Label-free, direct localization and relative quantitation of the RNA nucleobase methylations m6A, m5C, m3U, and m5U by top-down mass spectrometry

Heidelinde Glasner, Christian Riml, Ronald Micura and Kathrin Breuker*

Abstract

Nucleobase methylations are ubiquitous posttranscriptional modifications of ribonucleic acids (RNA) that can substantially increase the structural diversity of RNA in a highly dynamic fashion with implications for gene expression and human disease. However, high throughput, deep sequencing does not generally provide information on posttranscriptional modifications (PTMs). A promising alternative approach for the characterization of PTMs, i.e. their identification, localization, and relative quantitation, is top-down mass spectrometry (MS). In this study, we have investigated how specific nucleobase methylations affect RNA ionization in electrospray ionization (ESI), and backbone cleavage in collisionally activated dissociation (CAD) and electron detachment dissociation (EDD). For this purpose, we have developed two new approaches for the characterization of RNA methylations in mixtures of either isomers of RNA or nonisomeric RNA forms. Fragment ions from dissociation experiments were analyzed to identify the modification type, to localize the modification sites, and to reveal the site-specific, relative extent of modification for each site.

Introduction

Decades of research have identified more than 140 chemically distinct types of posttranscriptional RNA modifications, of which roughly two thirds involve methylation (1–4). Methylated RNA nucleotides can be found in all kingdoms of life, and a large variety of biological processes rely on dynamic and reversible methylation of both coding and noncoding (nc) RNA (5,6). For example, recent studies have revealed vital roles of N6-methyladenosine (m6A) (7,8), 5-methylcytidine (m5C) (9,10) and N1-methyladenosine (m1A) (11,12) in posttranscriptional regulation of gene expression. RNA methylation is implicated in human disease (13–16) and can frequently be found in ncRNAs whose functions are still unknown (17–19). Research into ncRNAs progresses at a high rate, but is still critically hindered by lack of adequate methodology for the characterization of posttranscriptional modifications of RNA. In two recent review articles, Limbach (20) and Helm and Motorin (21) discuss the broad range of experimental methods available for the detection of RNA modifications, and point out mass spectrometry (MS) as ‘the only technology capable of directly identifying nearly all possible chemical modifications in RNA’ and ‘the apparently most straightforward approach’ to sequence-specific analysis of transcriptomes, respectively.

Although the development of MS methodology for the characterization of modified RNA has been slower than that of proteins (22–27), an increasing number of studies illustrates the high potential of both the ‘bottom-up’ (28) and ‘top-down’ (29) approaches that were originally developed for MS of proteins. For example, bottom-up MS, which relies on chemically or enzymatically catalyzed protein or RNA hydrolysis and separation of the products by liquid chromatography (LC) prior to MS analysis, has been used to identify and localize a novel form of N6-threonylcarbamoyladenosine (t6A), i.e. 2-methylthio cyclic N6-threonylcarbamoyladenosine (ms2ct6A), in transfer RNA (tRNA) (30), and, in combination with stable isotope labeling, for quantitative analysis of ribosomal RNA (rRNA) (31) and tRNA (32,33) modifications. LC–MS based methods for the identification and quantitation of modified nucleosides have been designed to probe tissue-specific tRNA modification levels (34), to simultaneously identify and quantify modified nucleosides from <200 nt RNA (35), to characterize ribosomal RNAs (36), and to explore the epitranscriptome of human neural stem cells with atomolme sensitivity (37). Recently, sophisticated, label-free MS methods were used without LC to profile RNA modifications at the full-transcriptome level (38), and label-free LC–MS analysis of messenger RNA capping efficiency has been reported (39). Bottom-up MS can also be used to detect and localize the mass-silent modification pseudouridine (Ψ).

*To whom correspondence should be addressed. Tel: +43 512 507 57740; Fax: +43 512 507 57799; Email: kathrin.breuker@uibk.ac.at

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
(40–42). Because top-down MS does not involve hydrolysis, it can be used to detect and characterize different forms of a protein or RNA whereby it can provide a complete description of the primary structure and reveal all mass-altering modifications, as well as any correlations that exist between these modifications (43). Top-down MS (43–45) has been used for the direct identification of human cellular microRNA (miRNA) (46), small interfering RNA (siRNA) (47), for the characterization of synthetic 21–23 nt RNA (48), tRNA (49) and RNA stem–loop motifs modified by structural probes (50), for de novo sequencing of tRNA (51), and for probing tat peptide binding to TAR RNA (52).

Here we explore the potential of top-down MS for the simultaneous identification, localization and relative quantitation of methylated RNA residues using electrospray ionization (ESI), low-energy collisionally activated dissociation (CAD), and electron detachment dissociation (EDD) (51,53,54). Our label-free, direct approach for this purpose builds on previous studies of modified bovine ribonuclease A (55), human histone H3 (56,57) and H4 (58), tropomyosin (59), and cardiac troponin I (60) proteins, and requires that RNA methylation does not affect RNA desorption, ionization and dissociation. Moreover, the approach relies on high sequence coverage, so RNA backbone cleavage should be largely nonselective with respect to nucleobase identity. Although EDD is somewhat affected by nucleobase ionization energy (53), it provided 80% sequence coverage for highly modified 76 nt tRNA in a single spectrum, and low-energy CAD of the same tRNA provided 89% sequence coverage (51).

