRNA and miRNA, obtaining a substantial amount of RNA is laborious and costly.

Scheduled selected-reaction monitoring (SRM) method relying on the use of normalized retention time (iRT) for the prediction of retention times of analytes under given chromatographic conditions has been employed for high-throughput analyses of peptides and modified 2'-deoxyribonucleosides.^{13,15} A scheduled SRM method is advantageous in limiting acquisition time for specific precursor ions to a few minutes, thereby allowing for shorter cycle time and providing more data points per chromatographic peak to enable robust and sensitive analyte detection.¹³

The objective of this study was to develop a nanoflow liquid chromatography-multistage MS (nLC-MS³) method to simultaneously monitor modified ribonucleosides with low nanogram quantities of RNA. We show that the scheduled SRM method can facilitate simultaneous assessment of a large number of ribonucleosides (27 modified nucleosides and 4 unmodified canonical nucleosides, structures shown in Figure S1) in a single LC-MS³ run without compromising detection sensitivity or accuracy.

Experimental Section

RNA sample preparation

Total RNA was isolated from HEK293T cells and digested with enzymes following previously published procedures.⁹ Briefly, total RNA was isolated using E.Z.N.A. Total RNA Kit I (Omega) according to the manufacturer's recommended procedures. The RNA samples were then enzymatically digested to mononucleosides with nuclease P1 and phosphodiesterase 2 at 37°C for 4 hrs. Alkaline phosphatase and phosphodiesterase 1 were subsequently added to the resulting solution, and the mixture was incubated at 37°C for 2 hrs. The enzymes were subsequently removed from the digestion mixture by chloroform extraction.

A mixture of ribonucleoside standards and 20 ng of enzymatically digested RNA was used for defining iRT values and for establishing iRT-RT correlation. A list of ribonucleoside standards employed in this study are shown in Table S1.

Nanoflow liquid chromatography-multistage mass spectrometry (nLC-MS³) analysis

Unscheduled and scheduled SRM experiments were performed on an LTQ-XL linear ion trap mass spectrometer coupled with an EASY-nLC II (Thermo Fisher Scientific, San Jose, CA). The mass spectrometer was operated in the positive-ion mode with the electrospray, capillary, and tube lens voltages being 2.0 kV, 12 V, and 100 V, respectively. Ion transport tube temperature was maintained at 275°C. Precursor ions for MS² and MS³ analyses of ribonucleosides are listed in Table S2.

To assign iRT values, a nucleoside mixture from the enzymatic digestion of total RNA and stable isotope-labeled ribonucleoside standards was loaded, at a flow rate of 2.5 µL/min, onto a 5-cm or 3.5-cm long in-house packed porous graphitic carbon (PGC, 5 µm particle size, Thermo Fisher Scientific) trapping column (150 µm i.d.) for linear and non-linear gradients, respectively. Analytes eluting from the trapping column were directed to and resolved on an 18-cm Zorbax SB-C18 (5 µm in particle size, 100 Å in pore size, Agilent) or Magic C18-AQ (5 µm in particle size, 100 Å in pore size, Michrom BioResources) analytical column (75 µm i.d.) at a flow rate of 300 nL/min. Mobile phases A and B were 0.1% (v/v) formic acid in water and 0.1% (v/v) formic acid in acetonitrile, respectively, and the mobile phase gradients are listed in Table S3. Briefly, linear gradients had a linear increase of 16%-70% mobile phase B over 65 min for the fast gradient (hereafter referred to as linear, long-fast gradient). To evaluate iRT-RT correlation with complex gradients, we also employed a gradient with a step-wise increase in %B, which is referred to as non-linear gradient. For instance, non-linear, long-fast gradient consisted of 0–16% B in 5 min, 16–22% B in 23 min, 22–50% B in 17 min, 50–90% B in 5 min and finally at 90% B for 30 min.

