

## **Permethylation of ribonucleosides provides enhanced mass spectrometry quantification of post-transcriptional modifications**

Yixuan Xie<sup>1,2</sup>, Kevin A. Janssen<sup>2</sup>, Alessandro Scacchetti<sup>3</sup>, Roberto Bonasio<sup>3</sup>, Benjamin A Garcia<sup>1,2</sup>

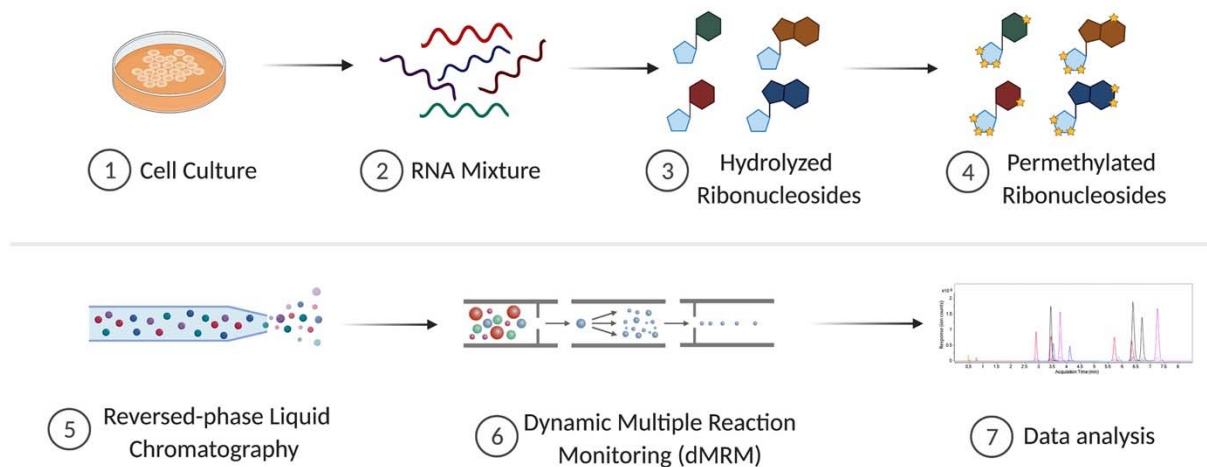
<sup>1</sup>Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, United States.

<sup>2</sup>Epigenetics Institute, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States.

<sup>3</sup>Epigenetics Institute, Department of Cell and Developmental Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States.

Correspondence and requests for materials should be addressed to B.A.G. (email: [bagarcia@wustl.edu](mailto:bagarcia@wustl.edu))

## Abstract



Chemical modifications of RNA are associated with fundamental biological processes such as RNA splicing, export, translation, degradation, as well as human disease states such as cancer. However, the analysis of ribonucleoside modifications is impeded due to the hydrophilicity of the ribonucleoside molecules. In this research, we used solid-phase permethylation to derivatize the ribonucleosides, and the permethylated ribonucleosides, which were then quantitatively analyzed using a liquid chromatography–tandem mass spectrometry (LC–MS/MS)-based method. The solid-phase permethylation efficiently derivatized the ribonucleosides, and more than 60 RNA modifications were simultaneously monitored using ultrahigh-performance liquid chromatography coupled with triple quadrupole mass spectrometry (UHPLC–QqQ–MS) performed in the dynamic multiple reaction monitoring (dMRM) mode. Because of the increased hydrophobicity of permethylated ribonucleosides, this method enhanced retention, separation, and ionization efficiency, resulting in improved detection and quantification when compared to existing analytical strategies of RNA modifications. We applied this new approach to measure the extent of cytosine methylation and hydroxymethylation in RNA obtained from mouse embryonic stem cells with genetic deficiencies in ten-eleven translocation (TET) enzymes. The results matched previously performed analyses and highlighted the sensitivity, efficacy, and robustness of the new method. The advantage of this method enables comprehensive analysis of RNA modifications in biological samples.

## Introduction

Ribonucleic acids (RNAs) are at the heart of the central dogma of biology. Messenger RNAs (mRNAs) are transcribed from the genome and direct protein synthesis in the cytosol. This process is highly regulated and aided by a number of noncoding RNAs, including transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs). In recent years, it has become evident that chemical modifications deposited on RNAs during or after transcription contribute to the regulation of gene expression, giving rise to a field sometimes referred to as “epitranscriptomics”.<sup>1-3</sup> Collectively, more than 150 distinct RNA modifications originating from the canonical adenosine (A), guanosine (G), cytidine (C), and uridine (U) have been identified across different organisms.<sup>4</sup> These RNA modifications are involved in many fundamental biological processes, including cell differentiation, alternative splicing, and stress responses.<sup>5</sup> At the molecular level, the modifications have been demonstrated to be crucial for proper RNA folding, topology, stability, high-order structure, and protein translation.<sup>6, 7</sup> Translation is highly dependent on RNA modifications, which are critical for proper rRNA and tRNA function, and can regulate the rate and fidelity of translation when decorating mRNA.<sup>8</sup> Additionally, RNA modification can affect the ability of proteins to recognize certain RNA.<sup>9, 10</sup> Dysregulation of RNA modifications is associated with several diseases. For example, one of the most prevalent RNA methylations, N6-methyladenosine (m6A), has been linked to the carcinogenesis and progression of colorectal cancer.<sup>11</sup>

Although the importance of RNA modifications has been well-recognized, their characterization and quantification has been a long-standing problem hampered by the inadequacy of existing analytical methods. Sequencing-based techniques are powerful; however, most of the existing methods are unable to identify the modifications directly.<sup>12</sup> While nanopore sequencing has been developed to directly analyze m6A, while this approach lacks sensitivity and it is challenging to detect other types of modifications.<sup>13</sup> Indirect identification has been robustly performed using modification-specific antibodies and chemical reagents that react with specific modifications, but these methods cannot be used to make comprehensive and quantitative measurements because each modification requires a distinct workflow.<sup>14</sup> Moreover, sequencing cannot reliably provide simultaneous quantitation and localization of multiple modifications. In addition to these genomic approaches, nuclear magnetic resonance (NMR) has been applied to RNA modification analysis.<sup>15</sup> However, the analysis of modifications using NMR is challenging due to chemical shift overlaps, and line broadening often leads to high background noise and signal loss.<sup>16</sup>

Mass spectrometry (MS) has emerged as an alternative for developing powerful methods to analyze the modifications of RNA.<sup>17</sup> Different modifications can be characterized in a single run, which enables high-throughput analyses. Although methods have been developed to analyze RNA at the oligonucleotide level to localize modifications, MS analysis of oligonucleotides is still far from established due to three limitations: digestion, instrumentation, and software.<sup>18-20</sup> Therefore, hydrolyzing RNAs into ribonucleosides and characterizing RNA modifications at the ribonucleoside level using MS is a more accessible approach.<sup>21, 22</sup> However, the commonly used reversed-phase chromatography is not optimal for native ribonucleosides due to their high polarity of most of these compounds. Although other stationary phases such as porous graphitic carbon (PGC) and hydrophilic interaction liquid chromatography (HILIC) have been utilized, these methods are not ideal for profiling all known ribonucleoside

modifications,<sup>23</sup> because these materials are not ideal for separating all modified ribonucleoside; for example, N6-threonylcarbamoyladenosine (t6A) cannot be reliably analyzed using PGC as a stationary phase.<sup>24</sup> Additionally, some of these chromatography approaches for RNA analysis often use ion-pairing agents, which are not compatible with mass spectrometry and are difficult to remove from the liquid chromatography (LC).<sup>25</sup>

Rather than changing the material in the stationary phase, chemical derivatization methods have been applied to alter the chemical and physical properties of ribonucleosides and make them compatible with a more desirable stationary phase. Patteson *et al.* developed a method using 1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide to label pseudouridine ( $\Psi$ ) residues.<sup>26</sup> Huang *et al.* also demonstrated a method to derivatize 5-methylcytosine (m5C) and its oxidation products using bromoacetyl-containing reagents.<sup>27</sup> These approaches were applied to exclusively monitor specific targets of interest depending on the selective chemical reactivities towards a particular ribonucleoside modification. Cai and co-workers converted ribonucleosides into acetonides using acetone for a more comprehensive study, and more than 50 derivatized ribonucleosides were identified.<sup>28</sup> Nevertheless, the chemistry requires *cis-diol* groups, and nucleosides containing 2'-*O*-methylation could not be identified by this approach.

Here, we report a novel method to detect RNA modifications by MS that takes advantage of permethylation as a derivatization strategy. Permethylation replaces the hydrogens on hydroxyl groups, amine groups, and carboxyl groups with methyl groups. It is an efficient strategy for derivatizing a wide array of polar molecules and has been widely employed for glycan analysis.<sup>29</sup> Permethylation of ribonucleosides were first attempted when the reaction was expected to yield volatile derivatives for gas chromatography–MS analysis.<sup>30,31</sup> As these studies using permethylation of canonical ribonucleosides were carried out before modern MS-based analytical methods were developed, the properties of the products were not well characterized using modern LC-MS. In addition, the unmodified ribonucleosides and the differently methylated ribonucleosides, such as A, m6A, and Am, cannot be unambiguously identified because the previous method only obtained MS1 spectra and lacked chromatographic separation. We therefore established a solid-phase permethylation method using isotopically labeled iodomethane to derivatize mono-ribonucleosides. The generated hydrophobic ribonucleoside derivatives were subsequently separated by reversed-phase C<sub>18</sub>-based liquid chromatography (RP-LC). Precursor and product ions were detected and quantified using triple quadrupole MS (QqQ MS) operated in the targeted dynamic multiple reaction monitoring (dMRM) mode, simultaneously. This method yielded a few advantages: (i) the solid-phase permethylation allowed efficient labeling in a short time period; (ii) the endogenous methylated ribonucleosides were distinguishable by different precursor and/or product ions due to derivatization with isotopically labeled iodomethane; (iii) the permethylation product of endogenous methylated and unmodified ribonucleosides had the same retention time, allowing improved the quantitative analysis; (iv) this method allowed spontaneous distinguishment for some ribonucleoside isomers (e.g., m3U and m5U) due to their unique precursor and product ions after permethylation; (v) the method yielded higher sensitivity (sub-femtomole level) compared to analyzing the underivatized ribonucleosides; and (vi) the chemical reaction was highly predictable, so that this method could be easily extended to currently unknown RNA modifications as they continue to be discovered. Consequently, more than 60 transitions were built as an example using the combination of ribonucleoside standards and digested

ribonucleosides from cell extracts. To demonstrate the method's viability, we examined 5-methylcytosine (m5C) levels in wild-type (WT) mouse embryonic stem cells (mESCs) and compared them to those found in cells deficient in enzymes of the ten-eleven translocation (TET) family. As previously reported, we observed the abundance of 5-hydroxymethylcytosine (hm5C) decreased in cells lacking one or more TET enzymes.<sup>32</sup> In addition, our improved quantification method revealed an expected but previously unreported increase in m5C levels, whereas the level of 2'-O-methylcytosine (Cm) were similar in WT and mutant cells. This work demonstrates the advantage of permethylation as a derivatization strategy for highly accurate quantification of RNA modifications *via* mass spectrometry.

## Experimental procedures

### *Samples and Materials.*

The ribonucleoside standards were purchased from Carbosynth (San Diego, CA). Iodomethane-*d*<sub>3</sub>, dichloromethane (DCM), zinc chloride (ZnCl<sub>2</sub>), sodium acetate (NaOAc), sodium hydroxide beads (NaOH), nucleosides test mix, nuclease P1, phosphodiesterase I, and phosphodiesterase II were purchased from Sigma-Aldrich (St. Louis, MO). Anhydrous dimethyl sulfoxide (DMSO) was purchased from Biotium (Fremont, CA). Acetonitrile (ACN), formic acid, recombinant shrimp alkaline phosphatase, porous graphitic carbon (PGC) stage tip, and micro spin column were purchased from Thermo Scientific (Waltham, MA).