The chemical structures of the nucleobase methylations investigated, m⁶A, m⁵C, m³U and m⁵U, are shown in Scheme 1. Because all previously proposed mechanisms for RNA backbone cleavage into e and y fragments (Scheme 2) by low-energy CAD involve the ribose 2′-OH group but not the nucleobase (61–64), nucleobase modifications should generally not affect the yield of e and y fragments. However, we have recently found increased backbone cleavage into e and y fragments on the 5′ side of adenosine and guanosine in CAD of RNA (M+nH)⁺⁺ and (M−nH)⁻⁻ ions, respectively, which we attributed to hydrogen bonding between

nucleobases and phosphates (65); nucleobase methylation may affect hydrogen bonding and thus backbone cleavage. Likewise, nucleobase methylation can affect nucleobase ionization energy and thus backbone cleavage into d and w fragments by EDD (53).

In the following, we refer to RNA isomers and forms as RNAs with the same basic sequence but different methylation patterns. As exemplified for a 10 nt sequence in Scheme 3A, RNA isomers (left) comprise the same number of methylated residues but at different positions, whereas RNA forms (right) differ by the number of methylated residues. RNA isomers, by definition, have the same elemental composition and therefore the exact same mass, are frequently inseparable by chromatographic means, and cannot be resolved by single-stage mass spectrometry. RNA forms differ in elemental composition and can be distinguished from each other by single-stage mass spectrometry as each methylation results in a 14.0157 Da mass increase, but isolation of individual forms in the mass spectrometer is only possible for RNAs of up to ~20 nt due to overlap of isotopic profiles (Scheme 3B).

**MATERIALS AND METHODS**

Experiments were performed on a 7 T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (Bruker, Austria) equipped with an ESI source for an internal calibrant for accurate mass measurements (Table 1). Methanol was HPLC grade (Acros, Austria), H2O was purified to 18 MΩ·cm at room temperature using a Milli-Q system (Millipore, Austria), ammonium acetate (≥99.0%, Na ≤ 5 mg/kg, K ≤ 5 mg/kg), ammonium citrate (2.5 M in H2O, Na ≤ 20 mg/kg, K ≤ 20 mg/kg), piperidine (≥99.5%), imidazole (≥99.5%, Na ≤ 50 mg/kg, K ≤ 50 mg/kg), and quinuclidine (97.0%, used as additive solely for accurate mass measurement of RNAs 14 and 15) were from Sigma-Aldrich (Austria). The (M−nH)⁻⁻ ions under

---

**Scheme 1.** Chemical structure of methylated nucleoside residues studied: N⁶-methyladenosine (m⁶A), 5-methylcytidine (m⁵C), 3-methyluridine (m³U), and 5-methyluridine (m⁵U).

**Scheme 2.** Nomenclature of fragment ions from RNA backbone cleavage according to reference (66).

---

**Scheme 3.** A: RNA isomers (left) comprise the same number of methylated residues but at different positions, whereas RNA forms (right) differ by the number of methylated residues. RNA isomers, by definition, have the same elemental composition and therefore the exact same mass, are frequently inseparable by chromatographic means, and cannot be resolved by single-stage mass spectrometry. RNA forms differ in elemental composition and can be distinguished from each other by single-stage mass spectrometry as each methylation results in a 14.0157 Da mass increase, but isolation of individual forms in the mass spectrometer is only possible for RNAs of up to ~20 nt due to overlap of isotopic profiles (Scheme 3B).
study were isolated in a linear quadrupole prior to dissociation by CAD or EDD; for a more detailed description of the experimental setup for CAD and EDD, see (67) and (53), respectively. For statistical reasons, between 25 and 500 scans were added for each spectrum (50–100 for ESI, 25–50 for (M–nH)− ion isolation, 200–500 for CAD and EDD), and data reduction utilized the SNAP2 algorithm (Bruker, Austria). RNAs 1, 7, 11 and 13–15 (Table 1) were prepared by solid-phase synthesis (68,69) and purified by HPLC. All other RNAs (Table 1) were custom synthesized by commercial suppliers: 2 by BioSpring GmbH (Frankfurt am Main, Germany), 3, 8–10 and 12 by Microsynth AG (Balgrach, Switzerland), 4–6 and 16–20 by Dharmacon GE Healthcare (Lafayette, CO, USA), and purified by HPLC. For desalting, 400 μl of an ammonium salt solution (100 mM ammonium acetate in H2O for RNAs 1–7, 11–13 and 16–20, 100 mM ammonium citrate in 1:1 H2O/CH3OH for RNAs 8–10 and 14–15) was added to 100 μl RNA solution (2.5–10 nmol in H2O) and concentrated to 50 μl using Vivaspin 500 centrifugal concentrators (Sartorius, Germany), PES membrane, MWCO 3000 for RNAs 1–7, 11–13 and 16–20, and MWCO 5000 for RNAs 8–10 and 14–15. The process was repeated five to seven times, followed by six to seven cycles of concentration and dilution with H2O (RNAs 4–7, 11–13 and 16–20) or 1:1 H2O/CH3OH (RNAs 1–3, 8–10 and 14–15). To minimize possible effects of RNA higher order structure on electrospray ionization and UV absorption, the samples in H2O were diluted with CH3OH to 1:1 H2O/CH3OH (70–72). RNA concentration was subsequently determined by UV absorption at 260 nm using a NanoPhotometer (Implen, Germany). In first experiments in which we studied mixtures of isomers and forms comprising m6A (23 nt RNAs 4–6, 16 and 17) and m5U (52 nt RNAs 14 and 15), nucleotide absorption coefficients from reference (73) were used because comparable values for unmethylated and methylated bases were not available. However, nucleotide absorption coefficients of A and m6A published later differed by only 3.2% (37), which amounts to negligible differences of <0.5% for RNAs 4–6; the stoichiometry of RNA isomers 4, 16 and 17 should not be affected. Values for m5U are still not available, but a single methylation (U: 8912.32 l mol−1 cm−1, m5U: 7743.31 l mol−1 cm−1, m5U: 9651.15 l mol−1 cm−1 (37)) should not significantly affect the absorption coefficient of the 52 nt RNAs 14 and 15. For all other experiments, absorption coefficients for m5C (43% smaller than that of C) were used (37). Mixtures of RNA forms or isomers in H2O/CH3OH were prepared immediately before each measurement. The amount of RNA sample consumed in the top-down MS experiments here varied between 0.0375 and 0.3 nM, but sensitivity on more modern instruments (nano-ESI source, rf ion transfer, stronger magnets, etc.) can be estimated to be higher by a factor of at least 200, corresponding to 0.1875–1.5 pM.