To assign iRT values for modified ribonucleosides, the iRT values for cytidine (rC) and adenosine (rA) were arbitrarily assigned to 10 and 100, respectively, and the iRT values for the nucleoside of interest (X) was determined using the following equation: $iRT_X = 100 - [(RT_2 - RT_X)/(RT_2 - RT_1)] \times 90$, where iRT_X represents the iRT value for nucleoside X, and RT_X , RT_1 , and RT_2 designate the observed retention times for X, rC and rA, respectively. The predicted retention time for nucleoside X was calculated from its iRT (iRT_X) and the actual retention times observed for rC (RT_1) and rA (RT_2) in the calibration run with the following equation: $RT_X = RT_2 - (100 - iRT_X) \times (RT_2 - RT_1)/90$.

Quantifications of 5-methylcytidine and 2'-O-methylcytidine

For all measurements, 2.5 ng of digested RNA spiked with isotopically labeled standards was loaded onto a 4.1-cm PGC trapping column and eluted to an 18-cm Zorbax SB-C18 analytical column. The same sample was injected and analyzed using the scheduled LC-MS/MS/MS method, with precursor ions for the stable isotope-labeled standards being incorporated into the transition list.

Results

To assign iRT values and to assess iRT-RT correlation, we injected a solution containing nucleoside mixture from the digestion of 5 ng of total RNA isolated from HEK293T cells and synthetic standard ribonucleosides. For accurate prediction of retention times, we

employed the four canonical ribonucleosides (i.e. cytidine, uridine, guanosine and adenosine) as standards, where we assigned the iRT values of cytidine and adenosine as 10 and 100, respectively. The long-fast gradient was employed for calculating iRT scores for uridine, guanosine and modified ribonucleosides using retention time and linear regression analysis (Figures 1, S1, and S2, Table S3). We also determined the iRT scores with the use of linear and non-linear gradients to evaluate the reproducibility of iRT values under these gradient conditions (Table S4).

Traditionally, linear gradients were employed for establishing RT-iRT correlation.^{13,15} To assess the reliability of iRT-RT correlation under different chromatography settings, we considered different gradients speeds (long, long fast, and short fast), and examined both linear and non-linear gradients. Our results showed that all different gradients yielded linear iRT-RT relationships with good correlation coefficients: linear gradients (long, $R^2 = 0.992$; long-fast, $R^2 = 0.997$; short-fast, $R^2 = 0.987$) and non-linear gradients (long, $R^2 = 0.969$; long-fast, $R^2 = 0.995$; short, $R^2 = 0.975$; short-fast, $R^2 = 0.972$) (Figure 2). Therefore, normalized retention time values can also be employed for more complex gradients. Typically, non-linear gradients allow for more efficient separation of closely eluting species.

We also observed that short gradients exhibited poor separation efficiencies, and did not allow for the elution of a few modified hydrophobic nucleosides. Long gradients, especially those with higher content of mobile phase B, including long and long-fast gradients, facilitate the elution of all or nearly all target nucleosides. Consistent with the previous study,¹³ the number of analytes eluted in a single run depends on the length of the gradient and the maximum percentage of mobile phase B.

Next, we examined if a similar iRT-RT correlation could be attained using an analytical column packed with a different stationary phase material, Magic C18-AQ. We found that, with the use of different linear gradients, the iRT values exhibit excellent linear correlations with the observed RTs on this analytical column (Figures 3 and S3). Moreover, iRT values did not differ substantially with the use of Magic C18-AQ or Zorbax SB-C18 analytical column (Figure 3, Table S3), suggesting that iRT values and RT predictions are transferable to other reversed-phase C18 stationary phase materials.

On average, the RTs predicted from iRTs were within 2 and 2.5 min of the actual RTs with the use of the linear and non-linear gradients, respectively (Figure 4A, Table S4). The iRT values were assigned based on the long-fast gradient; thus, the elution times of analytes with the use of long-fast gradient were more accurately predicted and overall displayed smaller variations from the predicted RTs than those with other modified gradients (Figure 4A). In addition, the analytes exhibiting the largest uncertainty in RT prediction eluted toward the end of the gradients, e.g. rG and its mono-methylated derivatives (Figure 1 and Figure 4A). Expanding the acquisition time window for later-eluting analytes will compensate for drift in RT without compromising measurement efficiency because a relatively small number of modified nucleosides elute from the column toward the end of the gradient.