### *Culturing Mouse Embryonic Stem Cells (mESCs).*

*Tet2* KO and *Tet1/2/3* triple KO mESC cell lines were previously described.<sup>32</sup> All the cells were cultured on gelatin-coated dishes in KnockOut DMEM supplemented with 15% FBS, 0.1 mM MEM nonessential amino acids, 0.1 mM 2-mercaptoethanol, 1 mM L-glutamine, 0.5% penicillin streptomycin, 100 U/mL leukemia inhibitory factor (LIF), 3 μM CHIR99021, and 1 μM PD0325901, and the cells were maintained in a humidified cell culture incubator with 5% CO<sub>2</sub> at 37 °C.

### *Preparation of ribonucleosides from cell culture.*

Total RNA was extracted with Direct-zol RNA Kits (Thermo Scientific, Waltham, MA) based on the manufacture protocols. 100 ng of RNA samples were digested into ribonucleosides with 5 mU/μL of nuclease P1, 5 mU/μL of recombinant shrimp alkaline phosphatase, 500 μU/μL of phosphodiesterase I, and 6.25 μU/μL of phosphodiesterase II in 20 μL of digestion buffer (1 mM ZnCl<sub>2</sub>, 30 mM NaOAc, pH 7.5) overnight at room temperature. The digested ribonucleosides were purified using PGC stage tips and dried in a Savant SpeedVac concentrator (Thermo Scientific, Waltham, MA).

### *Solid-phase permethylation of ribonucleosides.*

The digested ribonucleoside samples were permethylated using solid-phase permethylation, as previously described, with some optimization.<sup>33</sup> Briefly, NaOH beads were packed into the empty spin columns (about 2 cm height), and the beads were washed with 100 μL of DMSO twice. The purified and dried ribonucleoside samples were reconstituted in a mixture of 1 μL of water, 50 μL of DMSO, and 30 μL of iodomethane-*d*<sub>3</sub>. The samples were loaded into the spin column and spun down at 200 x g, followed by reloading the samples into

the column four times. Next, 20  $\mu\text{L}$  of iodomethane- $d_3$  was added to the sample and incubated at room temperature for another 10 min. The column was washed with 50  $\mu\text{L}$  of DMSO twice, 500  $\mu\text{L}$  of ice-cold water was added into the sample, and the mixture was incubated at room temperature for at least 1 min to quench the permethylation. Afterward, 300  $\mu\text{L}$  of DCM was added, the liquid-liquid extraction was repeated at least five times for each sample, and the organic layer was dried using a Savant SpeedVac concentrator.

#### *Ultrahigh-Pressure Liquid Chromatography/Triple Quadrupole Mass Spectrometry (UHPLC/QqQ-MS) Analysis.*

Separation and characterization of the ribonucleosides were carried out on a Thermo Scientific Vanquish Flex binary UHPLC system coupled to a Thermo Scientific TSQ Altis QqQ mass spectrometer (Thermo Scientific, Waltham, MA). For the analysis, 2  $\mu\text{L}$  of the sample was injected onto a Thermo Scientific Accucore Vanquish C18 column (150  $\times$  2.1mm, 1.5 $\mu\text{m}$ ) and separated using a 20 min binary gradient with a constant flow rate of 0.2 mL/min at 60  $^\circ\text{C}$ . Mobile phase A was water with 0.1% formic acid, and mobile phase B was 80% ACN/water (v/v) with 0.1% formic acid. To analyze the underivatized ribonucleosides, the following binary gradient was used: 0–3 min, 0% B; 3–10 min, 0–2% B; 10–11.5 min, 2%–99% B; 11.5–15 min, 99% B; 15–15.5 min, 99–0% B; 15.5–20 min, 0% B. For analysis of permethylated ribonucleosides, the following binary gradient was used: 0–7 min, 20%–40% B; 7–10 min, 40–70% B; 10–11 min, 70%–99% B; 11–15 min, 99% B; 15–16 min, 99–20% B; 16–20 min, 20% B. Samples were introduced into the mass spectrometer using electrospray ionization (ESI) source operated in the positive ion mode at 3500 V. Nitrogen sheath gas, auxiliary gas, and sweep gas flow rates were set at 30, 5, and 2 psi, respectively. The ion transfer tube temperature and vaporizer temperature were set at 350  $^\circ\text{C}$  and 175  $^\circ\text{C}$ , respectively. The precursor ions were fragmented using collision-induced dissociation (CID) with optimized energy. Data acquired from the UHPLC/QqQ-MS was collected using Thermo Scientific Xcalibur software (v4.1), and data analysis was performed using Thermo Scientific FreeStyle software (v2.1).

#### *nanoLC-MS/MS analysis.*

The samples were characterized using an EASY-nLC™ 1200 system coupled with a Q Exactive mass spectrometer (ThermoFisher Scientific). 3  $\mu\text{L}$  of the sample was injected, and the analytes were separated on self-packed C18 column (3  $\mu\text{m}$ , 0.150 mm  $\times$  250 mm) at a flow rate of 700 nL/min. Water containing 0.1% formic acid and 80% acetonitrile containing 0.1% formic acid were used as solvents A and B, respectively. MS spectra were collected with the mass range 200–600  $m/z$  in positive ionization mode. The filtered precursor ions in each MS spectrum were fragmented via high collisional dissociation (HCD) at 30% normalized collision energy (NCE) with nitrogen gas.

## **Results and discussions**

### *Optimization of Permethylation Reaction.*

During the process of permethylation, all hydrogens on hydroxyl, amine, and carboxyl groups in ribonucleoside molecules are replaced with methyl groups. Conventionally, permethylation with DMSO and NaOH in solution is used, but the reaction efficiency hindered the application of this derivatization approach. Kang *et al.* demonstrated a solid-phase

permethylation technique by packing sodium hydroxide beads in microspin columns.<sup>33</sup> High derivatization efficiency was achieved in a short time using this technique and it has been extensively applied to improve the characterization of glycans and glycoproteins. Therefore, we employed this solid-phase-based technique to maximize the permethylation efficiency for the ribonucleoside samples (**Figure 1a**). The amount of water is critical for the permethylation reaction. To determine the optimal condition, the reaction was carried out on a mixture of four canonical ribonucleoside standards (2.5 µg/mL each) using varying volumes of water. Reactions performed in presence of 1 µL of water produced more permethylated ribonucleosides compared to reactions with 0.5 µL of water and 2 µL of water (reaction 1-3 in **Figure S1a**). We speculate that water improved the solubility of ribonucleosides in DMSO, while the extra amount of water led to undesired side reactions resulting in reduced efficiency of the permethylation. In addition, Mechref and co-workers demonstrated that reaction efficiency could be improved by adding extra iodomethane in the middle of the reaction.<sup>34</sup> As the reaction 4 shows in **Figure S1a**, the yields of permethylated ribonucleosides increased dramatically and reached a maximum after the second aliquot of iodomethane-*d*<sub>3</sub> was added to the reaction mixture.

We also monitored the unreacted ribonucleosides in the mixture that were extracted from the aqueous phase during the liquid-liquid extraction (**Figure S1b**). The results showed that there were limited unreacted ribonucleosides after the reaction in all four conditions, and the unreacted species were even less abundant after adding the second aliquot of iodomethane-*d*<sub>3</sub> (less than 0.005% for adenosine). Furthermore, the permethylation reaction may not fully replace all the active hydrogens, generating partially methylated products, which would hamper accurate quantification. For example, adenosine was noted to have an incomplete derivatization product using the conventional in-solution method.<sup>35</sup> To ensure that the permethylation goes to completion, we used LC-MS/MS to characterize the adenosine products after solid-phase derivatization. As shown in **Figure S2**, the signal of fully permethylated adenosine was exceedingly abundant, while the partially methylated adenosine abundance was at the noise level. This result demonstrated the high efficiency of the optimized solid-phase permethylation method (>99.9%). To this end, 1 µL of water and the post-reloading supplemental aliquot of 20 µL of iodomethane-*d*<sub>3</sub> was utilized for the complete derivatization of the ribonucleosides in our experiments.

#### *Construction of the dMRM Transitions.*

Commercial ribonucleoside standards were employed to create the basic transitions. Standards that were used to construct the transitions include A (adenosine), Am (2'-*O*-methyladenosine), m6A (N6-methyladenosine), t6A (N6-threonylcarbamoyladenosine), io6A (N6-(*cis*-hydroxyisopentenyl)adenosine), i6A (N6-isopentenyladenosine), I (inosine), C (cytidine), ac4C (N4-acetylcytidine), s2C (2-thiocytidine), m5C (5-methylcytidine), f5C (5-formylcytidine), G (guanosine), m7G (7-methylguanosine), U (uridine), m5U (5-methyluridine), s2U (2-thiouridine), D (dihydrouridine), m5D (5-methylidihydrouridine), and Ψ (pseudouridine). For analysis, permethylated ribonucleoside standards were prepared, and dMRM transitions of the permethylated standards were obtained by scanning their respective fragment ions using the product ion mode of the QqQ.

To distinguish the endogenous methylated molecules from the methylated molecules after derivatization process, deuterium-labeled iodomethane was used for the reaction. As shown in **Figure 1b**, five  $d_3$ -methyl ( $-CD_3$ ) groups replaced hydrogen atoms from A, including three from the hydroxyl group on ribose ring and two from the amine group on nucleobase, while m6A and Am were labeled with four molecules of  $-CD_3$ . Notably and as expected,  $-CD_3$  groups labeled hydroxyl and amine groups on m6A and Am differently. Two molecules of  $-CD_3$  on the ribose and two on the nucleobase replaced hydrogens for Am, while three hydrogens on ribose and one hydrogen on nucleobase were replaced by  $-CD_3$  for m6A. Only the hydrogens on ribose were replaced by  $-CD_3$  for N6, N6-dimethyladenosine (m6,6A). Similar to the underivatized form, the primary fragmentation of these permethylated ribonucleosides was produced by ribose ring loss, hence, three different transitions were able to be constructed,  $m/z$  353.28  $\rightarrow$  170.13 for A,  $m/z$  350.26  $\rightarrow$  167.11 for m6A,  $m/z$  350.26  $\rightarrow$  170.13 for Am, and  $m/z$  347.24  $\rightarrow$  164.09 for m6,6A. Notably, the permethylation product of m6A and m1A had the same mass with  $m/z$  350.26. To distinguish these two isomers, a further (pseudo)-MS<sup>3</sup> is required and the methyl position can be differentiated with different abundance distribution at  $m/z$  119.03 and  $m/z$  120.04 (**Figure S3**). Importantly, permethylation yielded the unmodified and methylated ribonucleosides with similar structures which have the same retention time, allowing the transition list to be built more readily. For example, after the transition and retention time of inosine were monitored based on its standard, the other methylated species, including Im, m1I, and m1Im, could also be spontaneously created without those standards.

In addition, the permethylation also improved the ability to distinguish isomers. For example, the isomer of canonical uridine,  $\Psi$ , is a crucial RNA modification and is involved in regulation of gene expression.<sup>36</sup> Different techniques have been developed for the analysis of  $\Psi$ , but the same precursor mass and poor separation of the two isomers using RP-LC have hindered their analysis. Notably, these two isomers generated two different precursor and product ions after derivatization, which suggests that U and  $\Psi$  can be unambiguously characterized using our method. As mentioned above, most of the primary fragmentation of ribonucleosides came from ribose ring loss; however,  $\Psi$  was an exception due to the carbon-carbon linkage between the ribose and the nucleobase. In order to find the fragments of  $\Psi$ , we analyzed permethylated U and  $\Psi$  standards using a high-resolution Orbitrap MS. As shown in **Figure 2a**, U lost the methylated ribose ring (-180.14) and yielded its unique reporter ion of  $m/z$  130.07. Interestingly,  $m/z$  149.11 was the most abundant fragment for U, while it was not selected as the product ion. This is because it was generated from the cross-ring fragmentation of permethylated ribose and could also be produced by all other permethylated ribonucleosides. Additionally, the fragment ions of  $m/z$  206.15 and 225.11 were uniquely produced by  $\Psi$  due to the cross-ring fragmentation (**Figure 2b**). Therefore, the transitions  $m/z$  313.22  $\rightarrow$  130.07 and  $m/z$  330.25  $\rightarrow$  206.15 were created for U and  $\Psi$ , respectively. It is also challenging to distinguish the two methylated uridine isomers, m5U and m3U, using the conventional method. After permethylation, m5U had one more  $-CD_3$  molecule labeled compared to m3U, which yielded differences in precursor ion mass, product ion mass, and retention time.