RESULTS AND DISCUSSION

ESI MS of methylated RNA

The first question that we have addressed is whether or not nucleobase methylation affects the yield and net charge of (M–nH)n− ions from ESI. For this purpose, mixtures of different forms of methylated RNA with the same basic sequence were electrosprayed from denaturing solutions (1:1 H2O/CH3OH) (70–72) using different ESI additives that affect the average net charge of the (M–nH)n− ions (74). Although all RNAs studied here were purified by HPLC, we first recorded separate ESI MS spectra of each RNA sample to verify their purity. For RNAs 4, 5 and 6, which were simultaneously purchased from the same supplier, the ESI MS spectra indicated similar purities of 85, 80 and 81%, respectively; the various ions of generally low abundance that did not correspond to RNAs 4, 5 and 6 presumably originated from coelution of by-products formed during RNA synthesis or partial hydrolysis. As illustrated in Figure 1 for a 1:1:1 mixture of the samples of RNAs 4, 5 and 6 with one, two, and three m6A residues, respectively, the yield of (M–nH)n− ions of the three forms (normalized to the total abundance of all (M–nH)n− ions of RNAs 4, 5 and 6) was the same within ±2% for each net charge n. Moreover, no significant correlation between total ion yield or average net charge and the extent of methylation was found, regardless of the ESI additive used (Table 2), suggesting that the small deviations from nonstoichiometric (M–nH)n− ion yields are unsystematic errors and not caused by nucleobase methylation. Yields corrected for the above purity values were similarly unsystematic and agreed with uncorrected values to within ±1.2% (Table 2).

However, for other RNA forms, our MS data revealed substantial differences in purity of up to ∼20%. For example, ESI of a supposedly equimolar mixture of 27 nt RNAs 7, 8, 11, 12 and 13 with up to four methylated m5C residues showed differences in (M–nH)n− ion yield (normalized to
Table 1. RNA studied

