Effect of the terminal phosphate group on negative-ion CID/HCD fragmentation behaviors of RNA oligonucleotides

Mei-Qing Zuo[1, 2], Ge Song[1, 2], Meng-Qiu Dong[1, 2, *], Rui-Xiang Sun[1, 2, *]

[1] National Institute of Biological Sciences, Beijing 102206, China
[2] Tsinghua Institute of Multidisciplinary Biomedical Research,
Tsinghua University, Beijing 100084, China

***Correspondence**
Meng-Qiu Dong: dongmengqiu@nibs.ac.cn
Rui-Xiang Sun: sunruixiang@nibs.ac.cn

ABSTRACT

In recent years, research and application of mass spectrometry (MS) on RNA molecules have begun to flourish. However, MS/MS fragmentation behaviors of RNA oligonucleotides (oligos) are far from being fully understood. In this study, we have investigated the effect of the terminal phosphate group on the fragmentation behaviors of RNA oligos using high-resolution mass spectrometry. Specifically, we synthesized eight RNA oligos containing a terminal phosphate group on either or both ends, or neither. Negative-ion mode collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD) of those oligos revealed that the terminal phosphate group impacted the fragmentation behaviors of RNA oligos in a way that depended on the precursor charge state and the oligo length. Firstly, a terminal phosphate group trends to impart higher negative charges to precursor ions of RNA oligos. Secondly, a terminal phosphate group(s) bestows a predominant loss of phosphoric acid ($-H_3PO_4$) or metaphosphoric acid anion ($-[PO_3]^-$) from RNA oligo precursors in their CID or HCD spectra, especially for precursors of intermediate charge states. This uninformative loss decreases the intensity of sequencing ions (*a-, a-B*，*b-, c-, d-, w-, x-, y-, z-* ions), hindering the sequencing and characterization of RNA oligos by CID/HCD. The terminal phosphate groups can be removed using calf intestinal alkaline phosphatase (CIP), and we show that CIP treatment improved MS analysis of RNAs with a terminal phosphate group on either or both ends. Additionally, we found that the peak intensity of *m/z* 158.925, which we have identified as a dehydrated pyrophosphate anion, $[HP_2O_6]^-$, is markedly increased by the presence of a terminal phosphate group on RNA oligos. This study contributes to the knowledge base on which software tools could be developed for high-throughput characterization of RNA oligos using tandem mass spectrometry.

Keywords: Negative-ion mass spectrometry; RNA oligonucleotide; Collision-induced dissociation (CID); Higher-energy collisional dissociation (HCD); Fragmentation behavior

1. Introduction

Ribonucleic acids (RNA) play critical roles in a variety of signaling and regulatory processes across all three domains of life[1]. Decoding the compositions of RNA—both sequence and modification—will provide fundamental insights and help to unlock the mechanisms of those processes. Mass spectrometry (MS) has become an indispensable analytical method to identify and characterize proteins and small molecules, by virtue of its high sensitivity, high accuracy and high throughput. In recent years, MS has also been increasingly employed in RNA analysis as an expanding field, including RNA sequencing[2,3], single-nucleotide polymorphism (SNP) identification[4], post-transcriptional modifications (PTxMs) characterization[5,6], quality inspection of synthetic RNA[7], thanks to the comprehensive investigations on MS fragmentation behaviors of nucleic acids during the last 40 years[8-15].

Beginning with fast-atom bombardment (FAB)[16], fragmentation behaviors of nucleic acids have been studied using different ionization methods including electrospray ionization (ESI)[10,17] and matrix assisted laser desorption/ionization (MALDI)[10,18]. A wide range of fragmentation methods have been applied to nucleic acids, including collision-induced dissociation (CID), higher-energy collisional dissociation (HCD)[19], electron detachment dissociation (EDD)[20], negative electron transfer dissociation (NETD)[21], ultraviolet photodissociation (UVPD)[22]. Among them, CID/HCD is most widely utilized to analyze RNA molecule. In CID/HCD spectra, the predominant fragment ions of RNA oligos are *c*- and *y*-ions, arising from cleavage of 5' P-O bond at the phosphodiester backbone, according to the nomenclature of oligonucleotide fragment ion proposed by McLuckey et al[23]. The (a-B)- and *w*-ions, arising from cleavage of 3' C-O bond, which are the primary fragment pathway in CID/HCD spectra of DNA, can also be observed in RNA, but of less abundance[8-10]. There are several factors that affect RNA fragmentation behaviors, such as the precursor charge state[12,15,24], the fragmentation energy deposited[12], the metal ion adduction[13,24,25], the modification on nucleotides[11,19], or phosphates[26]. Our previous study on CID fragmentation patterns of RNA oligos with both terminal hydroxyl groups showed that the sequencing ions (*a, a-B, b, c, d, w, x, y, z-type*) tend to decrease in overall abundance while the nucleobase losses increased markedly when the precursor charge varies from a low, intermediate to a high state[15].

Generally, there are three different groups at the termini of RNA, hydroxyl (-OH), phosphate (-P), and 2',3'-cyclic phosphate (cP). -OH and -P can present on both 5'- and 3'-ends while cP can only be found on 3'-end[27]. Although the fragmentation behaviors have been published in several literatures, the effects of the termini groups to RNA fragmentation were seldom studied. Krivos et al. reported the loss of phosphoric acid ($-H_3PO_4$) in CID spectra of RNase digested RNA containing a 3'-phosphate group, which would reduce the abundance of sequencing ions[28]. However, the effect of 5'-end group was not involved in their study.

To fill this gap and expand the fragmentation behaviors of RNA oligos, this study characterized eight RNA oligos we synthesized with two different sequences (GUCA and AUCGAUCG). Each of them contains one of four types of terminal group combinations, (1)

5'-OH and 3'-OH, (2) 5'-OH and 3'-P, (3) 5'-P and 3'-OH, and (4) 5'-P and 3'-P. We used all these RNA oligos to systematically investigate the effect of the terminal phosphate group on fragmentation behaviors in negative-ion CID/HCD spectra. We found that the terminal phosphate group on RNA oligos influence CID/HCD fragmentation in a way that relies on both the precursor charge and the oligo length. The negative-ion CID/HCD spectra of RNA oligos with a terminal phosphate group(s) are dominated by precursors with the loss of phosphoric acid or metaphosphoric acid anion, especially for precursors of intermediate charge states. Furthermore, our data sets revealed that calf intestinal alkaline phosphatase (CIP) can effectively remove terminal phosphates and CIP treatment is highly recommended to improve MS analysis of RNA oligos with a terminal phosphate group on either or both ends.