We further monitored three adenosine-derived ribonucleoside standards i6A, io6A, and t6A, which have more complicated nucleoside base structures than methylated adenosine.



Similar to the unmodified adenosine, the  $-CD_3$  groups labeled hydroxyl groups in the ribose and amines in the nucleobase of i6A (**Figure S4a**), while an extra  $-CD_3$  group reacted with io6A due to its hydroxyisopentenyl group (**Figure S4b**). Although permethylated io6A also generated the i6A ion in high intensity due to fragmentation of the  $-OCD_3$ , both standards still generated high abundance fragments from ribose loss. Threonylcarbamoyl-modified adenosine, t6A, contains an acidic structure. The carboxyl group reacted with iodomethane- $d_3$  and generated a product with  $m/z$  532.38 as the precursor mass, and the expected fragment at  $m/z$  349.23 was monitored in positive ion mode (**Figure S4c**). One exception is positively charged m7G, where a hydroxyl group was added from the hydroxide (**Figure S5a**). As a result, one extra methoxy group was yielded, and m7G was monitored using  $m/z$  435.37 as the precursor ion and  $m/z$  252.22 as the product ion (**Figure S5b**). In general, same as the glycans,  $-CD_3$  group can completely replace all the hydrogens on hydroxyl, amine, and carboxyl groups for ribonucleosides. This suggested that the products of derivatization are able to be easily monitored for structurally complex RNA modifications. Overall, more than 60 transitions were monitored simultaneously and shown in **Supplementary Table 1**, and the qualifying fragment ions of the permethylated ribonucleosides were also chosen to validate the method.

#### *Optimization of Collision Energies.*

For underivatized ribonucleosides, low collision energy ( $\sim 10$  eV) was utilized for the identification and quantification due to the unstable nature of the ribonucleosides. Because the permethylation can enhance the stability of the precursor ions in the gas phase, the energy required for the fragmentation needed to be optimized. Therefore, we optimized the collision energies ranging from 10 to 35 eV for the permethylated ribonucleosides to obtain the best signal for the quantifying fragment ions. As shown in **Figure S6**, responses of most permethylated ribonucleosides first increased with rising collision energy, caused by more efficient fragmentation of the precursor ions. Signals decreased when the collision energy was set greater than 20 eV, which may be due to the larger percentage of over-fragmentation under the high-energy collision conditions. As a result, we found that 15-20 eV yielded the best signals for the majority of the permethylated ribonucleosides.

#### *Comparison with conventional ribonucleoside analysis.*

The ribonucleosides have similar structures, which contain hydrophilic ribose and a nucleobase. Therefore, it is a challenge to profile and quantify these compounds by traditional RP-LC-MS methods. To evaluate the advantages of the derivatized ribonucleosides compared to the underivatized form, equal amounts of standards were injected into the instrument and analyzed. As shown in **Figure 3a** and **b**, all four unpermethylated ribonucleoside standards were eluted within 4 mins under 1% of ACN. Notably, the chromatogram of these ribonucleosides had problems in peaks overlapping during the analysis. For example, the more hydrophilic ribonucleosides, cytidine and uridine had poor retention on C18 and were eluted in 0% ACN, which risks co-elution with hydrophilic contaminants. On the other hand, although purines including adenosine and guanosine retained on the reversed-phase, they bound weakly, so they also had poor chromatographic resolution. Ionization suppression therefore can consequently reduce sensitivity of hydrophilic ribonucleosides, so separation of the nucleosides is critical. Meanwhile, all the permethylated ribonucleoside peaks were evenly distributed across the

chromatogram with a higher percentage of ACN (~30% ACN) (**Figure 3c and d**), which indicated the chromatographic resolution of ribonucleosides on C18 was improved after labeling with the  $-CD_3$  groups. The theoretical plate number was calculated based on the retention time and the full peak width and shown in **Supplementary Table 2**. As a result, the number of theoretical plates for conventional ribonucleoside analysis was around 4000, while it was greatly improved for permethylated ribonucleosides even when a much higher percentage of ACN was applied. For example, the theoretical plate number was more than 26000 for permethylated guanosine, and nearly 16000 for permethylated adenosine. This demonstrated that C18 can continue to be the column resin of choice for nucleoside analysis without suffering poor resolution and analyte retention. Therefore, permethylation would allow users to avoid cumbersome HILIC chromatography for the analysis of RNA modifications.

Due to the hydrophilicity of the ribonucleosides, the compounds normally had low positive mode ionization efficiencies that resulted in difficulties in quantitative measurement.<sup>37</sup> It has been demonstrated that permethylation can enhance the signals of hydrophilic analytes.<sup>38</sup> Indeed, the signal of permethylated ribonucleosides was largely boosted compared to the underivatized ribonucleosides (**Figure 3 e and f**). For example, the peak area of the adenosine increased as much as 20-fold. This is likely due to the high hydrophilicity of underivatized ribonucleosides leading to lower proton affinity. At the same time, the incorporation of a large number of methyl groups with increased hydrophobicity can provide higher ionization efficiency in positive electrospray ionization. Uridine still had the lowest ionization efficiency among four canonical ribonucleosides. Nonetheless uridine signal after the permethylation increased 5-fold compared to the unpermethylated form, which is particularly valuable for analyzing uridine-based ribonucleosides that are normally difficult to detect. In general, the process of permethylation improved ribonucleoside quantitative analysis with great retaining performance, better separation, and higher sensitivity.

#### *Method validation.*

Quantitative analysis of canonical and modified ribonucleosides was performed to validate the derivatization method. The standard ribonucleoside mixtures were prepared by mixing the standards together and diluting solutions to the concentration range of 0.01–1000 ng/L. Working solutions with different dilution factors were derivatized according to the optimized solid-phase permethylation workflow. The calibration curve, linear regression coefficient ( $R^2$ ), linear range, limits of detection (LOD), limits of quantification (LOQ), and coefficients of variance (CV) for the permethylated ribonucleosides were calculated and shown in **Supplementary Table 3**. As a result, adenosine had the highest responses among all the standards due to the higher ionization efficiency of the molecule. All the correlation coefficients for the analytes were between 0.992-0.998, which indicated very good linearity of this method. Besides, both LOD and LOQ values were mostly at sub-femtomole levels, which was one order of magnitude lower compared to the analysis for underivatized ribonucleosides.<sup>39</sup> Notably, LOD and LOQ for dihydrouridine and 5-methylhydrouridine were higher than other ribonucleosides, which might be because the dihydrouridine is sensitive to the alkaline conditions wherein the pyrimidine ring opened and yielded by-products.<sup>40</sup> This result suggested that this method is suitable for analyzing broad RNA modifications found in low abundance in biological samples. Furthermore, the linear calibration ranges for the permethylated

ribonucleosides demonstrated good linearity within three orders of magnitude, while the technical replicate CVs were nearly 4% for most ribonucleosides. Collectively, this method showed an improved quantitative analysis compared to the ribonucleoside analysis without permethylation.<sup>21</sup>

#### *Analysis of ribonucleosides from mESCs.*

To demonstrate the applicability of the method, we quantified the levels of C, Cm, hm5C, and m5C in RNAs extracted from mESCs carrying mutations in genes encoding TET dioxygenases as well as wildtype (WT) controls. TET dioxygenases, which include TET1, TET2, and TET3, play important roles in epigenetic regulation because they catalyze the oxidation of 5-methylcytosine in DNA to form oxidized products.<sup>41</sup> Recently, He et al. demonstrated that TET2 is also necessary for the conversion of m5C to hm5C on tRNAs, which regulates the formation of tRNA fragments.<sup>32</sup> In this previous study, hm5C was measured using native ribonucleoside MS analysis and was shown to be substantially decreased in TET mutants. However, our previous analyses were not sufficiently accurate to detect the expected corresponding increase in m5C.

We repeated these measurements using our improved permethylation-based method (**Figure 4a and b**). Consistent with our published results, we detected a significant decrease in hm5C in RNA obtained from TET2-deficient mESCs, which was even more pronounced in RNA from triple *Tet1/2/3* knockout (KO) cells (**Figure 4c**).<sup>32</sup> Notably, our improved method detected 10-fold higher levels of hm5C in these samples compared to our previous results based on native ribonucleoside analyses, likely due to improved sensitivity and CVs, as discussed above. We were also able to detect the expected increase in m5C in both *Tet2* KO and *Tet1/2/3* triple KO as compared to WT cells (**Figure 4d**). As a control, levels of Cm, which is not known to be targeted by TET enzymes, were not affected by the mutations (**Figure 4e**).

These observations were in agreement with our previous findings and provided more comprehensive information, including the level of m5C and Cm, simultaneously. Overall, these results exemplify the advantage of improved detection and quantification when performing the analysis of ribonucleoside modifications using the permethylation workflow.

#### **Conclusions**

The analysis of the RNA modifications suffers from the hydrophilicity of the molecules, which can result in low sensitivity and inaccurate quantification. In this study, we derivatized ribonucleosides to hydrophobic molecules by using solid-phase permethylation. A dynamic MRM method was developed to provide better quantitative characterization for RNA modifications. The high reproducibility of the quantitative results suggests that this method essentially improves upon the conventional analysis and provides a powerful tool to monitor the ribonucleoside modifications on RNAs. Overall, this method is capable of profiling more than 60 modified ribonucleosides in RNAs with high throughput and the transition list can be readily extended as more RNA modifications are discovered. Similar to applications in glycan research, different isotopically labeled iodomethane can be further applied as multiplexing reagents to enhance the quantification of multiple samples simultaneously.<sup>42</sup> This method can also be potentially employed for detecting canonical and modified deoxyribonucleosides in DNA samples, which provides a tool for facilitating both epigenetic and epitranscriptomic

analysis. Additionally, the quantitative results of ribonucleosides extracted from WT and TET KO mESCs demonstrate that this approach can detect differences in low-abundant, but important, RNA modifications in biological samples. As ribonucleoside modifications are associated with several diseases, this method provides a new reliable platform for biomarker research and discovery.

### **Conflict of interest**

The authors declare that they have no conflicts of interest with the contents of this article.

### **Author contributions**

Y.X. and K.A.J. designed and performed the experiments, analyzed data, produced all figures, and drafted the manuscript. A.S. and R.B. prepared TET KO mESCs, extracted RNAs for the analysis, analyzed data, and edited the manuscript. R.B. supervised A.S. B.A.G. planned the overall experimental project and co-wrote the manuscript.

### **Acknowledgements**

The authors thank Dr. Juan J. Castillo and Siyu Chen (University of California, Davis) for their suggestions on permethylation reaction. Research reported in this publication was supported by grant to BAG from the NIH AI118891, CA196539 and AG031862.