The major motivation for developing this method was to achieve simultaneous quantifications of modified ribonucleosides with a scheduled SRM method. By using an

LTQ-XL linear ion trap mass spectrometer, we divided data acquisition time into 10-min retention time windows and monitored the transitions for ribonucleosides in individual windows based on their predicted retention times. In doing so, we were able to monitor a small number of analytes per acquisition time segment rather than monitoring all analyte transitions throughout the entire gradient. However, some scan events are redundant between segments because several analytes have similar transitions. For example, *N*¹-methyladenosine (*m*¹A) and *N*⁶-methyladenosine (*m*⁶A) share the same MS/MS transition, but differ in MS/MS/MS (data not shown) and retention time (Figure 1). Thus, common precursor ions may be monitored in more than one retention time segment for isomeric ribonucleosides; nevertheless, they are easily distinguishable from one another based on retention time and MS³.⁹

Next, we utilized the method to identify all possible modified ribonucleosides in the nucleoside mixture of 2.5 ng of total RNA isolated from HEK293T cells. Conventional LC-MS/MS-based global screening methods require considerable amount of RNA digestion mixture per injection.^{1,2,14} Another drawback of these methods resides in the throughput, where a single or few analytes of interest are monitored per run to enable short cycle time and robust quantification.^{6,9} Our scheduled LC-MS³ method incorporated more analytes per run while ensuring cycle time in each segment is short enough for robust quantification. Furthermore, the use of isotopically labeled standards provides unambiguous identification and reliable quantification, as isotopically labeled standards exhibit identical chromatographic behaviors and fragmentation pathways to the unlabeled analytes.⁹ Our results revealed the presence of 20 modified ribonucleosides and the 4 canonical ribonucleosides in the nucleoside mixture of 2.5 ng of total RNA (Figure S4).

We also compared the quantification results for 5-methylcytidine (*m*⁵C) and 2'-*O*-methylcytidine (*C*_m) obtained from the scheduled LC-MS/MS/MS method with a previously published stable-isotope dilution method based on unscheduled SRM analysis of two modified ribonucleosides.⁹ The scheduled method allowed for monitoring significantly more types of modifications, including isotopically labeled standards. In this vein, the limits of detection of the scheduled SRM method for *m*⁵C and *C*_m, which is defined as the amounts of analyte needed to yield a signal-to-noise ratio of 3.0 in the selected-ion chromatogram, were (88 ± 14) amol and (152 ± 25) amol, respectively, with the linear gradient (n = 3), and (91 ± 3) and (145 ± 27) amol, respectively, with the non-linear gradient (n = 3). Meanwhile, the quantification results for *m*⁵C and *C*_m are as reliable as those obtained from the previously published low-throughput method, as manifested by the lack of statistically significant differences in the quantification results between the two methods (Figure 4B and Table S5). Therefore, the scheduled SRM method offers high-throughput analysis without diminishing quantification efficiency.

Discussion

LC-MS/MS is a widely used tool for global RNA modification screening. Recently published LC-MS/MS-based methods have overcome some long-standing analytical challenges, including assessing the levels of modified nucleosides of low abundance. Nano-flow LC with a PGC column afforded improved sensitivity and global screening capabilities

for the analyses of modified ribonucleosides.^{14,16} Another need in the analysis of modified ribonucleosides resides in the differentiation of regioisomers. MS/MS-based methods, through monitoring the neutral loss of a ribose, do not allow for the differentiation of regioisomeric mono-methylated nucleobase modifications, e.g. m¹A and m⁶A. Nevertheless, MS/MS/MS on a linear ion-trap mass spectrometer can provide diagnostic fragment ions for distinguishing some of these regioisomers.⁹ In addition, retention time provides another dimension of information to corroborate identification of modified ribonucleosides based on MS/MS and/or MS/MS/MS.

Future studies would benefit from expanding the number of modified ribonucleosides to be monitored. The method presented here targeted modified ribonucleosides with elution time between rC and m⁶A. Other more hydrophobic modified ribonucleosides are known to exist, including N²,N²-dimethylguanosine and N⁶-isopentenyladenosine.^{1,17} The analyses of these nucleosides may entail the use of alternative stationary phase materials for efficient elution, as well as using other more hydrophobic nucleosides as standards for iRT determination.