2. Materials and methods

2.1 Materials and sample preparation

Eight RNA oligos with two sequences (GUCA and AUCGAUCG) and four types of terminal-group combinations (from 5' to 3': HO-X-OH, HO-X-P, P-X-OH, P-X-P, X denotes one of the two sequences) were synthesized by Biolino Acid Technology (Tianjin, China). The detailed information of these RNA oligos is provided in Supplemental Table S1. Four AUCGAUCG oligos were freshly synthesized while four GUCA oligos were synthesized earlier for other purpose and had been stored in -80℃ for about one year. For direct infusion MS analysis, RNA oligos were dissolved in water to make 0.15 mM solutions (measured by NanoDrop 1000 Spectrophotometer, Thermo Fisher Scientific) and then diluted to 5 μM in 7.5 mM hexylamine (Energy chemical, Shanghai, China), 50% acetonitrile (Fisher Scientific) before being sprayed into mass spectrometry using a syringe pump. For LC-MS/MS analysis, 250 pmol of each RNA oligo was treated with 10 units calf intestinal alkaline phosphatase (CIP, New England Biolabs Inc., M0290S) in 37℃ for 30 minutes. Then, the digests were desalted using ZipTip Piette Tips (Merck Millipore) following the manufacturer's protocol before LC-MS/MS analysis. 250 pmol of each RNA oligo without CIP and ZipTip treatment was also conducted LC-MS/MS analysis as the negative control.

2.2 MS parameters

All experiments were performed on an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific) in the negative-ion mode. For direct infusion analysis, the sample solutions were sprayed into the electrospray source at a flow rate of 2 μl/min by a syringe pump. The ion spray voltage was -2.5 kV. The MS1 mass range was 150-1500 m/z for GUCA and 180-2700 m/z for AUCGAUCG. An inclusion list including all theoretical precursor *m/z* values from charge 1^- to charge 9^- for each RNA oligo was used. The fragmentation method was CID or HCD. The normalized collisional energy (NCE) was varied from 10% to 50%, with a step of 5%. For MS/MS, auto mass range modes were selected for all samples. The AGC target was 1E6 for MS1 and 1E5 for MS2. The maximum injection time was automatically adjusted by the system for both MS1 and MS2.

Orbitrap was used as the mass analyzer for both MS1 and MS2 at a resolution of 120k. The isolation window was 0.4 m/z to reduce the proportion of mixed spectra. LC-MS/MS analysis was performed on an Easy-nLC 1200 UHPLC (Thermo Fisher Scientific) coupled to the Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer. RNA samples were loaded and separated on a precolumn (75 μm inner diameter, 7 cm long, packed with Xbridge BEH C18 2.5 μm beads from Waters Corporation) and sprayed by an empty column (75 μm inner diameter, 3 cm long) with an in-housed pulled 5 μm tip. The column end fittings (Valco Instruments Company Incorporated, C-NEF.5XFPK.35S1) were used for the connection of liquid chromatography, the precolumn and the empty column to avoid the dislodge or fracture of the in situ cast silicate-based frits[29]. A linear gradient made with buffer A (7.5 mM triethylamine, Thermo Fisher Scientific, 25108) and buffer B (80% acetonitrile and 7.5 mM triethylamine) as follows: 0-22 min, 0-30% B; 22-23 min, 30-90% B; 23-27 min, 90% B; 27-28 min, 95-0% B; 28-30 min, 0% B; 30-45 min, 0% B. The flow rate was 350 nl/min. Cycle time was set to 3s. The resolution was 120 k for MS1 and 60k for MS2.

2.3 Data analysis

Data analysis was conducted using in-house python and MATLAB scripts. Theoretical *m/z* values of *a-, a-B*，*b-, c-, d-, w-, x-, y-, z*-ions, internal fragments and the precursors with a single nucleobase loss were calculated and matched with the observed peaks. For RNA oligos with one or two terminal phosphate groups, the precursors with the loss of phosphoric acid or metaphosphoric acid anion were also calculated and matched with the observed peaks. Cyanate anion (NCO^-) loss was not included here since the terminal pyrimidine bases were not contained in all oligos we used[15,30]. The observed monoisotopic peak was assigned to a theoretical ion within a mass tolerance of 10 ppm. For CID spectra, NCE 35 was chosen as the representative energy although the spectra under different NCEs were acquired. This is because our data analysis (not shown) indicated that the sequencing ions are optimal under NCE 35 in terms of their total intensity. For all HCD spectra, we selected the spectra under NCE 40 to study the pattern of low-mass ions (100 to 220 Da) because of their relative high abundance.