## References

1. Li, S.; Mason, C. E., The Pivotal Regulatory Landscape of RNA Modifications. *Annual Review of Genomics and Human Genetics* **2014**, *15* (1), 127-150.
2. Peer, E.; Rechavi, G.; Dominissini, D., Epitranscriptomics: regulation of mRNA metabolism through modifications. *Current Opinion in Chemical Biology* **2017**, *41*, 93-98.
3. Xiong, X.; Yi, C.; Peng, J., Epitranscriptomics: Toward A Better Understanding of RNA Modifications. *Genomics, Proteomics & Bioinformatics* **2017**, *15* (3), 147-153.
4. Boccaletto, P.; Machnicka, M. A.; Purta, E.; Piątkowski, P.; Bagiński, B.; Wirecki, T. K.; de Crécy-Lagard, V.; Ross, R.; Limbach, P. A.; Kotter, A.; Helm, M.; Bujnicki, J. M., MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Research* **2018**, *46* (D1), D303-D307.
5. Yi, C.; Pan, T., Cellular Dynamics of RNA Modification. *Accounts of Chemical Research* **2011**, *44* (12), 1380-1388.
6. Tuorto, F.; Liebers, R.; Musch, T.; Schaefer, M.; Hofmann, S.; Kellner, S.; Frye, M.; Helm, M.; Stoecklin, G.; Lyko, F., RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nature Structural & Molecular Biology* **2012**, *19* (9), 900-905.
7. Helm, M., Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Research* **2006**, *34* (2), 721-733.
8. Frye, M.; Harada Bryan, T.; Behm, M.; He, C., RNA modifications modulate gene expression during development. *Science* **2018**, *361* (6409), 1346-1349.
9. Nechay, M.; Kleiner, R. E., High-throughput approaches to profile RNA-protein interactions. *Current Opinion in Chemical Biology* **2020**, *54*, 37-44.
10. Roundtree, I. A.; Evans, M. E.; Pan, T.; He, C., Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **2017**, *169* (7), 1187-1200.
11. Fang, Z.; Hu, Y.; Hu, J.; Huang, Y.; Zheng, S.; Guo, C., The crucial roles of N6-methyladenosine (m6A) modification in the carcinogenesis and progression of colorectal cancer. *Cell & Bioscience* **2021**, *11* (1), 72.
12. Krogh, N.; Nielsen, H., Sequencing-based methods for detection and quantitation of ribose methylations in RNA. *Methods* **2019**, *156*, 5-15.
13. Garalde, D. R.; Snell, E. A.; Jachimowicz, D.; Sipos, B.; Lloyd, J. H.; Bruce, M.; Pantic, N.; Admassu, T.; James, P.; Warland, A.; Jordan, M.; Ciccone, J.; Serra, S.; Keenan, J.; Martin, S.; McNeill, L.; Wallace, E. J.; Jayasinghe, L.; Wright, C.; Blasco, J.; Young, S.; Brocklebank, D.; Juul, S.; Clarke, J.; Heron, A. J.; Turner, D. J., Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods* **2018**, *15* (3), 201-206.
14. Helm, M.; Motorin, Y., Detecting RNA modifications in the epitranscriptome: predict and validate. *Nature Reviews Genetics* **2017**, *18* (5), 275-291.
15. Yoluç, Y.; Ammann, G.; Barraud, P.; Jora, M.; Limbach, P. A.; Motorin, Y.; Marchand, V.; Tisné, C.; Borland, K.; Kellner, S., Instrumental analysis of RNA modifications. *Critical Reviews in Biochemistry and Molecular Biology* **2021**, *56* (2), 178-204.
16. Becette Owen, B.; Zong, G.; Chen, B.; Taiwo Kehinde, M.; Case David, A.; Dayie, T. K., Solution NMR readily reveals distinct structural folds and interactions in doubly <sup>13</sup>C- and <sup>19</sup>F-labeled RNAs. *Science Advances* **6** (41), eabc6572.
17. Wetzel, C.; Limbach, P. A., Mass spectrometry of modified RNAs: recent developments. *Analyst* **2016**, *141* (1), 16-23.
18. Sutton, J. M.; Guimaraes, G. J.; Annavarapu, V.; van Dongen, W. D.; Bartlett, M. G., Current State of Oligonucleotide Characterization Using Liquid Chromatography–Mass Spectrometry: Insight into Critical Issues. *Journal of the American Society for Mass Spectrometry* **2020**, *31* (9), 1775-1782.
19. Hagelskamp, F.; Borland, K.; Ramos, J.; Hendrick, A. G.; Fu, D.; Kellner, S., Broadly applicable oligonucleotide mass spectrometry for the analysis of RNA writers and erasers in vitro. *Nucleic Acids Research* **2020**, *48* (7), e41-e41.
20. Wein, S.; Andrews, B.; Sachsenberg, T.; Santos-Rosa, H.; Kohlbacher, O.; Kouzarides, T.; Garcia, B. A.; Weisser, H., A computational platform for high-throughput analysis of RNA sequences and modifications by mass spectrometry. *Nature Communications* **2020**, *11* (1), 926.

21. Su, D.; Chan, C. T. Y.; Gu, C.; Lim, K. S.; Chionh, Y. H.; McBee, M. E.; Russell, B. S.; Babu, I. R.; Begley, T. J.; Dedon, P. C., Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nature Protocols* **2014**, *9* (4), 828-841.
22. Gregorova, P.; Sipari, N. H.; Sarin, L. P., Broad-range RNA modification analysis of complex biological samples using rapid C18-UPLC-MS. *RNA Biology* **2021**, *18* (10), 1382-1389.
23. Tuytten, R.; Lemièrre, F.; Van Dongen, W.; Witters, E.; Esmans, E. L.; Newton, R. P.; Dudley, E., Development of an On-Line SPE-LC–ESI-MS Method for Urinary Nucleosides: Hyphenation of Aprotic Boronic Acid Chromatography with Hydrophilic Interaction LC–ESI-MS. *Analytical Chemistry* **2008**, *80* (4), 1263-1271.
24. Sarin, L. P.; Kienast, S. D.; Leufken, J.; Ross, R. L.; Dziergowska, A.; Debiec, K.; Sochacka, E.; Limbach, P. A.; Fufezan, C.; Drexler, H. C. A.; Leidel, S. A., Nano LC-MS using capillary columns enables accurate quantification of modified ribonucleosides at low femtomol levels. *RNA* **2018**, *24* (10), 1403-1417.
25. Wu, J.; Zhang, Y.; Wiegand, R.; Wang, J.; Bepler, G.; Li, J., Quantitative analysis of intracellular nucleoside triphosphates and other polar metabolites using ion pair reversed-phase liquid chromatography coupled with tandem mass spectrometry. *Journal of Chromatography B* **2015**, *1006*, 167-178.
26. Patteson, K. G.; Rodicio, L. P.; Limbach, P. A., Identification of the mass-silent post-transcriptionally modified nucleoside pseudouridine in RNA by matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Research* **2001**, *29* (10), e49-e49.
27. Huang, W.; Lan, M.-D.; Qi, C.-B.; Zheng, S.-J.; Wei, S.-Z.; Yuan, B.-F.; Feng, Y.-Q., Formation and determination of the oxidation products of 5-methylcytosine in RNA. *Chemical Science* **2016**, *7* (8), 5495-5502.
28. Li, S.; Jin, Y.; Tang, Z.; Lin, S.; Liu, H.; Jiang, Y.; Cai, Z., A novel method of liquid chromatography–tandem mass spectrometry combined with chemical derivatization for the determination of ribonucleosides in urine. *Analytica Chimica Acta* **2015**, *864*, 30-38.
29. Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wührer, M., Glycan labeling strategies and their use in identification and quantification. *Analytical and Bioanalytical Chemistry* **2010**, *397* (8), 3457-3481.
30. Pantarotto, C.; Martini, A.; Belvedere, G.; Bossi, A.; Donnelly, M. C.; Frigerio, A., Application of gas chromatography–chemical ionization mass fragmentography in the evaluation of bases and nucleoside analogues used in cancer chemotherapy. *Journal of Chromatography A* **1974**, *99*, 519-527.
31. Von Minden, D. L.; McCloskey, J. A., Mass spectrometry of nucleic acid components. N,O-Permethyl derivatives of nucleosides. *Journal of the American Chemical Society* **1973**, *95* (22), 7480-7490.
32. He, C.; Bozler, J.; Janssen, K. A.; Wilusz, J. E.; Garcia, B. A.; Schorn, A. J.; Bonasio, R., TET2 chemically modifies tRNAs and regulates tRNA fragment levels. *Nature Structural & Molecular Biology* **2021**, *28* (1), 62-70.
33. Kang, P.; Mechref, Y.; Klouckova, I.; Novotny, M. V., Solid-phase permethylation of glycans for mass spectrometric analysis. *Rapid Communications in Mass Spectrometry* **2005**, *19* (23), 3421-3428.
34. Zhou, S.; Wooding, K. M.; Mechref, Y., Analysis of Permethyated Glycan by Liquid Chromatography (LC) and Mass Spectrometry (MS). In *High-Throughput Glycomics and Glycoproteomics: Methods and Protocols*, Lauc, G.; Wührer, M., Eds. Springer New York: New York, NY, 2017; pp 83-96.
35. Dolhun, J. J.; Wiebers, J. L., Mass spectra of nucleoside derivatives. *Organic Mass Spectrometry* **1970**, *3* (5), 669-681.
36. Zhao, Y.; Dunker, W.; Yu, Y.-T.; Karijovich, J., The Role of Noncoding RNA Pseudouridylation in Nuclear Gene Expression Events. *Frontiers in Bioengineering and Biotechnology* **2018**, *6* (8).
37. Lan, M.-D.; Yuan, B.-F.; Feng, Y.-Q., Deciphering nucleic acid modifications by chemical derivatization-mass spectrometry analysis. *Chinese Chemical Letters* **2019**, *30* (1), 1-6.
38. Ruhaak, L. R.; Xu, G.; Li, Q.; Goonatilake, E.; Lebrilla, C. B., Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chemical Reviews* **2018**, *118* (17), 7886-7930.
39. Fang, Z.; Hu, Y.; Chen, J.; Xu, K.; Wang, K.; Zheng, S.; Guo, C., Mass Spectrometry-Based Targeted Serum Monomethylated Ribonucleosides Profiling for Early Detection of Breast Cancer. *Frontiers in Molecular Biosciences* **2021**, *8* (833).
40. Motorin, Y.; Muller, S.; Behm - Ansmant, I.; Branlant, C., Identification of Modified Residues in RNAs by Reverse Transcription - Based Methods. In *Methods in Enzymology*, Academic Press: 2007; Vol. 425, pp 21-53.
41. Wu, H.; Zhang, Y., Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes & Development* **2011**, *25* (23), 2436-2452.

42. Dong, X.; Peng, W.; Yu, C.-Y.; Zhou, S.; Donohoo, K. B.; Tang, H.; Mechref, Y., 8-plex LC–MS/MS Analysis of Permethylated N-Glycans Achieved by Using Stable Isotopic Iodomethane. *Analytical Chemistry* **2019**, *91* (18), 11794-11802.