<table>
<thead>
<tr>
<th>RNA</th>
<th>Sequencea</th>
<th>nt</th>
<th>(M_{\text{measured}}^b)</th>
<th>(M_{\text{calculated}}^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GAAGG UUUC CUUCG</td>
<td>15</td>
<td>4729.619</td>
<td>4729.621</td>
</tr>
<tr>
<td>2</td>
<td>GAAGG m5Um5U UUUC CUUCG</td>
<td>15</td>
<td>4757.651</td>
<td>4757.652</td>
</tr>
<tr>
<td>3</td>
<td>GAAGG m5Um5Um5U UUUC CUUCG</td>
<td>15</td>
<td>4785.679</td>
<td>4785.684</td>
</tr>
<tr>
<td>4</td>
<td>pAUUAU AAm6AACC AAACA AAAAA UAA</td>
<td>23</td>
<td>7439.107</td>
<td>7439.101</td>
</tr>
<tr>
<td>5</td>
<td>pAUUAU m6A AAcC Am6AACA AAAAA UAA</td>
<td>23</td>
<td>7453.125</td>
<td>7453.117</td>
</tr>
<tr>
<td>6</td>
<td>pAUUAU m6A Am6Am6ACC AAACA AAAAA UAA</td>
<td>23</td>
<td>7467.122</td>
<td>7467.133</td>
</tr>
<tr>
<td>7</td>
<td>AUCUC CUUCG CCAUC GGGGC CGCGG AU</td>
<td>27</td>
<td>8615.163</td>
<td>8615.162</td>
</tr>
<tr>
<td>8</td>
<td>AUCUC CUUCGm6C CCAUC GGGGC CGCGG AU</td>
<td>27</td>
<td>8629.177</td>
<td>8629.177</td>
</tr>
<tr>
<td>9</td>
<td>AUCUC CUUCG CCAUC GGGGCm6C CGCGG AU</td>
<td>27</td>
<td>8629.175</td>
<td>8629.177</td>
</tr>
<tr>
<td>10</td>
<td>AUCUC CUUCG CCAUC GGGGC m6C CGCGG AU</td>
<td>27</td>
<td>8629.176</td>
<td>8629.177</td>
</tr>
<tr>
<td>11</td>
<td>AUCUC CUUCGm6C Cm6CAUC GGGGC CGCGG AU</td>
<td>27</td>
<td>8643.192</td>
<td>8643.193</td>
</tr>
<tr>
<td>12</td>
<td>AUCUC CUUCGm6C Cm6Cm6CAUC GGGGC CGCGG AU</td>
<td>27</td>
<td>8671.201</td>
<td>8671.209</td>
</tr>
<tr>
<td>13</td>
<td>AUCUC CUUCGm6C m6C Cm6CAUC GGGGm6C CGCGG AU</td>
<td>27</td>
<td>8671.219</td>
<td>8671.224</td>
</tr>
<tr>
<td>14</td>
<td>UCAGC CUAGU UUAAC CGUAU UGCAA GCGCG UGAUA AAGU AGC UA AAAAG GC</td>
<td>52</td>
<td>16764.267</td>
<td>16764.267</td>
</tr>
<tr>
<td>15</td>
<td>UCAGC CUAGU UUAAC CGUAU UGCAA GCGCG UGAUA m6C UGAUA AAG CG</td>
<td>52</td>
<td>16778.285</td>
<td>16778.283</td>
</tr>
<tr>
<td>16</td>
<td>pAUUAU m6A AACC AAACA AAAAA UAA</td>
<td>23</td>
<td>7439.105</td>
<td>7439.101</td>
</tr>
<tr>
<td>17</td>
<td>pAUUAU Am6AACC AAACA AAAAA UAA</td>
<td>23</td>
<td>7439.104</td>
<td>7439.101</td>
</tr>
<tr>
<td>18</td>
<td>pAUUAU m6C CCCCCC AAACA AAAAA UAA</td>
<td>23</td>
<td>7367.075</td>
<td>7367.068</td>
</tr>
<tr>
<td>19</td>
<td>pAUUAU m6Cm6C CCCCCC AAACA AAAAA UAA</td>
<td>23</td>
<td>7367.066</td>
<td>7367.068</td>
</tr>
<tr>
<td>20</td>
<td>pAUUAU Cm6Cm6C CCCCCC AAACA AAAAA UAA</td>
<td>23</td>
<td>7367.066</td>
<td>7367.068</td>
</tr>
</tbody>
</table>

aFrom 5'-3' terminus, OH-terminated unless indicated (p: phosphate)
bIn Da; \(M_n\) refers to the mass of the most abundant isotope.

Table 2. Average charge and yield of (M–nH)+ ions from ESI of RNAs 4, 5 and 6

<table>
<thead>
<tr>
<th>ESI additive</th>
<th>RNA</th>
<th># of m6A residues</th>
<th>Average charge</th>
<th>Yield [%]</th>
<th>Corrected yield [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piperidine</td>
<td>4</td>
<td>1</td>
<td>11.99</td>
<td>33.5</td>
<td>32.4</td>
</tr>
<tr>
<td>(25 mM, pH ~10)</td>
<td>5</td>
<td>2</td>
<td>12.04</td>
<td>33.2</td>
<td>34.1</td>
</tr>
<tr>
<td>(25 mM, pH ~10)</td>
<td>6</td>
<td>3</td>
<td>12.04</td>
<td>33.2</td>
<td>33.5</td>
</tr>
<tr>
<td>Piperidine/imidazole</td>
<td>4</td>
<td>1</td>
<td>7.44</td>
<td>32.0</td>
<td>30.9</td>
</tr>
<tr>
<td>(25 mM each, pH ~10)</td>
<td>5</td>
<td>2</td>
<td>7.46</td>
<td>35.9</td>
<td>36.8</td>
</tr>
<tr>
<td>(25 mM each, pH ~10)</td>
<td>6</td>
<td>3</td>
<td>7.50</td>
<td>32.1</td>
<td>32.3</td>
</tr>
<tr>
<td>Ammonium acetate</td>
<td>4</td>
<td>1</td>
<td>4.99</td>
<td>34.5</td>
<td>33.3</td>
</tr>
<tr>
<td>(20 mM, pH ~6)</td>
<td>5</td>
<td>2</td>
<td>4.99</td>
<td>33.2</td>
<td>34.1</td>
</tr>
<tr>
<td>(20 mM, pH ~6)</td>
<td>6</td>
<td>3</td>
<td>4.99</td>
<td>32.3</td>
<td>32.6</td>
</tr>
</tbody>
</table>

Equations for calculation of average charge and yield can be found in the supporting information.

the total abundance of all (M–nH)+ ions of RNAs 7, 8, 11, 12 and 13 of up to 5% (Figure 2), but again, no significant correlation between ion yield or average net charge (6.47, 6.47, 6.47, 6.48, 6.49, respectively) and the extent of methylation was found. Instead, separate ESI MS spectra of each RNA indicated purities between ~63% and ~81% that correlated linearly (\(R^2 = 0.9985\)) with (M–nH)+ ions from ESI of the supposedly equimolar mixture. Corrected for purity, the (M–nH)+ ion yields were 20.3, 19.0, 20.4, 20.5 and 19.8% for RNAs 7, 8, 11, 12 and 13 (before correction: 21.8, 19.5, 21.7, 17.3 and 19.7%, respectively), which agrees to within 1% with a 1:1:1 stoichiometry in solution. Correspondingly large differences in (M–nH)+ ion yield were observed for 1:1 mixtures (RNAs 7 and 11), ~4% with 30 mM piperidine and 30 mM imidazole or 20 mM ammonium acetate as additive, but only 1.5% with 30 mM piperidine as additive (Supplementary Table S1).