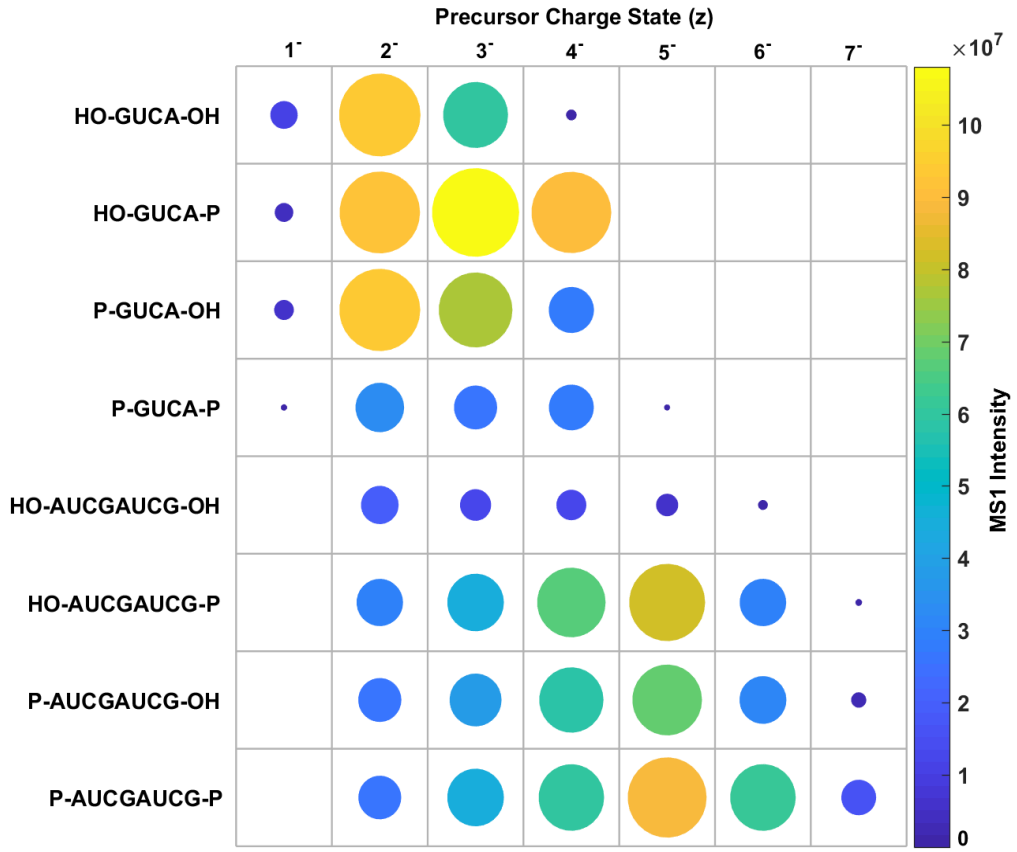
3. Results and discussion

3.1 A terminal phosphate group trends to impart higher negative charges to precursor ions of RNA oligos

Multiply charged precursor ions are a characteristic feature of RNA oligos ionized by ESI in the negative-ion mode[15]. In this section we will examine whether the different terminal groups have an impact on the charge state distribution of an RNA oligo. To address this issue, the concentrations of all eight RNA oligos we synthesized and introduced previously were adjusted to be consistent and then injected to MS analysis (Supplemental Table S1). The MS1 intensities and their fractions among all charge states observed for each RNA oligo were

shown in Figure 1(a) and Figure 1(b), respectively. In the first row of Figure 1b, the most abundant charge state of HO-GUCA-OH is charge 2^- (57.0%), followed by 3^- (36.0%), 1^- (6.2%), and 4^- (0.8%). For other three GUCA oligos containing terminal phosphate groups, relatively uniform abundances of charge 2^-, 3^- and 4^- (most around 30%) and low abundance of charge 1^- (<3%) were observed (Figure 1b). The *m/z* value of charge 5^- was only detected in P-GUCA-P, where the "charge level" (CL, the ratio of the actual charge state to the total number of phosphate groups within the oligo) is 100%[31]. The similar pattern of charge state distributions was observed for the oligos of AUCGAUCG. Here, we used an average charge (avg_z) to show the difference of charge state distributions among all terminal combinations for each oligo. avg_z was computed by the sum of all charge numbers detected, each multiplied by their corresponding intensity percentage, and the value shown as the red stars in Figure 1b. avg_z was shifted to a higher negative charge for RNA oligos with one or two terminal phosphates. All these comparisons of charge distribution observations revealed that a terminal phosphate group trends to impart higher negative charges to precursor ions of RNA oligos. To our best knowledge, no detailed mechanism of this trend has yet been reported. We hypothesize that the terminal phosphate enables to be deprotonated in the negative-ion MS mode.

(a)



(b)

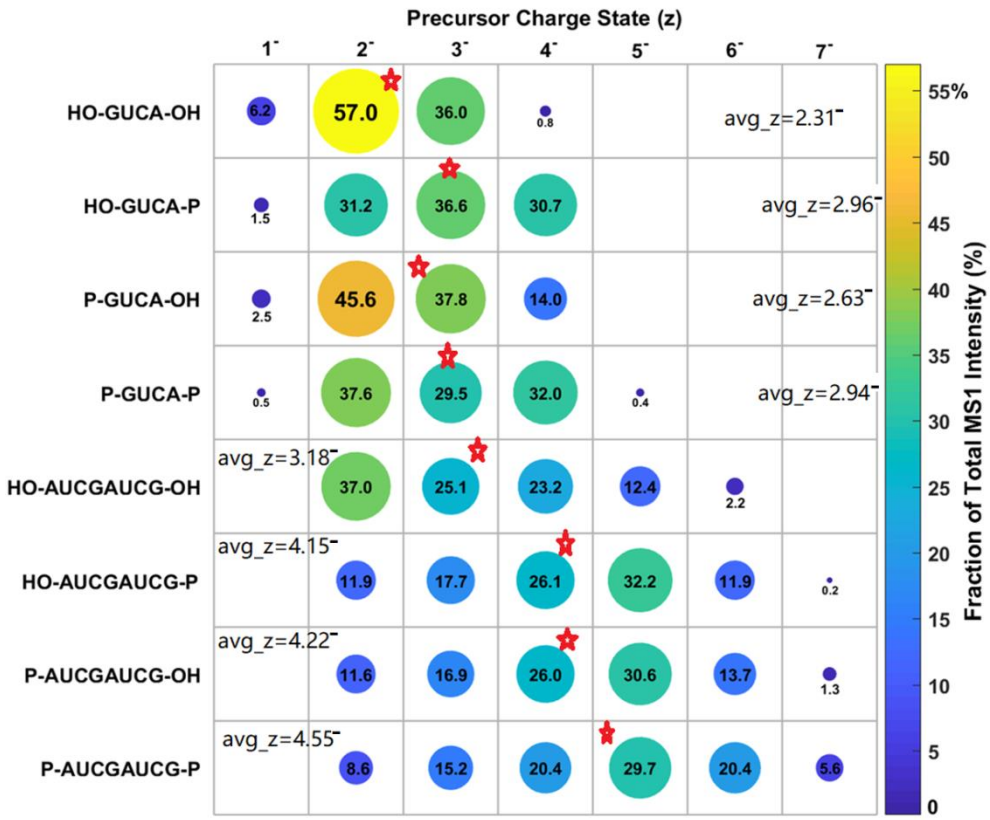


Figure 1. The charge state distributions of eight RNA oligonucleotides. (a) MS1 intensity, (b) Fraction of total MS1 intensity (%). In figure 1(b), the number inside/below the circle is the MS1 intensity percentage of the summed intensities of all charge states observed. The red star is the average charge state for each oligonucleotide, computed by the sum of all charge numbers multiplied by their intensity percentage.