## Figures

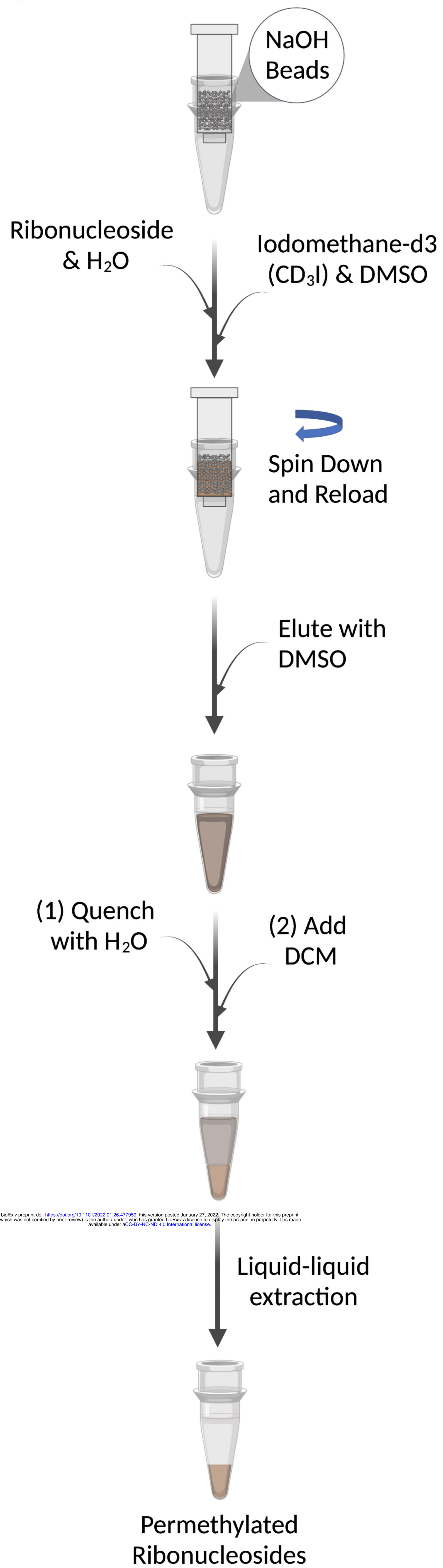
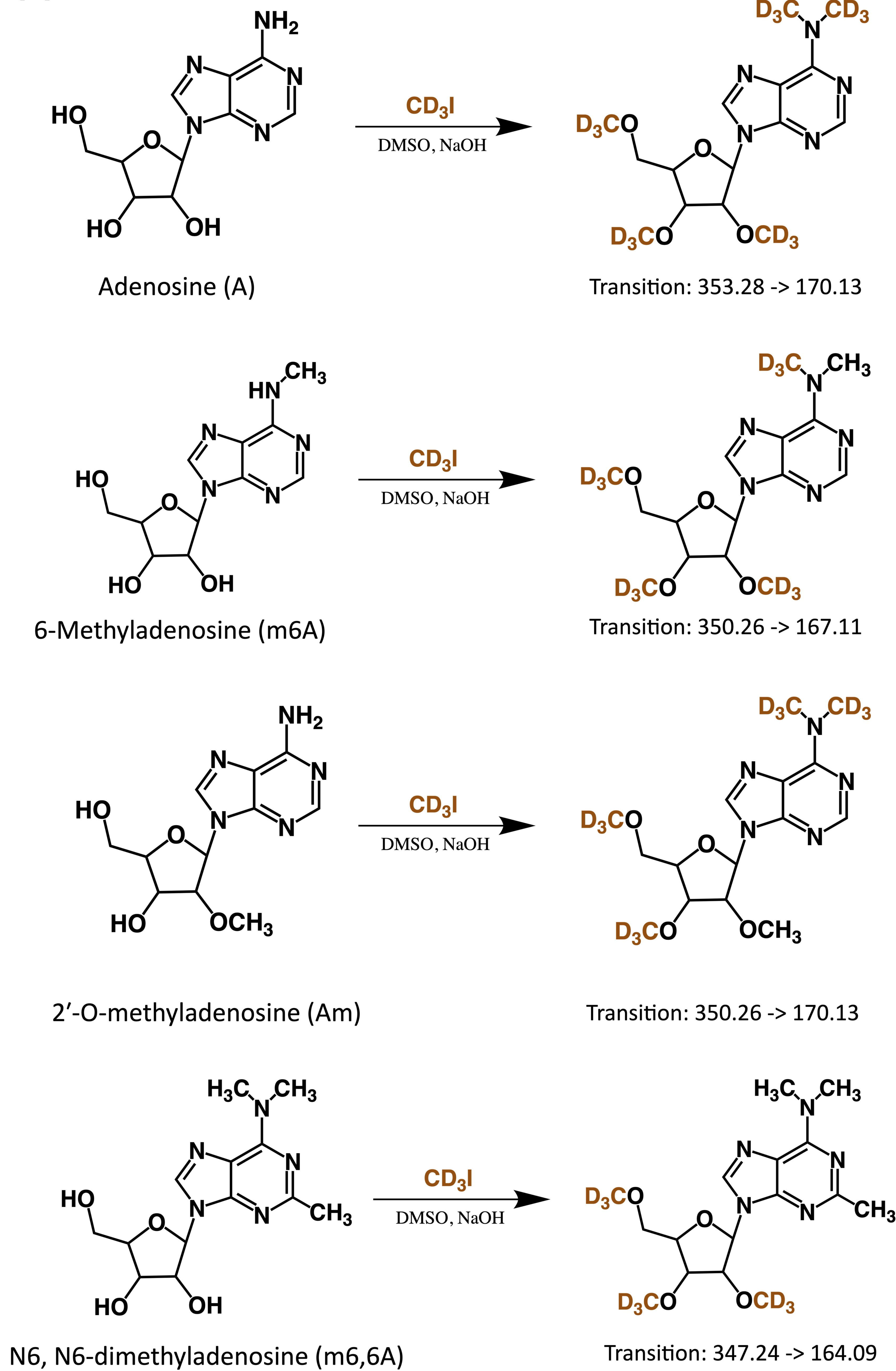
**Figure 1. Permethylation was used to derivatize the ribonucleosides. (a)** The workflow of the solid-phase permethylation method. **(b)** The example of permethylation of unmodified and methylated adenosine. Adenosine, 6-methyladenosine, and 2'-*O*-methyladenosine were labeled by the -CD<sub>3</sub> group differently, and their unique precursor and product ions were monitored using UHPLC-QqQ-MS under the dMRM mode.

**Figure 2. LC-MS/MS spectra of two permethylated isomers: uridine (U) and pseudouridine (Ψ).** MS/MS spectra of the methylated(*d*<sub>3</sub>)-labeled protonated precursor ion **(a)** *m/z* 313.22 for U and **(b)** *m/z* 330.25 for pseudouridine Ψ. Fragmentation by HCD of U resulted in signals *m/z* 149.11 and *m/z* 130.07 as the two most abundant fragment ions, while the fragmentation of Ψ produced signals *m/z* 206.15 and *m/z* 22.5.11 as the two most abundant fragment ions.

**Figure 3. Comparison of the analysis between permethylated and native ribonucleoside.** MRM chromatogram of **(a)** permethylated(*d*<sub>3</sub>) ribonucleoside standards **(c)** native ribonucleoside standards. The LC gradient was applied for analyzing **(b)** permethylated(*d*<sub>3</sub>) canonical ribonucleosides and **(d)** the native forms. The ACN% (B%) was much higher for permethylated(*d*<sub>3</sub>) ribonucleosides analysis and the separation of analytes was improved. **(e)** The comparison of peak area and peak height between permethylated and native ribonucleosides.

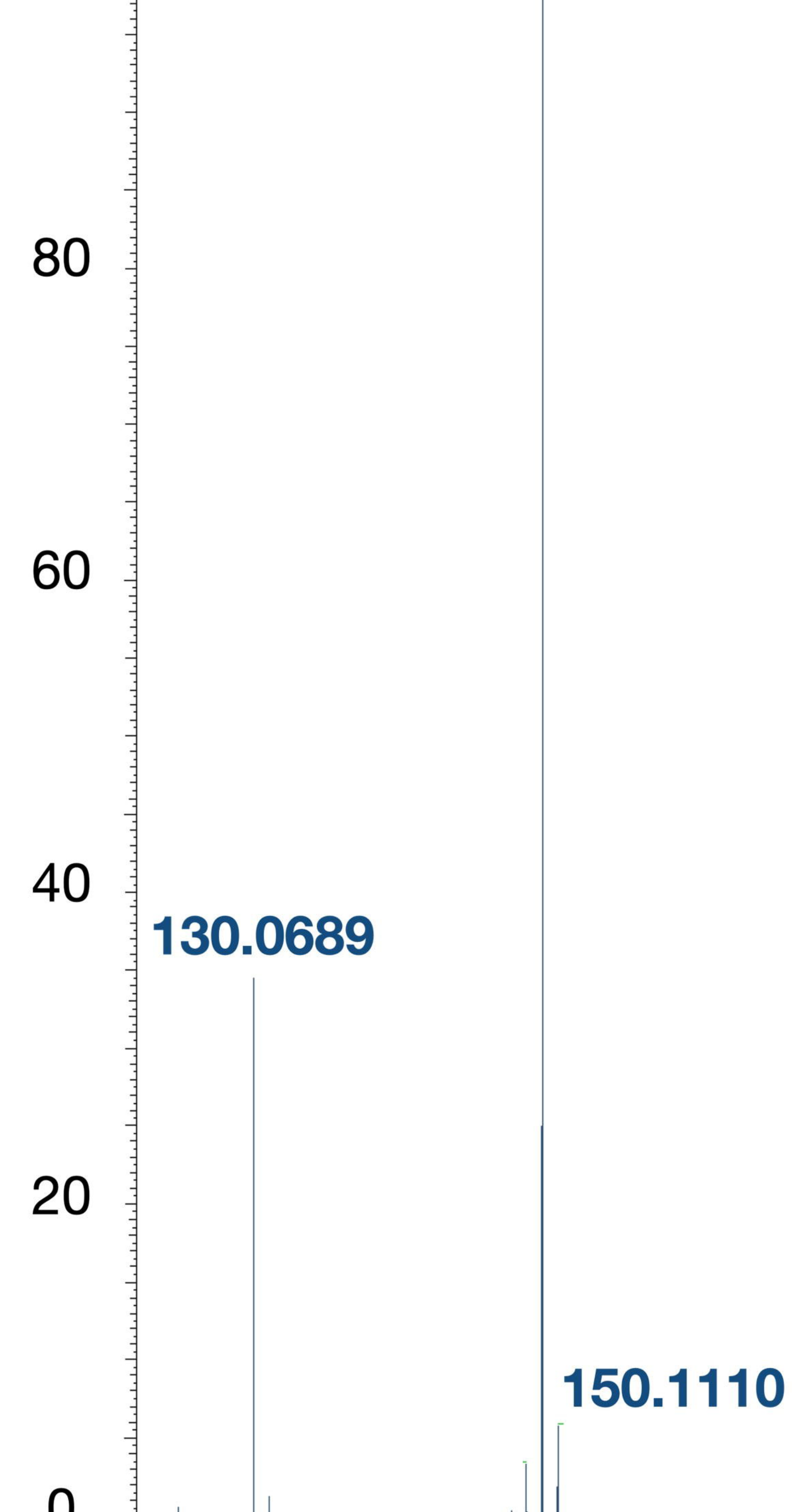
**Figure 4. The application of the method for analyzing modified cytosine in WT and TET KO mESCs.** The example chromatography of C, Cm, m5C, and hm5C in purified total RNAs under the condition of **(a)** WT and **(b)** TET2 KO. **(c)** The relative abundance of hm5C, m5C, and Cm. (p values were determined using two-tailed Student's t test for unpaired samples. Error bars represent mean ± s.d., n = 3, \*\* p < 0.01, \*\*\*\* p < 0.001, n.s. means p > 0.05.)