To investigate any possible effects of ESI additive on the yield of (M–nH)+ ions of RNA forms with differing extent of methylation more systematically, we studied 1:1:1 mixtures of the 15 nt RNAs 1, 2 and 3 with 0, 2 and 4 m5U residues, respectively. A statistical analysis of ~50 spectra from ESI of separately prepared solutions of 1:1:1 mixtures of RNAs 1, 2 and 3 revealed that RNA concentration rather than the additive used can cause differences in (M–nH)+ ion yield of different RNA forms. More specifically, the deviation of (M–nH)+ ion yields of RNAs 1, 2 and 3 from those expected for the 1:1:1 stoichiometry in solution and in the absence of any effects of methylation on ESI (i.e. 33.3%:33.3%:33.3%) consistently increased with increasing RNA concentration, but without any correlation between (M–nH)+ ions and the extent of methylation (Figure 3). Moreover, no correlation was found between average charge (piperidine: 4.03, 4.03 and 4.04 for 0, 2 and 4 m5U residues, respectively; imidazole: 4.73, 4.73 and 4.73 for 0, 2 and 4 m5U residues, respectively; ammonium acetate: 6.85, 6.81 and 6.78 for 0, 2 and 4 m5U residues, respectively) and the extent of methylation. This suggests that RNA forms 1, 2 and 3 have different propensity for dimer (or multimer) formation in solution, and/or for any concentration-dependent behavior in the ESI droplets, that in turn affect ion formation by ESI, but again without any correlation between (M–nH)+ ion yields and the extent of methylation. In further support of this hypothesis, the stoichiometry of RNA forms in solution was generally more accurately reflected (e.g. at a total RNA concentration of 0.75 μM, 32.1, 34.2 and 33.7% for RNAs 1, 2 and 3, respectively) in spectra from ESI of RNA solutions with a ~10-fold molar excess of polyethylene glycol 1000 that was found to destabilize RNA structure (75).
Figure 1. Mass spectra from ESI of 1:1:1 mixtures of 23 nt RNAs 4, 5 and 6 (0.5 or 1 μM each) in 1:1 H2O/CH3OH using (A) 25 mM piperidine, (B) 25 mM piperidine and 25 mM imidazole and (C) 20 mM CH3COONH4 as additives, insets show expanded regions with isotopic profiles of the most abundant (M−nH)− ions and (D) corresponding yield of (M−nH)− ions versus net charge. Color coding refers to the number of m6A residues: red for 1 (RNA 4), violet for 2 (RNA 5) and blue for 3 (RNA 6).

The first major conclusion from the ESI studies is that m6A (Figure 1), m5C (Figure 2) and m3U (Figure 3) have no systematic effect on RNA (M−nH)− ion yields, but nonsystematic, RNA concentration-dependent effects of methylation on (M−nH)− ion yields were observed for forms 1, 2 and 3. Further studies are needed to clarify whether these originate from RNA aggregation in solution and/or from RNA (M−nH)− ion formation by ESI, the mechanism of which is only poorly understood (74). Even so, any effects of methylation were relatively small, on the order of ±3% for 1:1:1 mixtures.

Figure 2. Yield of (M−nH)− ions from ESI of a 1:1:1:1:1 mixture of 27 nt RNAs 7, 8, 11, 12 and 13 (0.5 μM each in 1:1 H2O/CH3OH with 30 mM piperidine and 30 mM imidazole as additives) versus charge, left inset illustrates region of the corresponding mass spectrum with isotopic profiles of (M−6H)− ions, right inset shows a linear correlation of yield with sample purity. Color coding refers to the number of m5C residues: gray for 0 (RNA 7), red for 1 (RNA 8), violet for 2 (RNA 11), blue for 3 (RNA 12), and green for 4 (RNA 13).

Figure 3. Yield of (M−nH)− ions from ESI of a 1:1:1 mixture of 15 nt RNAs 1, 2 and 3 in 1:1 H2O/CH3OH versus total RNA concentration with 20 mM piperidine (circles), 20 mM piperidine and 20 mM imidazole (diamonds), or 20 mM ammonium acetate (triangles) as additive. Open symbols represent data from individual spectra and filled symbols show average values with standard deviations as error bars. Color coding refers to the number of m5U residues: gray for 0 (RNA 1), violet for 2 (RNA 2) and green for 4 (RNA 3).
Nevertheless, single-stage mass spectrometry as shown in Figures 1 and 2 does not provide any information about the sites or the site-specific extent of methylation. For this, a second stage of mass spectrometry is required in which the ions of interest are isolated and dissociated. As illustrated here in Figure 1 and shown in previous studies (51,74), different ESI additives can be used to produce \((M-nH)^{-}\) ions with either low, intermediate or high net charge. CAD of RNA produces primarily \(c\) and \(y\) ions from phosphodiester backbone bond cleavage (Scheme 2) when \((M-nH)^{-}\) ion net charge is low, ∼0.2 charges/nt, whereas EDD into \(d\) and \(w\) ions is most efficient when the net negative charge is high, ∼0.5 charges/nt (51,53,67). The fragment ions from CAD and EDD were previously shown to provide complementary information for de novo sequencing of highly modified RNA (51,76). In the following, we will first discuss the use of CAD for the localization and relative quantitation of RNA methylations.