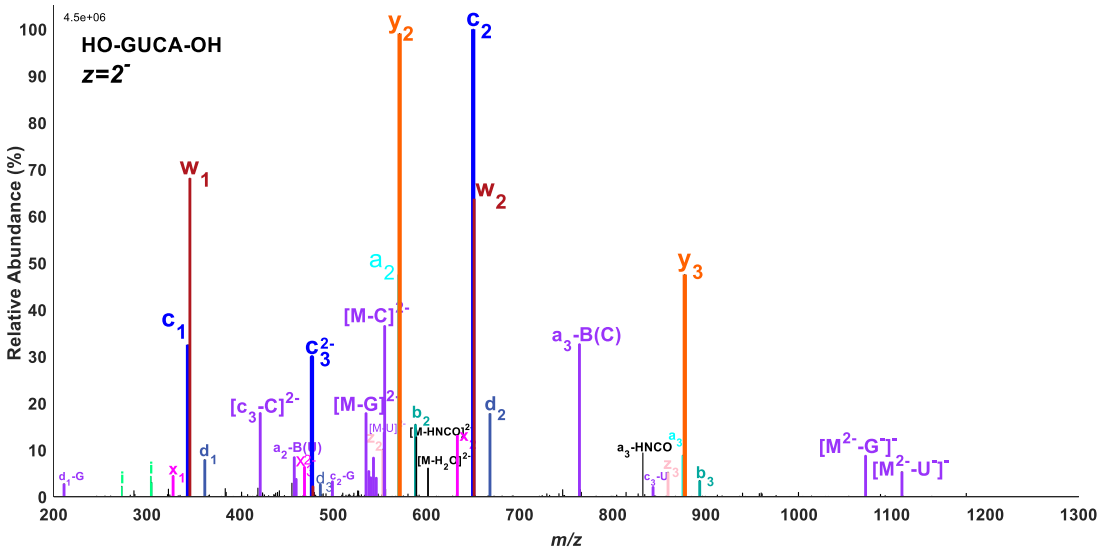
3.2 A terminal phosphate group(s) bestows a predominant loss of phosphoric acid or metaphosphoric acid from RNA oligo precursors with medium charges

In addition to the charge state distributions at MS1 level, the terminal phosphate groups also affect, at MS/MS level, the fragmentation pattern of CID and HCD when compared with the oligo with the same sequence but both –OH termini. Figure 2 shows the example of CID spectra from four GUCA oligos of charge 2^- , fragmented under NCE 35%. The peaks from precursor ions with a neutral phosphoric acid loss ($-H_3PO_4$) or its variations, include the charged phosphoric acid loss ($-[H_2PO_4]^-$) or metaphosphoric acid loss ($-[PO_3]^-$) are highlighted in red. No loss of phosphoric acid from precursor ions were observed in CID spectra of HO-GUCA-OH (Figure 2a). For other three GUCA oligos with terminal phosphate groups, the precursor ions with the loss of phosphoric acid were all the base peak in their CID spectra (Figures 2b, 2c, and 2d), leading to much lower abundance of sequencing ions when compared with the spectrum in Figure 2a. The same analysis was conducted for AUCGAUCG of charge 4^- and the similar trend was observed in Figure 3. The difference is that the base peak from three AUCGAUCG oligos with terminal phosphate group has become the charged metaphosphoric acid loss ($-[PO_3]^-$).

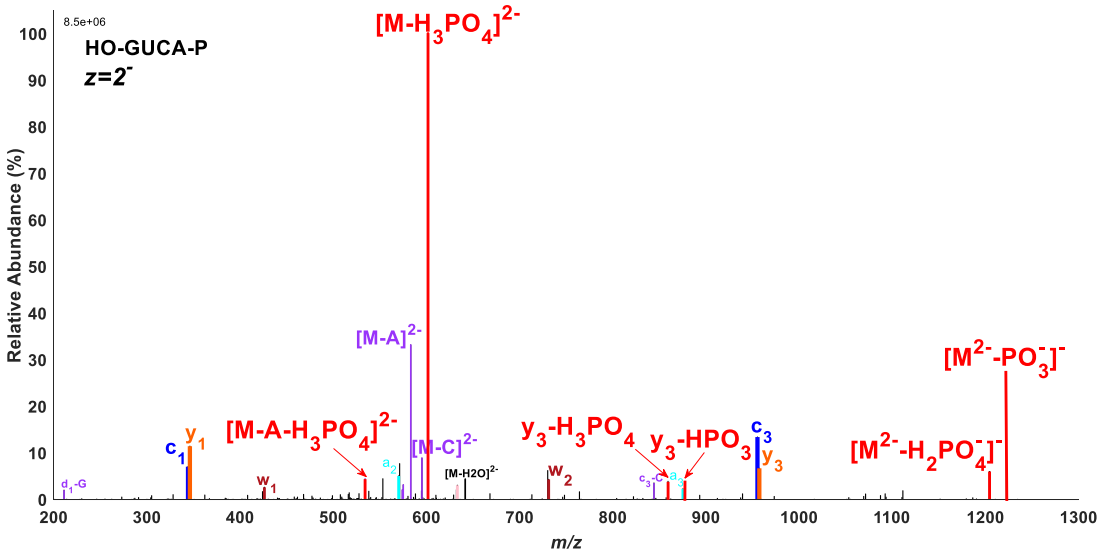
There are also some variations (red peaks annotated) observed although they are relatively low abundant in Figure 3.