**(a)****(b)**



**(a)** 149.1077



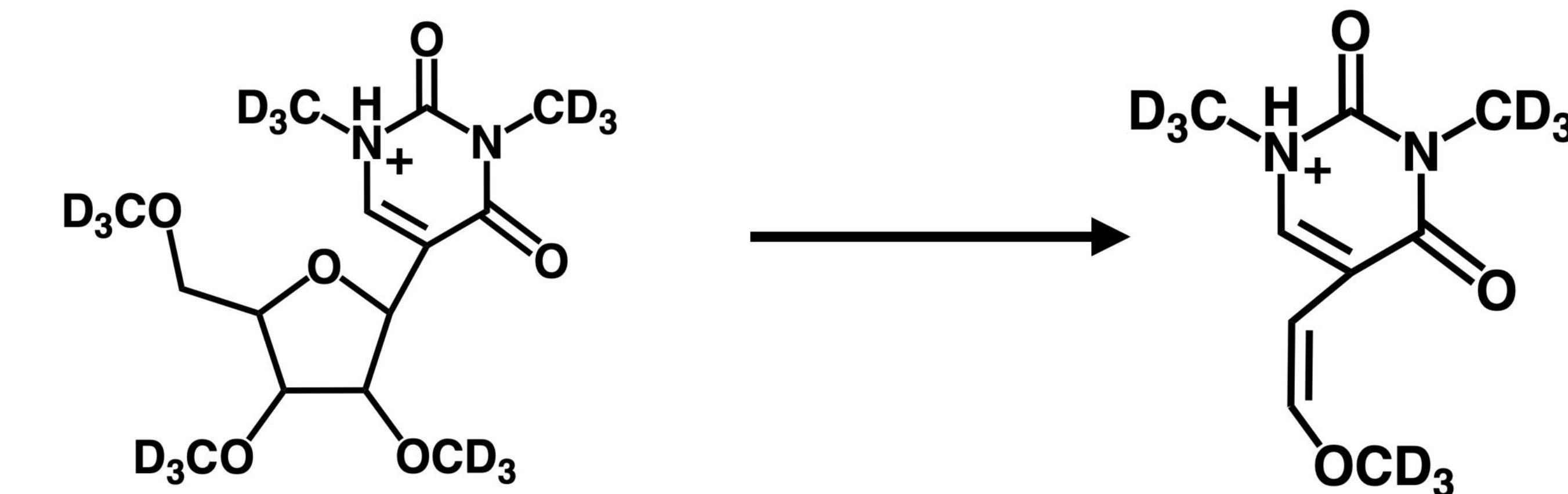
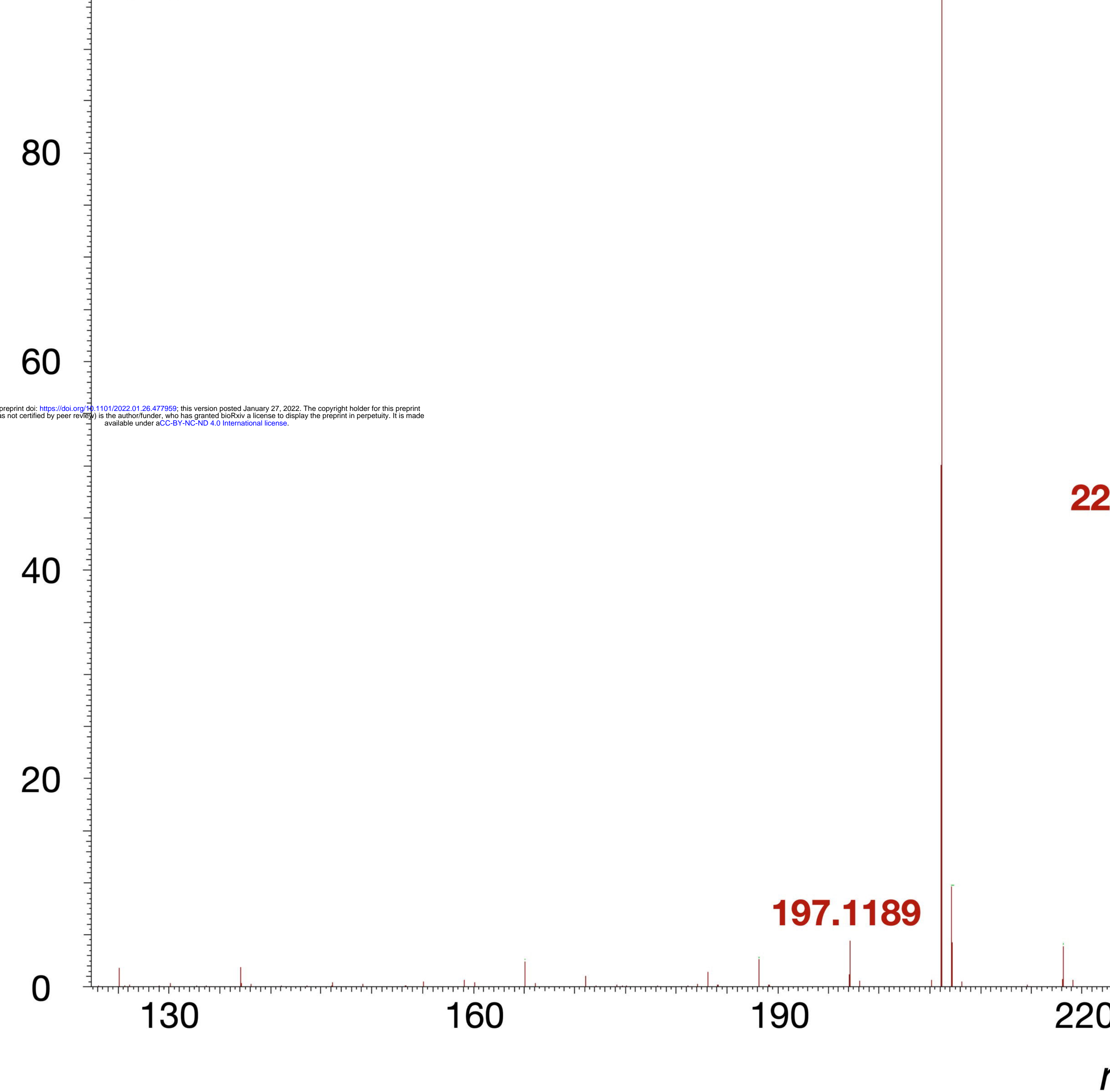
**uridine  
permethylated (d3)**