In conclusion, we developed a scheduled SRM method for targeted analyses of modified ribonucleosides. We established iRT values for 27 modified and 4 canonical ribonucleosides. We also observed consistent iRT scores for ribonucleosides with the use of two types of stationary phase materials, Zorbax SB-C18 and Magic C18-AQ, and with different mobile phase gradients. We reason that the iRT values should be easily transferable among different chromatographic settings and in different laboratories for reliable retention time prediction. We also showed that the precision and accuracy of the scheduled LC-MS/MS/MS method were comparable to those obtained previously with the low-throughput, unscheduled SRM method. Therefore, the scheduled SRM-based LC-MS/MS/MS method allows for robust and high-throughput analysis of modified ribonucleosides, which holds great potential for investigations into the dynamic regulation of the epitranscriptome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement.

This work was supported by the National Institutes of Health (R35 ES031707), and G.G. was supported in part by an NRSA T32 Institutional Training Grant (T32 ES018827).

References

- (1). Su D; Chan C; Gu C; Lim K; Chionh Y; McBee M; Russell B; Babu I; Begley T; Dedon P *Nat. Prot* 2014, 9, 828–841.
- (2). Chen Q; Shi J; Peng H; Zhang X; Zhang Y; Qian J; Duan E; Zhai Q; Zhou Q *Sci.* 2016, 351.
- (3). Ross R; Cao X; Yu N; Limbach PA *Methods* 2016, 107, 73–78. [PubMed: 27033178]
- (4). Calo E; Flynn RA; Martin L; Spitale RC; Chang HY; Wysocka J *Nature* 2015, 518, 249–253. [PubMed: 25470060]
- (5). Gaston KW; Limbach PA *RNA Bio.* 2014, 11, 1568–1585. [PubMed: 25616408]

- (6). Yamauchi Y; Nobe Y; Izumikawa K; Higo D; Yamagishi Y; Takahashi N; Nakayama H; Isobe T; Taoka M *Nucleic Acids Res.* 2015, 44.
- (7). Basanta-Sanchez M; Temple S; Ansari SA; D'Amico A; Agris PF *Nucleic Acids Res.* 2015, gkv971–gkv971.
- (8). Fu L; Guerrero CR; Zhong N; Amato NJ; Liu Y; Liu S; Cai Q; Ji D; Jin S-G; Niedernhofer LJ; Pfeifer GP; Xu G-L; Wang YJ *Am. Chem. Soc* 2014, 136, 11582–11585.
- (9). Fu L; Amato NJ; Wang P; McGowan SJ; Niedernhofer LJ; Wang Y *Anal. Chem* 2015, 87ha, 7653–7659.
- (10). Helm M; Motorin Y *Nat. Rev. Genet* 2017, 18, 275–291. [PubMed: 28216634]
- (11). Li X; Xiong X; Yi C *Nat. Meth* 2016, 14, 23–31.
- (12). Booth M; Ost T; Beraldi D; Bell N; Branco M; Reik W; Balasubramanian S *Nat. Prot* 2013, 8, 1841–1851.
- (13). Cui Y; Wang P; Yu Y; Yuan J; Wang Y *Anal. Chem* 2018, 90, 14111–14115. [PubMed: 30500177]
- (14). Giessing AMB; Scott LG; Kirpekar FJ *Am. Soc. Mass Spectrom* 2011, 22, 1242–1251.
- (15). Escher C; Reiter L; MacLean B; Ossola R; Herzog F; Chilton J; MacCoss M; Rinner O *Proteomics* 2012, 12, 1111–1121. [PubMed: 22577012]
- (16). Sarin L; Kienast S; Leufken J; Ross R; Dziergowska A; Debiec K; Sochacka E; Limbach P; Fufezan C; Drexler H; Leidel S *RNA* 2018, 24, 1403–1417. [PubMed: 30012570]
- (17). Iwanami Y; Brown G *Arch. Biochem. Biophys* 1968, 126, 8–15. [PubMed: 5671075]