Additionally, we wonder if the extent of the impact of RNA terminal phosphate group on the fragmentation is related to the precursor charge state. To this end, we analyzed CID spectra from all charge states detected from eight oligos. Figure 4 shows the fraction of the total peak intensity from the loss of phosphoric acid or its variations (PI) and the sequencing ions (SI) in CID spectra of NCE 35%. The total intensity fraction of these two types of ions, PI and SI, was computed by the summed intensity of each type divided by the summed intensity of all peaks in each CID spectrum. For CID spectra from three GUCA oligos with terminal phosphate group, the charge states with significant PI are 3^-, 4^- and 2^-. For CID spectra from three AUCGAUCG oligos with terminal phosphate group, the charge states with significant PI are 5^- and 4^-. Therefore, from the example of these two sequences, we observed that in the negative-ion CID spectra of RNA oligos with terminal phosphate group, the precursors with the loss of phosphoric acid or metaphosphoric acid anion are predominant, especially for precursors of medium charge states.

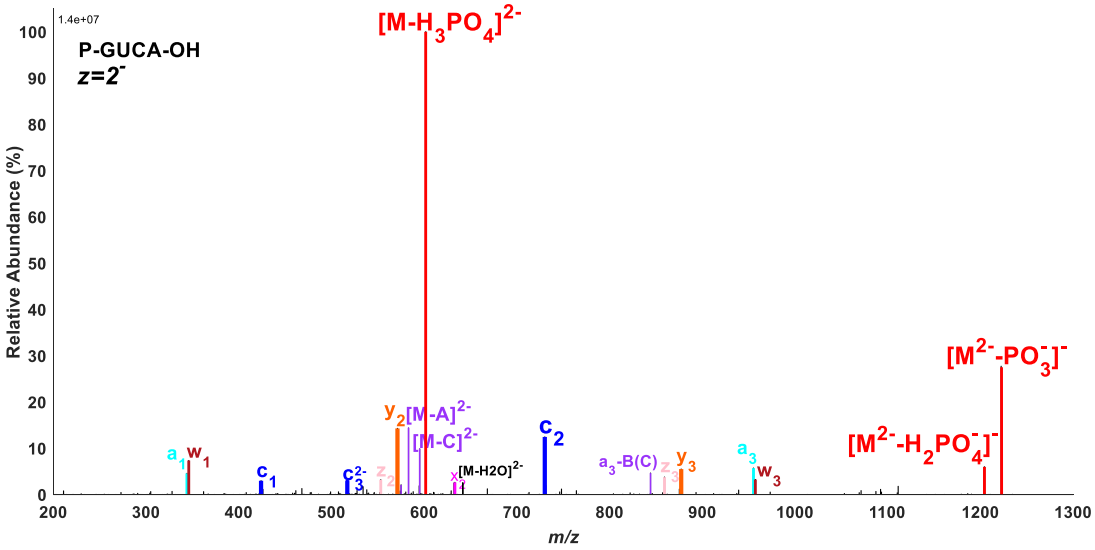
(a)



(b)


8.5e+06
HO-GUCA-P
z=2-
[M-H3PO4]2-
[M-A]2-
[M-A-H3PO4]2-
[M-C]2-
[M-H2O]2-
y3-H3PO4
y3-HPO3
[M2--PO3-]-
[M2--H2PO4-]-
Relative Abundance (%)
m/z

(c)


1.4e+07
P-GUCA-OH
z=2-
[M-H3PO4]2-
[M-A]2-
[M-C]2-
[M-H2O]2-
a3-B(C)
[M2--PO3-]-
[M2--H2PO4-]-
Relative Abundance (%)
m/z

(d)

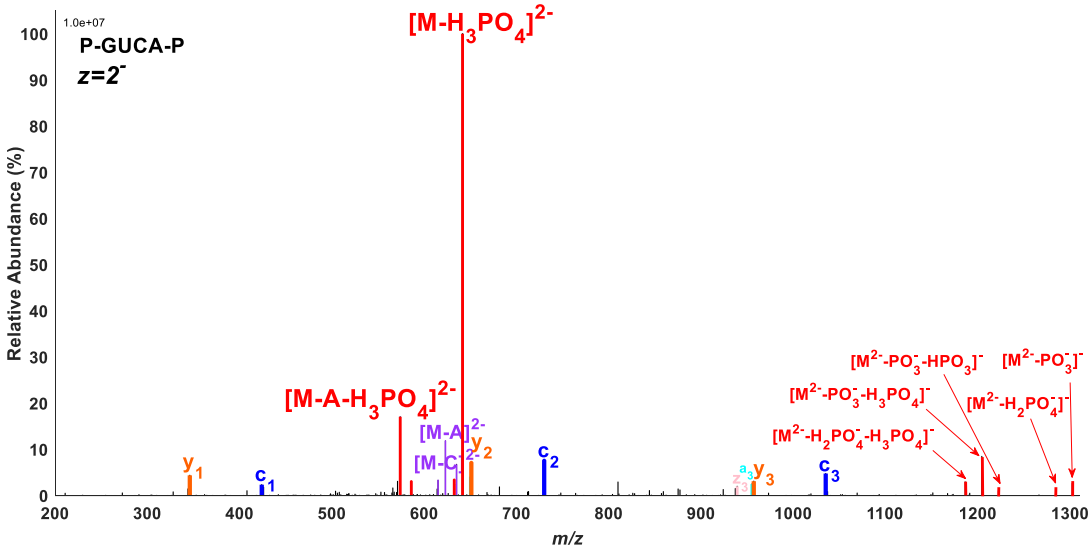
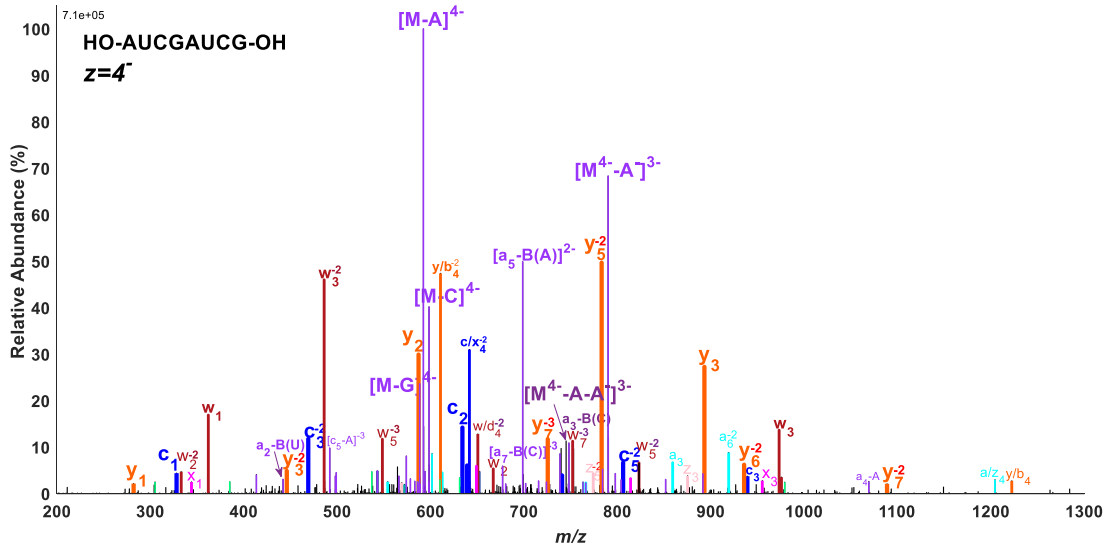
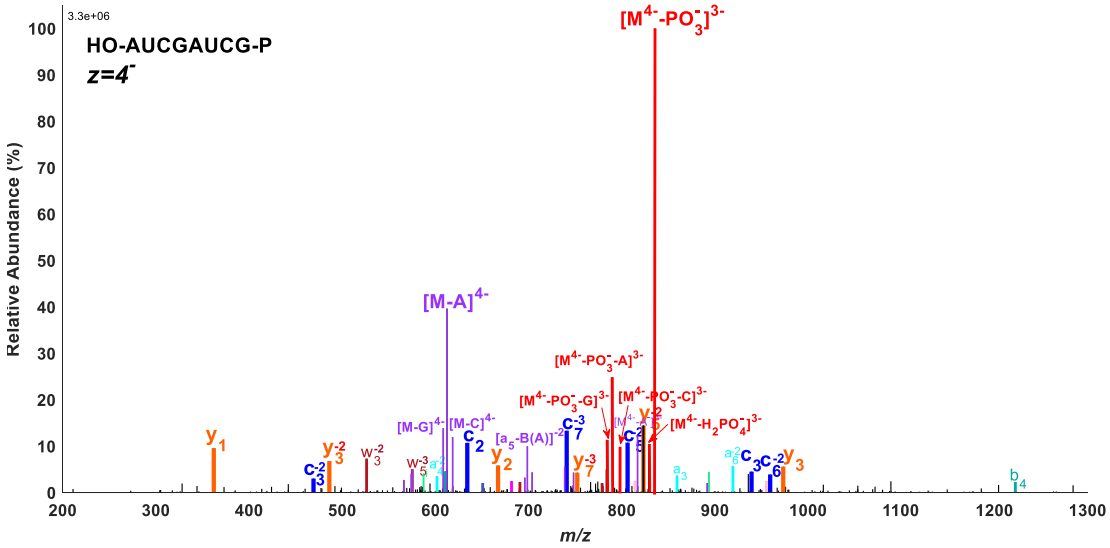