**CAD of \((M-nH)^{-}\) ions of methylated RNA**

For the characterization of RNA isomers, we studied equimolar mixtures of RNAs 8–10, 18–20 and 4, 16 and 17. Figure 4A shows mass spectra from ESI, isolation, and CAD of the \((M-6H)^{-}\) ions of the isomeric 27 nt RNAs 8–10 (mC) whose sequence corresponds to the spanning repeat 8 of XIST RNA that was recently identified as an in vivo target for site-specific 5-methylation of cytosines 10, 20 and 21 (19). Consistent with previous studies (51,53,67,77), CAD of the \((M-6H)^{-}\) ions (0.22 charges/nt) produced predominantly \(c\) and \(y\) fragments from phosphodiester backbone cleavage (>92% out of all fragments from backbone cleavage in the CAD spectrum in Figure 4A, see Supplementary Table S2) that were analyzed to reveal the sites and site-specific extent of methylation. For each cleavage site, the fraction of methylated fragments was calculated from relative abundances of \(c\) and \(y\) fragments with and without methylation, respectively (Figure 4B and C). Because dissociation into \(a\) and \(w\) fragments (Scheme 2) was only a minor channel (<8%), the corresponding signals were generally too small or even missing, and could therefore not be used for this purpose (Supplementary Table S2). Only unmethylated \(c\) and methylated \(y\) fragments were detected for sites 1–9, indicating that residues 1–9 are unmethylated. The stepwise increase in the fraction of methylated fragments at cleavage sites 10, 20 and 21 by 33.9, 33.4 and 32.7%, respectively, is consistent with methylation of residues 10, 20 and 21, and quantitatively agrees to within <1% with the 1:1:1 stoichiometry of RNAs 8–10 in the mixture. Apparently, 5-methylation of cytosine does not appreciably affect phosphodiester backbone cleavage into \(c\) and \(y\) fragments, regardless of the energy used for CAD (Figure 4C) and the position of mC within the sequence; similar data were obtained for isomeric mixtures of RNAs 18–20 (mC, Supplementary Table S3) and 4, 16 and 17 (mA, Supplementary Table S4).

The characterization of mixtures of different forms of methylated RNA by top-down MS is conceptually different from that of mixtures of different isomers. First, ESI MS immediately reveals the presence of different forms (Figures 1 and 2) but not isomers (Figure 4A) of methylated RNA.
Second, because the total number of methylations is the same for isomers (Scheme 3A), the fraction of methylated c and methylated y fragments from dissociation of mixtures of RNA isomers add up to 100% at each cleavage site such that data points fall on top of each other when fractions of methylated c and methylated y fragments are plotted with reversed ordinate axes (Figure 4C; in an alternative but equal representation, data points fall on top of each other when fractions of methylated c and unmethylated y fragments are plotted on the same ordinate axis), but this is not the case for mixtures of RNA forms. As illustrated in Figure 5 for CAD of a simple 1:1 mixture of the two 52 nt RNA forms 14 and 15 that differ only by methylation at site 31 (m^3U), the fraction of methylated y fragments increases from 0% at cleavage sites 31–50 (y_1 was not detected) to an average value of 56.6 ± 2.7% at sites 7, 16, 17, 20–23, 25 and 27–30; the 6.6% deviation from the theoretical value, 50%, can be attributed to differences in sample purity. The corresponding fractions of methylated c fragments were 0% at cleavage sites up to 30, and 51.3, 52.9 and 51.6% at sites 31, 35 and 38, respectively, and do not fall on top of the fractions of methylated y fragments when plotted with reversed ordinate axes (Figure 5). Values for most of the fragments larger than ~35 nt could not be determined because larger fragments have a higher probability to undergo secondary dissociation (53,67), nevertheless, the data from a single CAD spectrum revealed both the site of methylation and the stoichiometry of the two 52 nt RNA forms in the mixture.

Visualization of methylation sites and site-specific fractions of methylated c and y fragments such as in Figures 4C and 5 is less straightforward for mixtures of RNA isomers with more than one methylation and RNA forms that differ by more than one methylation because the fraction of methylated fragments is no longer directly correlated with the number of methylations of a fragment. For example, CAD of a 1:1 mixture of RNAs 1 (unmethylated) and 2 (m^3U at both sites 6 and 7) produced only unmethylated y_8–y_14 fragments (y_1 was not detected), ~50% each of unmethylated and singly methylated y_9, and ~50% each of unmethylated and doubly methylated y_10–y_14 fragments (Figure 6A, see the above discussion on ESI for deviations from the theoretical values of 50%). A plot of the fraction of methylated c fragments shows a transition at cleavage site 6, whereas that of methylated y fragments shows a transition at site 7 (Figure 6B), consistent with m^5U at sites 6 and 7.