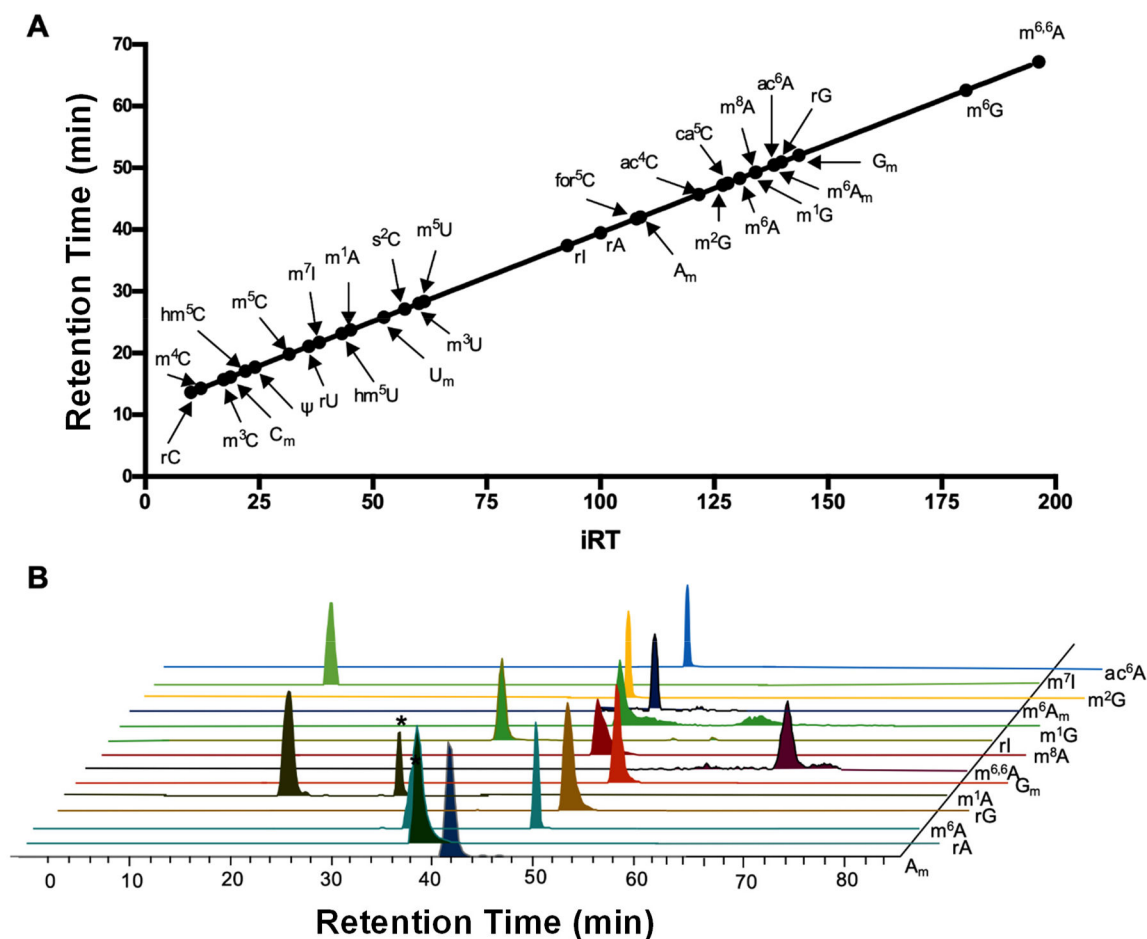


Figure 1.

Scheduled SRM-based LC-MS/MS/MS method for the analysis of modified ribonucleosides. A) The iRT-RT correlation for ribonucleosides on a Zorbax SB-C18 column with the use of long-fast gradient. Ribonucleosides are labeled on the line. Complete names and chemical structures are shown in Figure S1 and Table S1. B) Representative selected-ion chromatograms (SICs) for rA, rG, and their mono-methylated derivatives using a scheduled SRM-based LC-MS³ method with a non-linear, long-fast gradient. SICs for other modified ribonucleosides monitored are shown in Figure S2. Interference peaks are denoted with asterisks.