1.0e+07
P-GUCA-P
z=2-
[M-H3PO4]2-
[M-A-H3PO4]2-
[M-A]2-
[M-C]2-
[M2--PO3--HPO3]-
[M2--PO3--H3PO4]-
[M2--H2PO4--H3PO4]-
[M2--H2PO4-]-
[M2--PO3-]-
Relative Abundance (%)
m/z

Figure 2. The effect of terminal phosphate group on CID spectra of GUCA oligonucleotides (charge z=2$^-$, NCE=35%). (a) HO-GUCA-OH, (b) HO-GUCA-P, (c) P-GUCA-OH, (d) P-GUCA-P. The peaks assigned to precursor ions with phosphoric acid loss (-H_3PO_4) or its charged variations are highlighted in red. All green peaks with an 'i' letter on top of them in Figure (a) are the internal fragments, which are generated from cleavages of two phosphodiester bonds. They contain neither the 5' nor the 3' end.

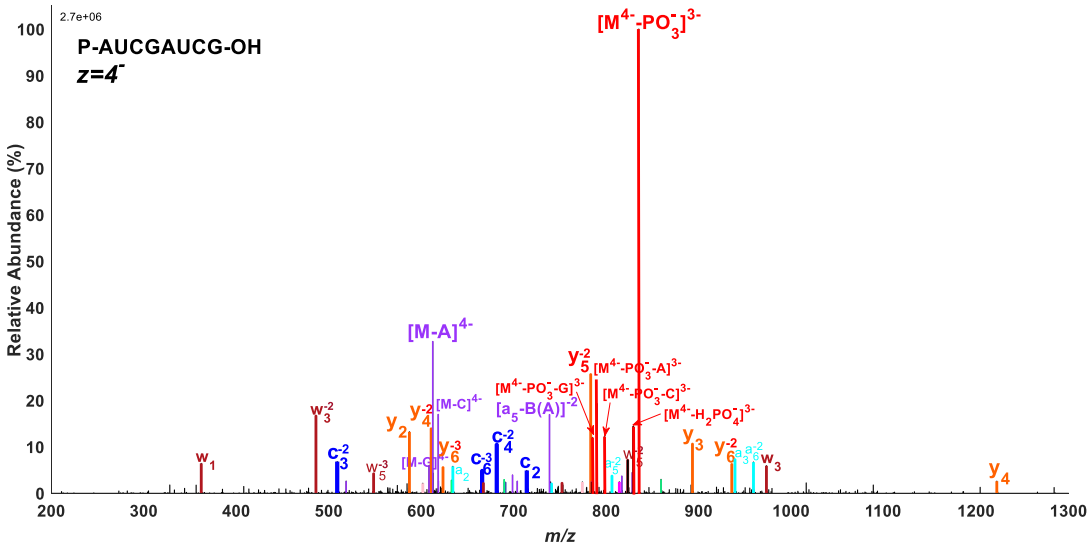
(a)



(b)



(c)



(d)