More generally, Scheme 4 illustrates theoretical fractions of methylated c and y fragments from CAD of hypothetical 1:1:1 mixtures of RNA isomers with more than one methylation and RNA forms that differ by more than one methylation. The fraction of methylated fragments from CAD of the hypothetical mixture of isomers in Scheme 4A shows two increases by 33.3%, similar to that in Figure 4C, even though the number of methylations in the hypothetical mixture is twice as high as that in Figure 4C. The difference between these two isomer mixtures is the number of methylations of each fragment that is the same for c or y fragments from a given cleavage site in Figure 4C but varies for the hy-
Scheme 4. Theoretical fractions of methylated c (circles) and y (triangles) fragments from CAD of hypothetical 1:1:1 mixtures of RNA (A) isomers with two methylations and (B) forms with 0, 1 and 2 methylations; the number of methylations of fragments in each fraction is illustrated as color coded bars (red for 1 and violet for 2 methylations).

Figure 7. Fractions of (A) c and (B) y fragments from CAD (average from three measurements at 75, 77.5 and 80 eV laboratory frame energy, which showed no significant effect on relative fragment abundances) of (M−5H)− ions from ESI of a 1:1:1 mixture of the 23 nt RNA forms 4, 5 and 6 (1 μM each in 1:1 H2O/CH3OH with 20 mM CH3COONH4 as additive, Figure 1C) with 1, 2 and 3 m5A residues, respectively, versus cleavage site; color coding indicates the number of methylations of a fragment: gray = 0, red = 1, violet = 2, blue = 3.

Apparent, for more complex mixtures of RNA isomers and forms, simply plotting the fraction of methylated fragments (Figures 4C, 5 and 6B) does not adequately illustrate the complexity of methylation patterns. Although Scheme 4 accounts for the fraction of methylated fragments, and its color coding for the number of methylations of the fragments, the methylation patterns of individual isomers or forms are still not obvious. A more comprehensive way of representing the MS data from dissociation of different isomers or forms of methylated RNA that immediately reveals individual methylation patterns are plots with colored blocks that represent fractions of unmethylated and methylated fragments, as illustrated in Figure 7 for a 1:1:1 mixture of the 23 nt RNA forms 4-6.

In Figure 7, the colored blocks were arranged such that the following criteria were fulfilled: (i) in each horizontal series, the number of methylations can only increase but not decrease with increasing size of the fragments (i.e. from left to right for c fragments, and from right to left for y fragments); (ii) the number of color transitions in each horizontal series must correspond to the total number of methylations of a form (for multiply methylated residues, this criterion is not applicable); (iii) patterns for c and y fragments must be consistent with each other regarding the sites and numbers of methylation; (iv) the total number of methylations of each form, indicated by the color on the 3′ terminus (right) for c ions and the 5′ terminus (left) for y ions in each horizontal series, must account for the total number of methylations in the ESI spectrum (Figure 1). The same criteria can be used for mixtures of isomers with adaption of (iv) to a fixed total number of methylations.

To test our approach for an even more complex mixture, we studied the 27 nt RNAs 7, 8, 11, 12 and 13 forms with 0 to 4 m5C residues by CAD (Figure 2). In treating the methylation pattern as unknown, site-specific fractions of c and y fragments were color coded according to the number of methylations and arranged by taking into account...
the above criteria. For the unmethylated form, gray blocks corresponding to unmethylated c and y fragments were horizontally arranged in the top rows of Figures 8 A and B, respectively. Because this left no gray blocks representing c fragments from cleavage at sites >10 and y fragments from cleavage at sites <10, the form with one methylation could be identified. Further, only one arrangement was possible for the colored blocks representing c fragments of the form with four methylations (color transitions from gray to red to violet to blue to green); blocks for y fragments were arranged accordingly. This left two possibilities for arranging the blocks representing c fragments of the forms with two and three methylations, but only one of them was consistent with arranging the blocks for y fragments according to the above criteria (Figure 8). Within error limits, fractions of the five different forms of RNAs 7, 8, 11, 12 and 13 calculated from the CAD data in Figure 8 showed good agreement with uncorrected (M−nH)w− ions from ESI (Figure 2) and (M−6H)w− ion yields after isolation (Table 3). Notably, our top-down MS approach revealed the correct methylation patterns of all five RNA forms in a mixture, which would not have been possible by bottom-up MS as correlations between modifications are generally lost after digestion (43).

To minimize nucleobase losses and the formation of internal fragments from secondary dissociation (53,67), both of which are detrimental to the identification, localization and relative quantitation of RNA modifications by top-down MS, and to suppress undesired RNA backbone cleavage into a and w fragments (a fragments frequently show extensive nucleobase loss) (67,77), low-energy CAD of (M−nH)w− ions with low (0.22–0.27) net charge per nucleotide was used in the experiments reported so far. For 1:1:1 mixtures of the 15 nt RNA forms 1, 2 and 3, we evaluated the range of (M−nH)w− ion charge density (for n = 4–7, corresponding to 0.27–0.47 charges per nt) suitable for relative quantitation by c and y fragments from CAD. As illustrated in Supplementary Figure S1, the most accurate quantitation data and the highest sequence coverage from c and y fragments were obtained for n = 4 and 5, corresponding to 0.27 and 0.33 charges/nt. Higher charge densities led to significantly decreased signals of c and y fragments as a result of competitive backbone cleavage into a and w fragments, and correspondingly larger errors in the fractions of RNAs 1, 2 and 3 that were calculated from c and y fragment yields (Supplementary Figure S1).