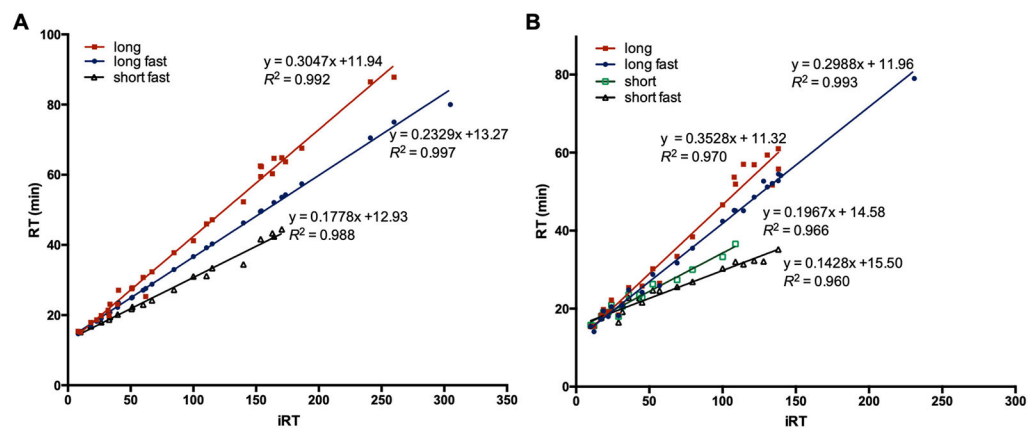


Figure 2. RT-iRT correlations acquired from the use of a porous graphitic carbon (PGC) trapping column and a Zorbax SB-C18 analytical column with different mobile phase gradients. A) linear gradients with modified gradient speeds (long, long-fast, and short-fast); and B) non-linear gradients with different gradient speeds (long, long-fast, short-fast, short).

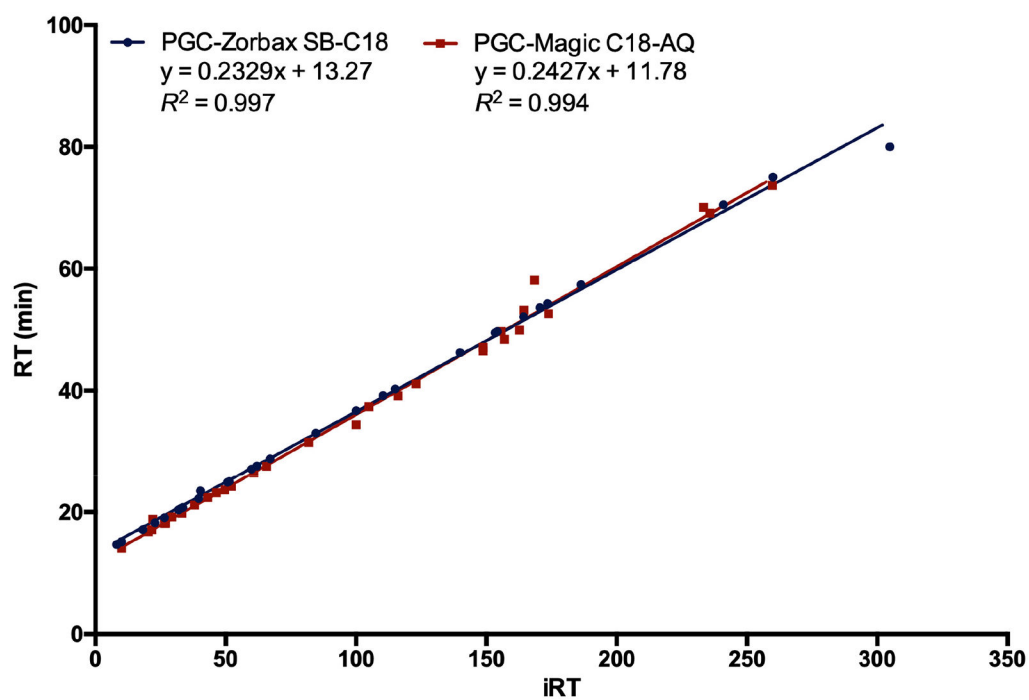


Figure 3.
A comparison of RT-iRT correlations obtained from the use of Zorbax SB-C18 and Magic C18-AQ analytical columns.