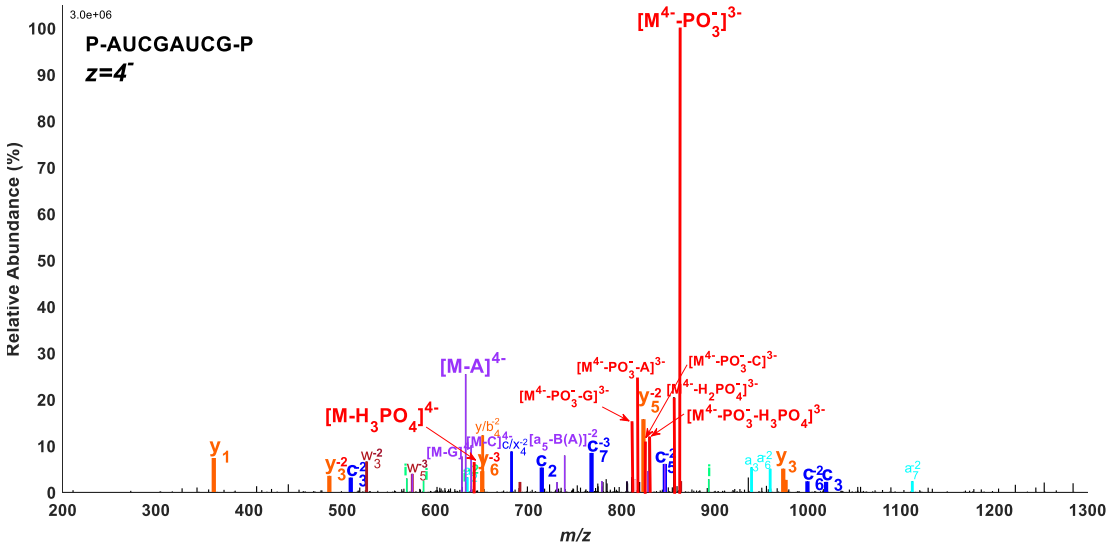


Figure 3. The effect of terminal phosphate group on CID spectra of AUCGAUCG oligonucleotides (charge z=4$^-$ and NCE=35%). (a) HO-AUCGAUCG-OH, (b) HO-AUCGAUCG-P, (c) P-AUCGAUCG-OH, (d) P-AUCGAUCG-P. The peaks assigned to precursor ions with metaphosphoric acid loss (-[PO_3]$^-$) or its variations are highlighted in red.

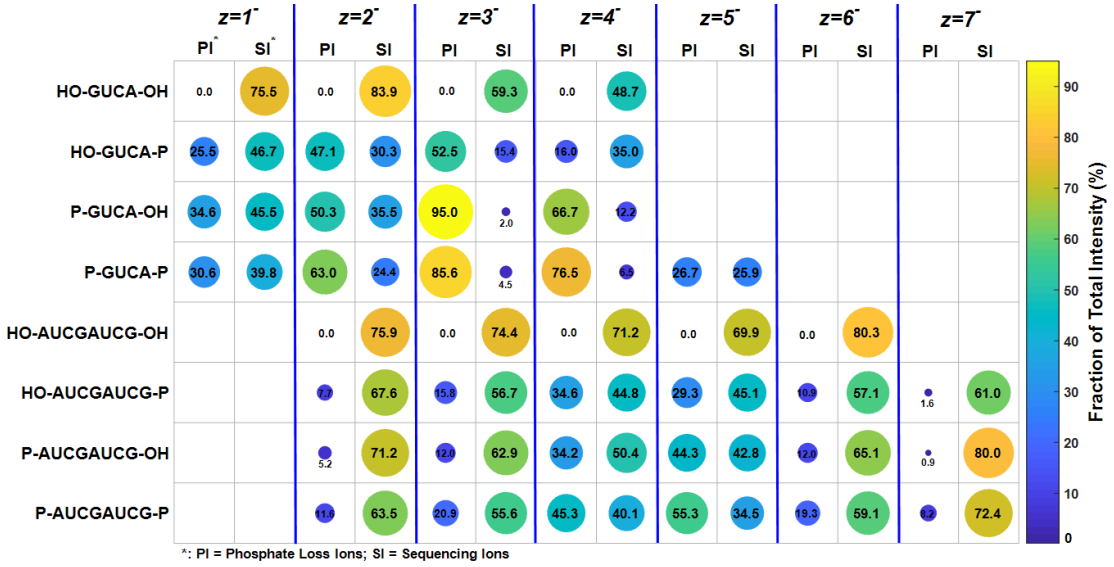


Figure 4. The fraction of the total intensity from phosphoric acid losses or its variations (PI) and the sequencing ions (SI) in CID spectra of NCE 35%. The intensity fraction, the number (%) shown inside/below the circle, was the percentage computed by the summed intensity of each class (PI or SI) divided by the summed intensity of all peaks in each CID spectrum.

3.3 The dehydrated pyrophosphate anion (*m/z* 158.925) in HCD spectra of RNA oligos is markedly increased by the presence of a terminal phosphate group

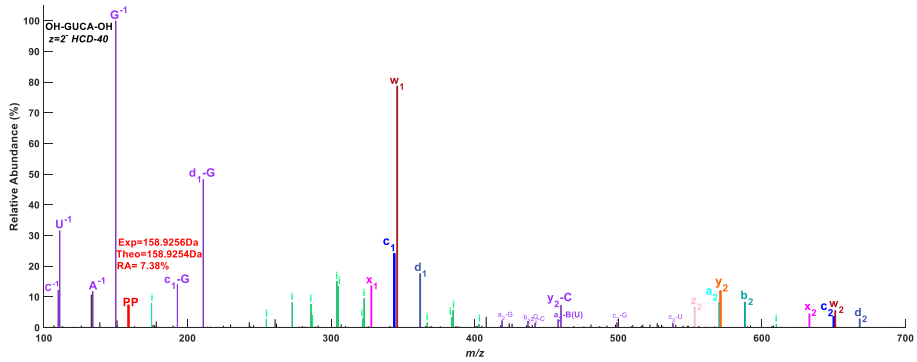
Unlike the limitation of 'one-third mass cutoff' in CID spectra, the peak pattern within low *m/z* region (*m/z* 100~220) in HCD spectra can be clearly observed. We took HCD spectra of GUCA (z=2$^-$) under NCE 40% as an example in Figure 5. Initially, we found there was an unknown peak, around *m/z* 158.93. We did not find any annotation or report on this peak after searching the literatures on MS fragmentation of RNA or DNA. We then tried to search it in a NIST MS database of small molecules and found that the peak 158.93 was typically observed in the negative-ion HCD spectra of many molecules with multiple phosphate groups, such as Uridine 5'-triphosphate, phytic acid, among others. After the comparison and accurate mass analysis, we have identified ultimately this peak as a dehydrated pyrophosphate anion with the formula of $[HP_2O_6]^-$. Actually, the abundant pyrophosphate anions have been reported in lipid A isolated from several bacteria[32] and other small molecules[33-36]. Nevertheless, how this ion changes in intensity from different terminal oligos, and how it generates during an HCD procedure, remain untouched.