Importantly, all CAD data from this study show that RNA phosphodiester backbone cleavage of (M−nH)w− ions into c and y fragments is not significantly affected by the presence of m6A, m5C, m3U and m5U residues unless (M−nH)w− ion charge is high (≥0.4 charges/nt). By contrast, preliminary data that will be published in a separate article suggest that methylation of the exocyclic amino group of guanosine substantially affects the yield of fragments from phosphodiester backbone cleavage. Low-energy CAD produces highly reproducible spectra provided that the energy available for dissociation is kept the same. However, as we have discussed in reference (67), the total energy available for dissociation is often unknown. We have suggested to instead consider the extent of molecular ion dissociation as a possible way of standardizing spectra obtained on different instruments (67). As shown here and in previous studies (51–53,65,67), undesired base loss, secondary fragmentation, and dissociation into a and w instead of c and y fragments can be minimized by use of energies that produce <40% dissociation of RNA (M−nH)w− ions with low net charge. Thus any mass spectrometer equipped for low-energy CAD experiments can in principle be used for label-free, direct localization and relative quantitation of the RNA nucleobase methylations m6A, m5C, m3U and m5U by top-down MS.

**EDD of (M−nH)w− ions of methylated RNA**

Finally, we have explored the potential of EDD for the relative quantitation of RNA methylations. A comparison of Figure 9, which shows fractions of d and w fragments (Scheme 2) from EDD of (M−12H)2− ions from a 1:1:1 mixture of the 23 nt RNA forms 4, 5 and 6 with 1, 2 and 3 m6A residues, respectively, with the CAD data of the same RNA mixture in Figure 7 immediately reveals that EDD provides less accurate data for the relative quantitation of RNA methylations than CAD. Apparently, methylation of the exocyclic amino group of adenosine strongly affects RNA dissociation into d and w fragments as their yields differ substantially, by up to 24% (Supplementary Figure S2), from theoretical values (33.3 or 66.7%) at cleavage sites next to the m6A residues at positions 6, 7, 8 and 12. Similar effects were observed with EDD of (M−7H)w− and (M−8H)w− ions from a 1:1:1 mixture of the 15 nt RNA forms 1, 2 and 3 with 0, 2 and 4 m5U residues, respectively (Supplementary Figure S3). Based on our recently proposed mechanism of RNA backbone cleavage into d and w frag-
bility of the intermediates involved in the dissociation reaction. Second, nucleobase methylation can affect the stabilization of nucleobase ionization energy, which is critical to the first step of methylated residues. First, methylation can alter the stoichiometric yields of fragments from cleavage next to phosphodiester backbone cleavage into $\text{[M-6H]}^-\text{ions}$ with 1, 2 and 3 m6A residues, respectively, versus cleavage site; color coding indicates the number of methylations of a fragment: gray = 0, red = 1, purple = 2, blue = 3.

Table 3. Yield of 27 nt RNA forms from ESI, isolation and CAD of (M–6H)− ions

<table>
<thead>
<tr>
<th>RNA</th>
<th># of m^3C residues</th>
<th>(M–nH)^±/ESI [%]</th>
<th>(M–6H)^−/isolation [%]</th>
<th>[%] e</th>
<th>[%] y</th>
<th>[%] e and y</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>0</td>
<td>21.8</td>
<td>22.6</td>
<td>23.9 ± 2.0</td>
<td>24.8 ± 1.7</td>
<td>24.2 ± 2.0</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>19.5</td>
<td>19.5</td>
<td>19.1 ± 1.0</td>
<td>18.5 ± 1.8</td>
<td>18.8 ± 1.5</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>21.7</td>
<td>21.8</td>
<td>22.3 ± 1.2</td>
<td>22.3 ± 1.9</td>
<td>22.3 ± 1.5</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>17.3</td>
<td>17.2</td>
<td>17.8 ± 1.2</td>
<td>15.8 ± 1.5</td>
<td>16.2 ± 1.7</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>19.7</td>
<td>18.9</td>
<td>18.9 ± 2.2</td>
<td>18.7 ± 2.0</td>
<td>18.8 ± 2.1</td>
</tr>
</tbody>
</table>

Figure 9. Fractions of (A) $d$ and (B) $w$ fragments from EDD (26 eV electron energy) of (M–12H)32− ions from ESI (0.5 μM each in 1:1 H2O/CH3OH with 25 mM piperidine as additive, Figure 1A) of a 1:1:1 mixture of the 23 nt RNA forms 4, 5 and 6 with 1, 2 and 3 m6A residues, respectively, versus cleavage site; color coding indicates the number of methylations of a fragment: gray = 0, red = 1, purple = 2, blue = 3.

both CAD and EDD provide extensive sequence information and can be used for the identification and localization of methylated residues, and can be combined for de novo sequencing of methylated RNA.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The authors thank Eva-Maria Schneeberger, Jovana Vusurovic and Matthias Halper for discussion, BioSpring GmbH (Frankfurt am Main, Germany) for providing RNA 2, Matthias Erlacher for providing RNAs 4–6 and 16–20 (from Dharmacon GE Healthcare, USA) and Veronika Schwarz for synthesis of RNAs 14 and 15.

FUNDING
Austrian Science Fund (FWF) [P27347 to K.B., P27947 to R.M.]; FFG [8005714]; University of Innsbruck ['Doktoratstipendium NEU aus der Nachwuchsfröderung der LFU' to H.G.]. Funding for open access charge: FWF. Conflict of interest statement. None declared.

REFERENCES