$m/z$ : 313.22

$m/z$ : 130.07

Relative Abundance

**(b)** 206.1482

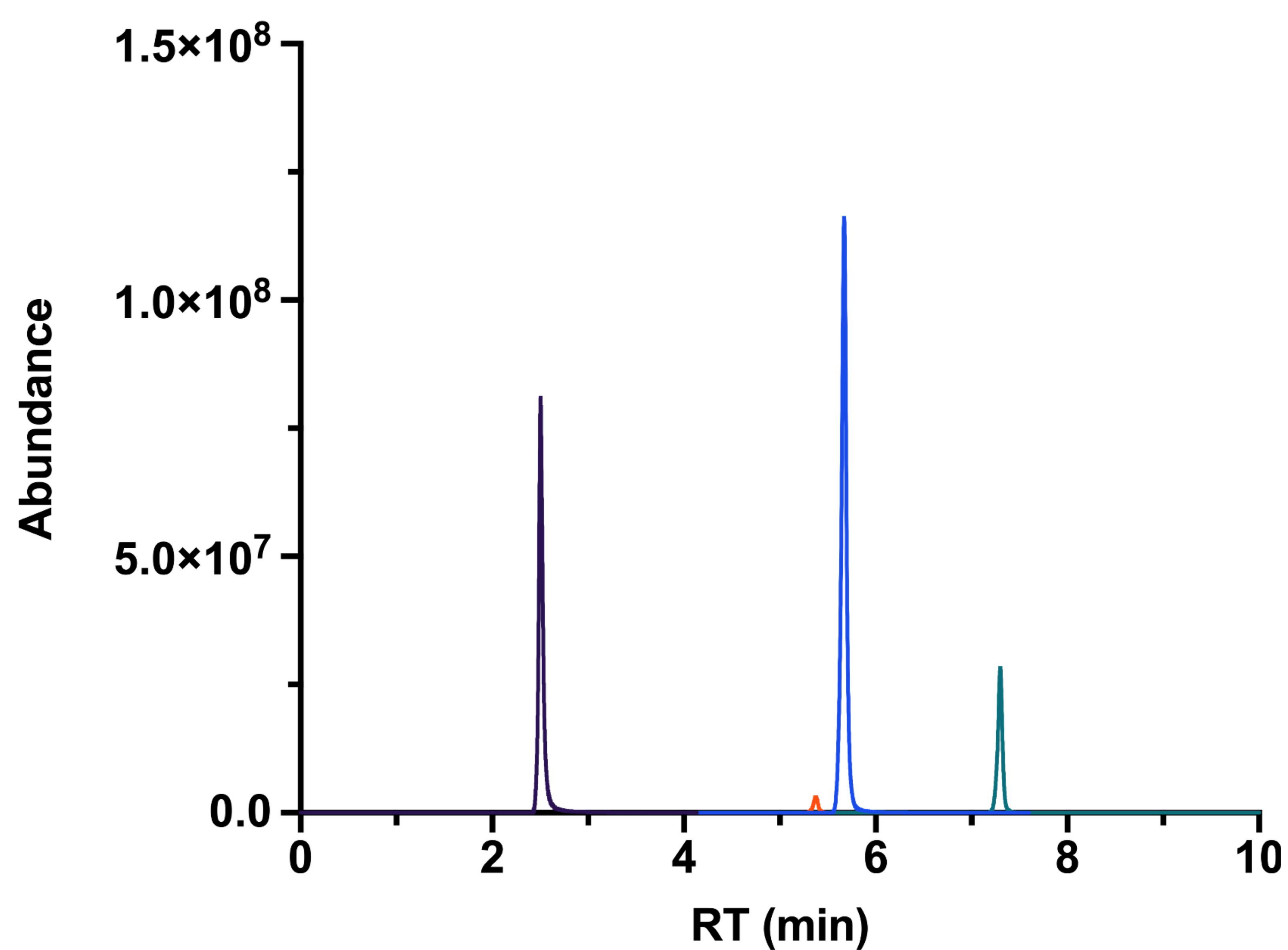


**pseudouridine  
permethylated (d3)**

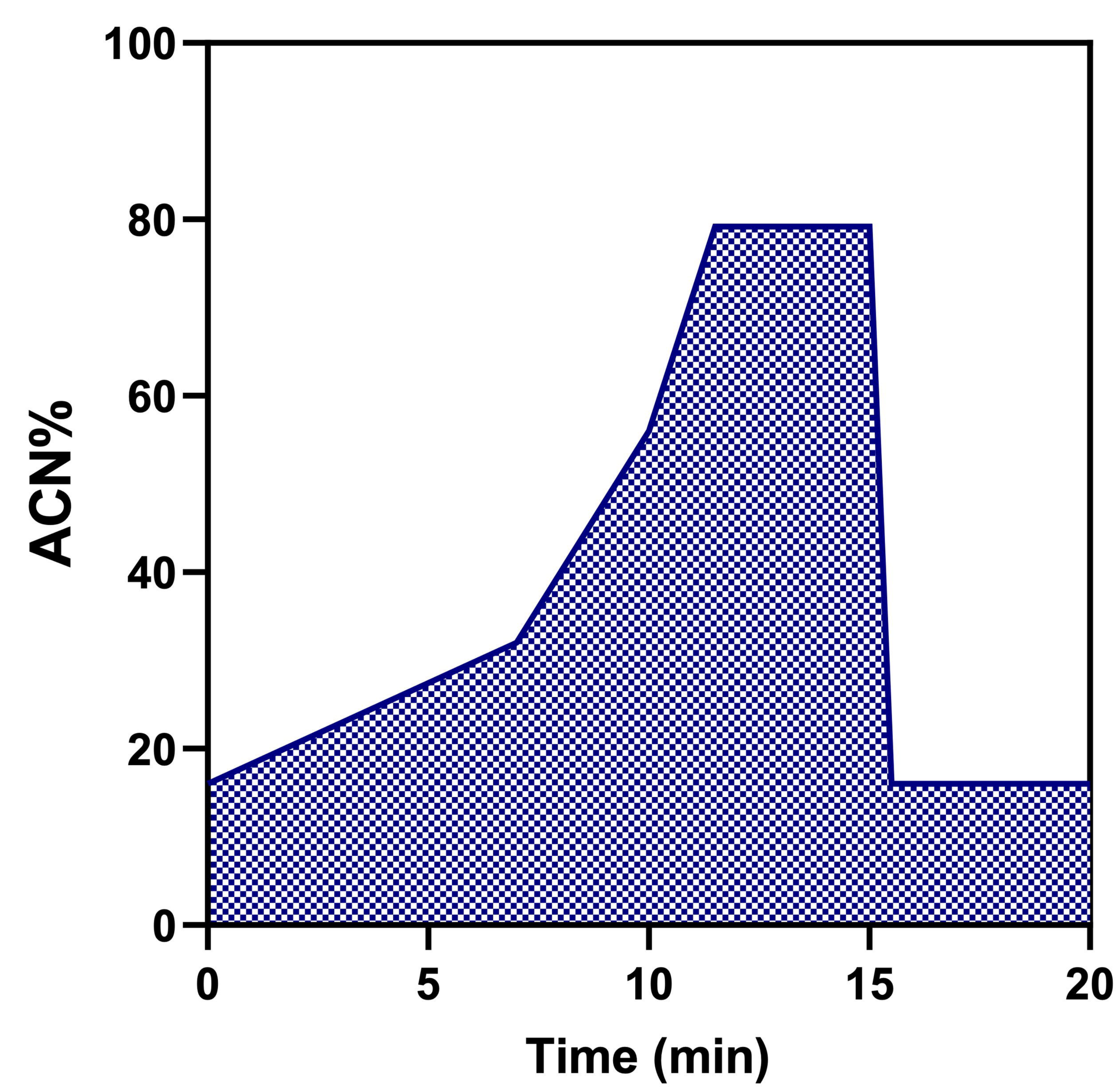
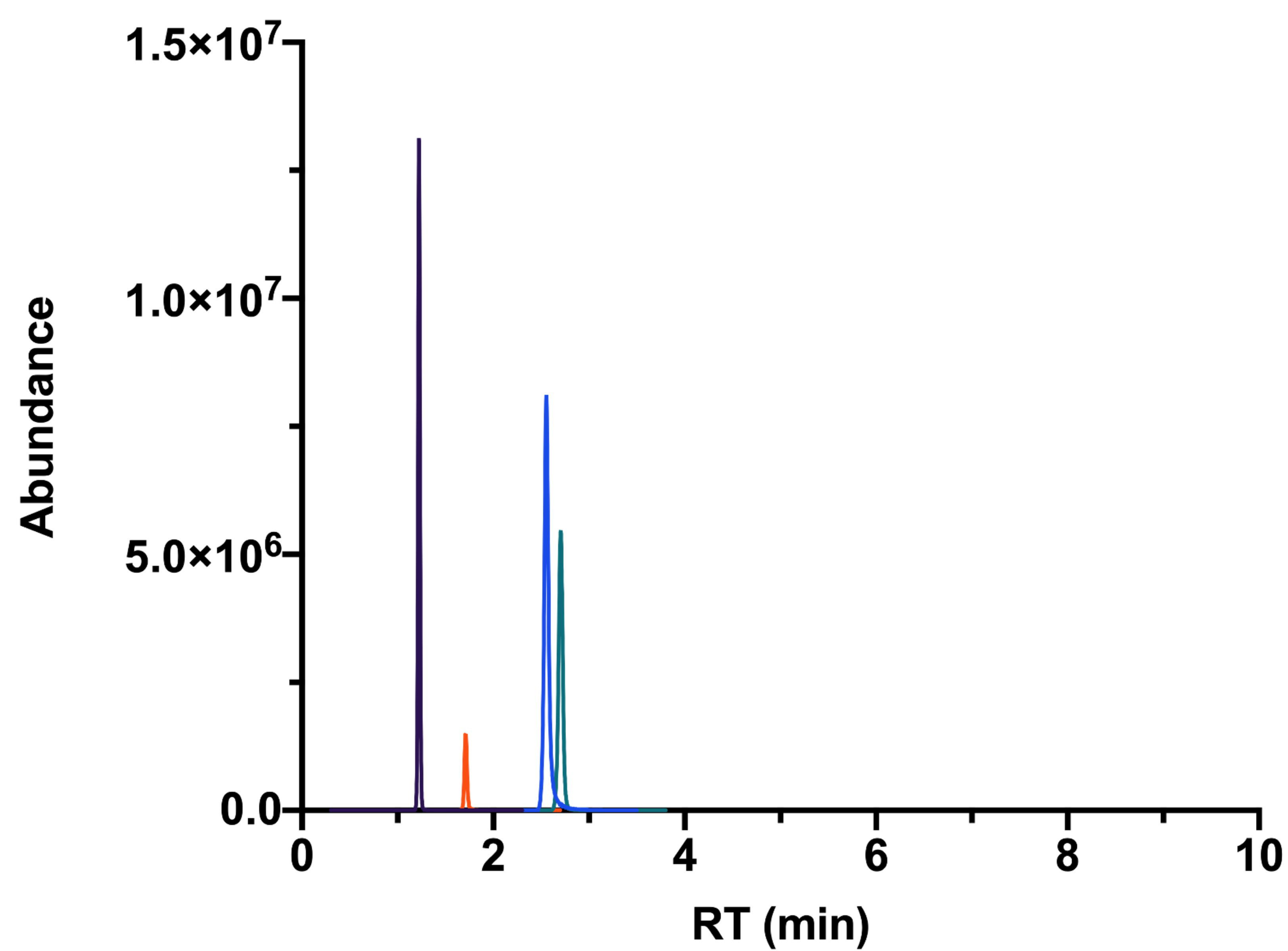
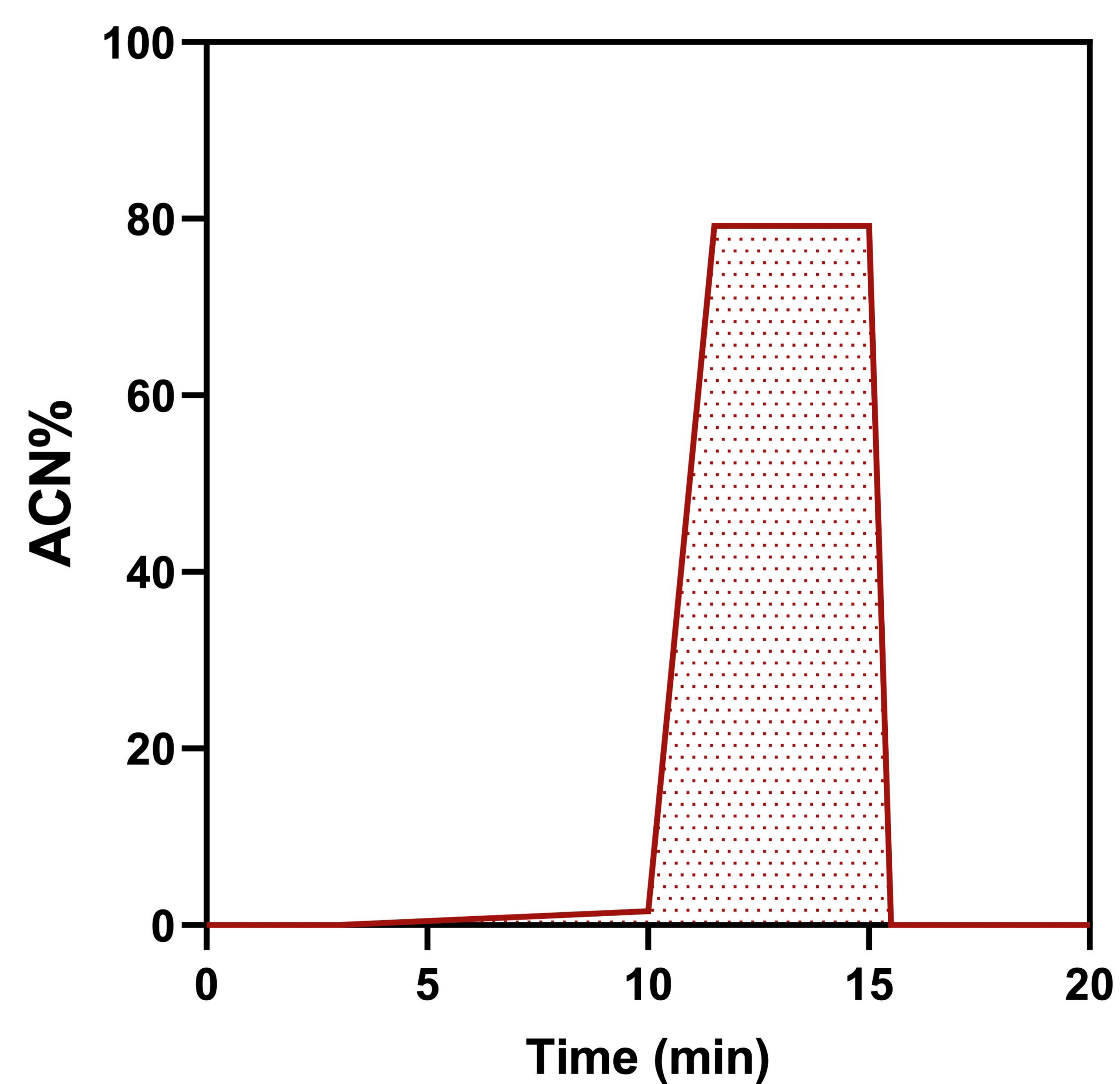
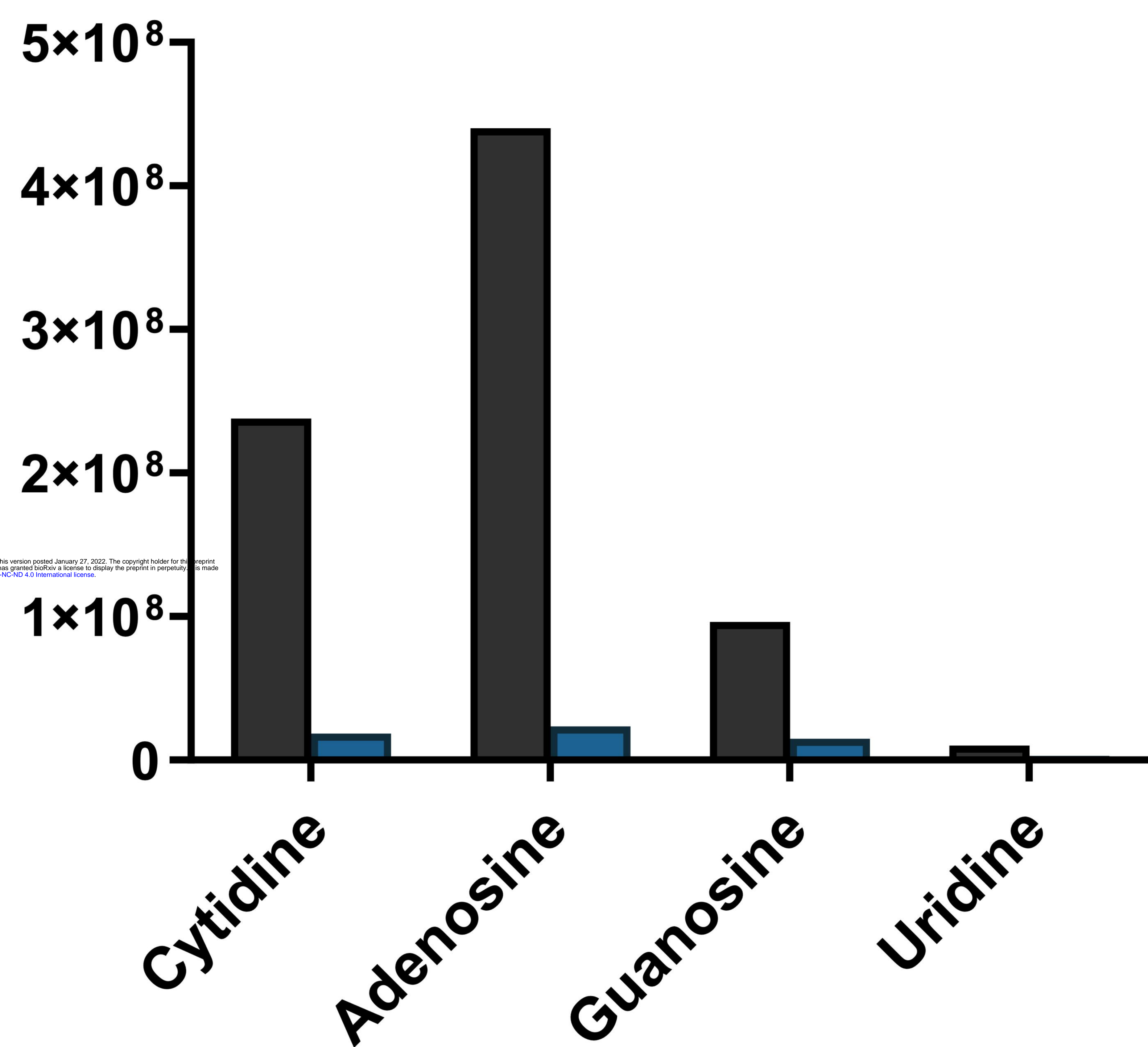
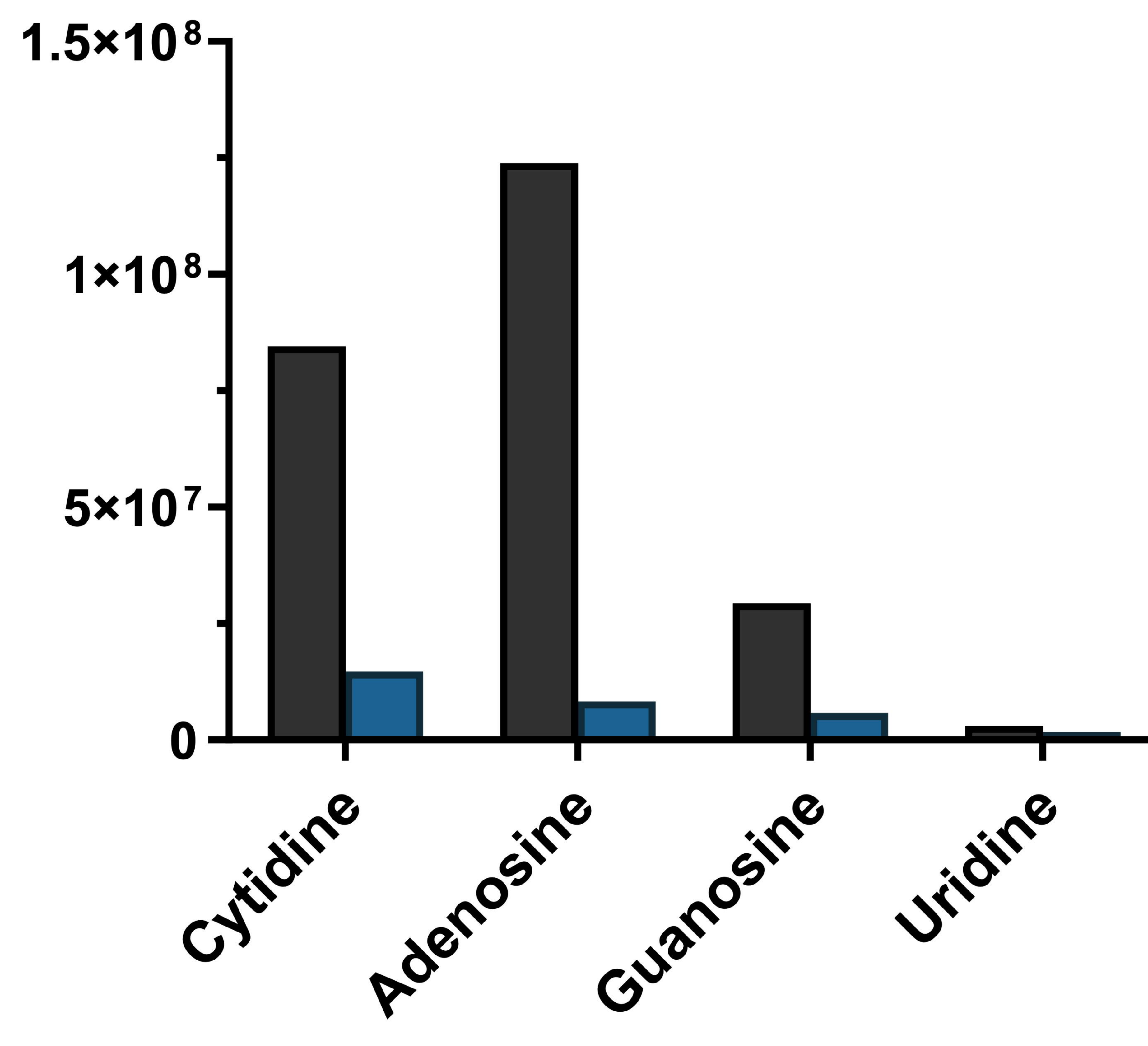
$m/z$ : 330.25