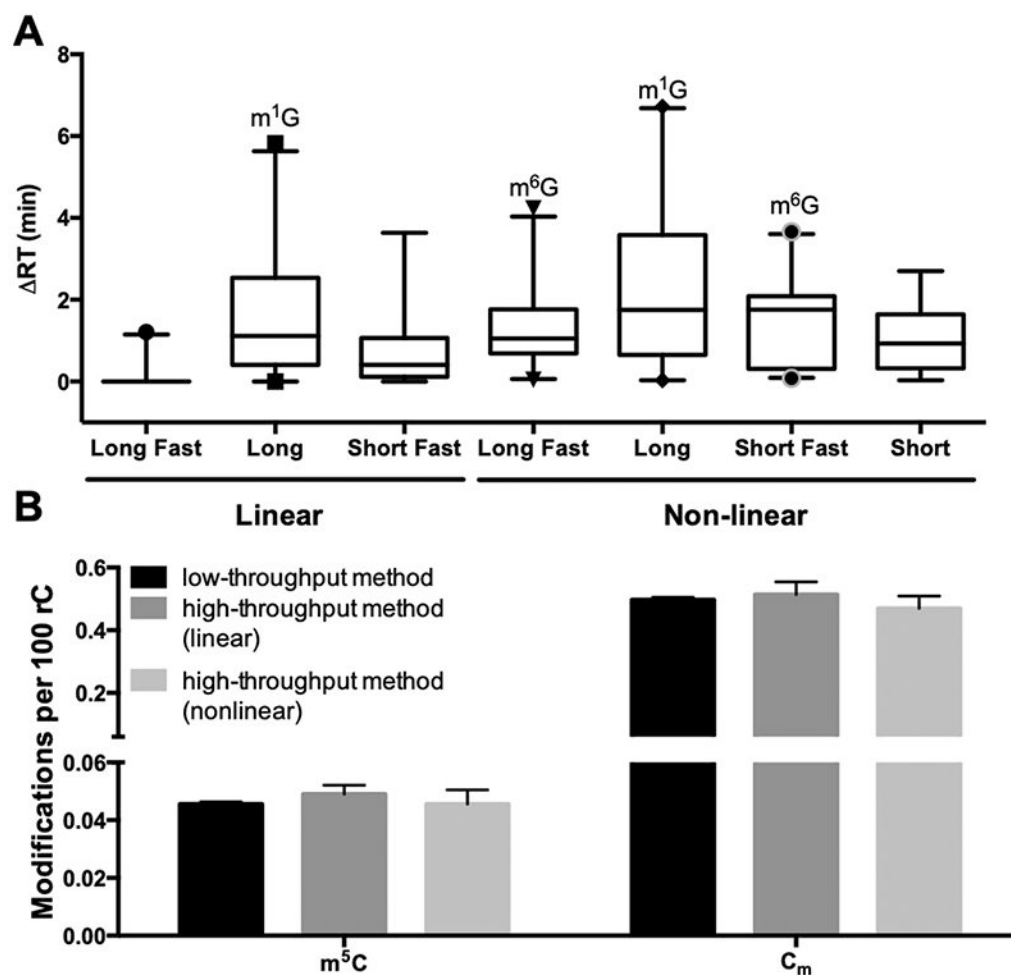


Figure 4. Precision and accuracy of the high-throughput scheduled SRM method determined by: A) absolute difference (in minutes) between the predicted and observed RT using linear and non-linear, modified gradients; and B) a comparison of quantification results from scheduled SRM LC-MS³ using long-fast, linear and non-linear gradients with those obtained with low-throughput method. The whiskers plotted in A) correspond to the 5th and 95th percentiles and representative outliers are shown based on data acquired from an average of four LC-MS³ runs. On average, divergences of the predicted RTs from the observed RTs were within 2 and 2.5 min for linear and non-linear gradients, respectively. Detailed average difference from predicted RT for all gradients are found in Table S4. Data in B) represent the mean and standard deviation ($n = 3$) between low-throughput method monitoring 6 precursor ions throughout the gradient and high-throughput scheduled SRM method monitoring 33 unlabeled and 3 labeled precursor ions. p -values were calculated using two-tailed Student's t -test are found in Table S5.