$m/z$ : 206.15



**(a)****Chromatogram of Permethylated Ribonucleosides**

— Cytidine      — Adenosine  
 — Guanosine    — Uridine

**(b)****LC Gradient Curve for Permethylated Ribonucleosides****(c)****Chromatogram of Native Ribonucleosides****(d)****LC Gradient Curve for Native Ribonucleosides****(e)****Comparison of Peak Area****Comparison of Peak Height**

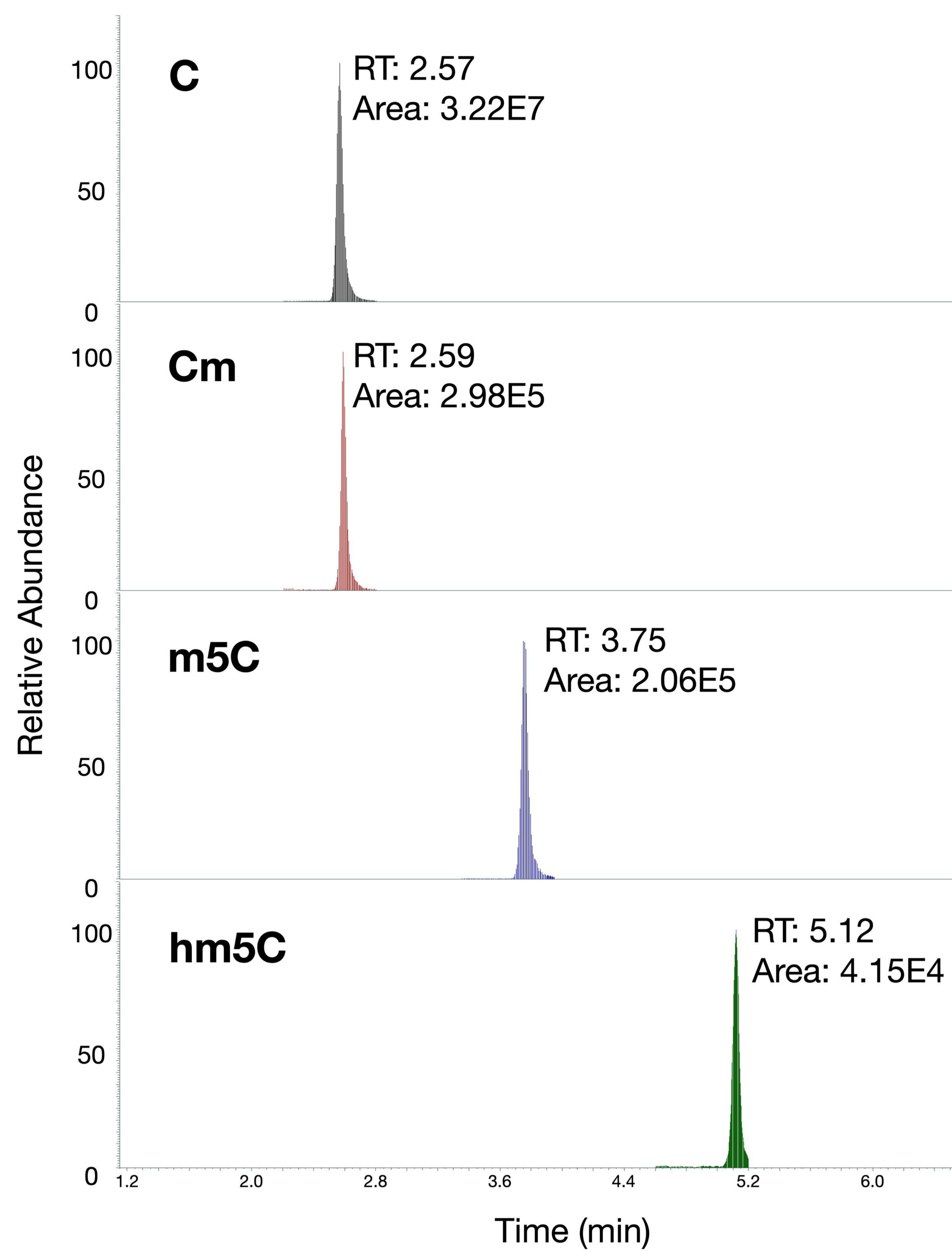
■ Permethylated Ribonucleoside

■ Native Ribonucleoside

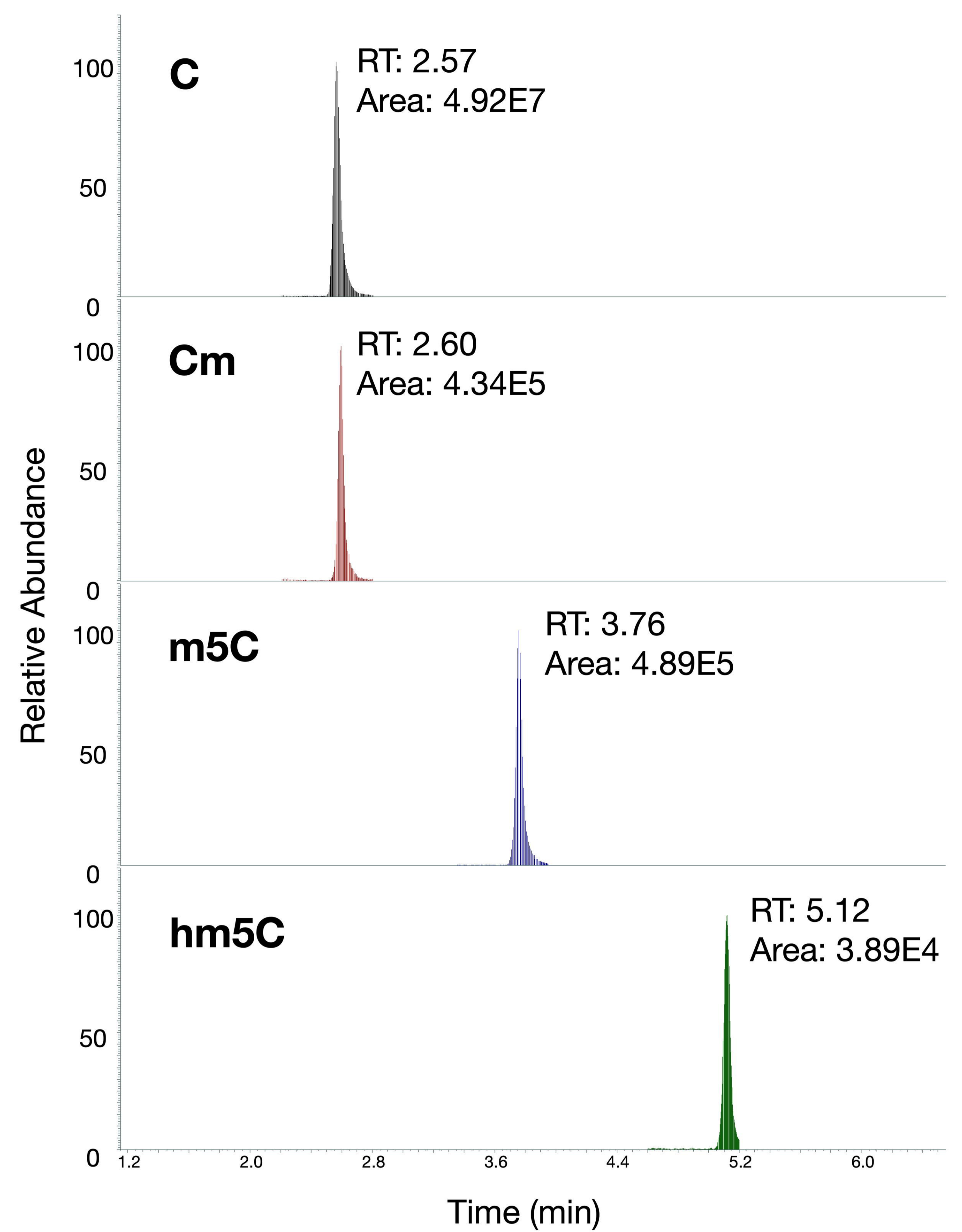


(a)

WT



(b)

*Tet2 KO*

(c)