Figure 5 showed the relative abundance (RA) of the peak of dehydrated pyrophosphate anion from oligos of GUCA with different terminal groups. For simplicity, we use 'PP' in this figure in red to denote this anion, *m/z* 158.925. PP's RA increases from 7.38% for both –OH termini in Figure 5a, to 32.60% in Figure 5b and 23.26% in Figure 5c for only one terminal phosphate group, then to 100% in Figure 5d for two terminal phosphate groups. The increasing number of terminal phosphate groups from zero, one

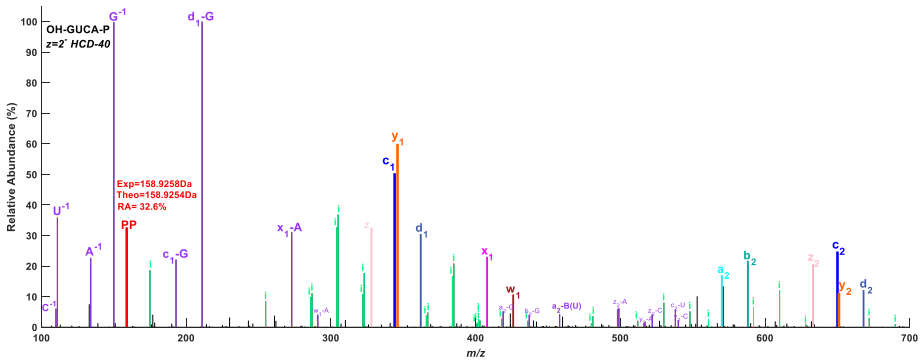
to two, inclines to enhance PP's RA in HCD spectra of NCE 40%.

In order to assess the contribution of the phosphorus atom to PP from terminal or backbone phosphates, we plotted the trend curve of PP's RA with an increasing NCE from 25% to 50%, an interval of 5%, in Figure 6. Under all these NCEs，PP from the oligo with two terminal phosphates is much higher than that from one or no phosphate terminus. The oligos HO-UUAAC-OH and HO-GGUAC-OH contain the same number (4) of phosphates with the oligos P-GUCA-OH and HO-GUCA-P, but PP of the latter two oligos are much higher than those of the former two under high NCEs (40~50). Similarly, the oligos HO-GGCAUU-OH and HO-GCUAAU-OH contain five phosphates, so does P-GUCA-P, but PP of the latter one is also much higher than those of the former two under all NCEs shown in Figure 6. Here, a part of HCD spectra of these oligos were analyzed from our previous research[15]. All these data reveal that RNA oligos with the terminal phosphate group could yield a higher PP ion in HCD spectra than those with only backbone phosphates. This characteristic can be attributed to the labile phosphate group at the termini of each RNA oligo.

(a)



(b)



(c)

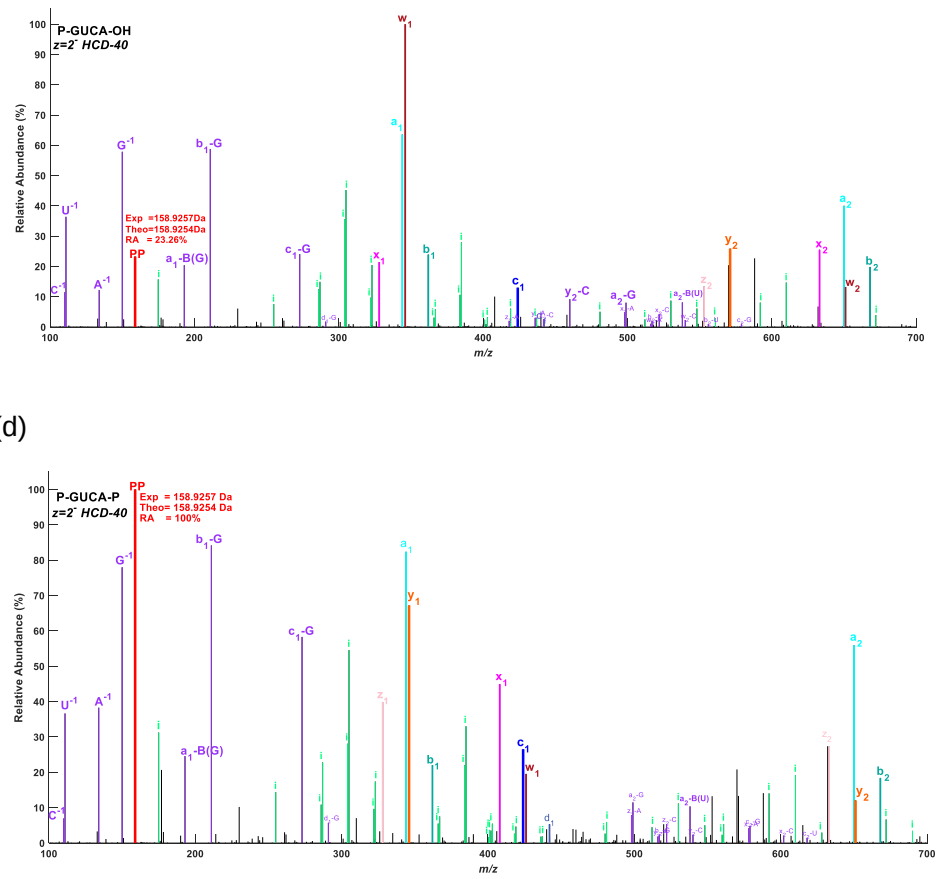


Figure 5. The dehydrated pyrophosphate anion (PP) in HCD spectra (NCE=40%, z=2$^-$) from the oligo GUCA. (a) HO-GUCA-OH, (b) HO-GUCA-P, (c) P-GUCA-OH, (d) P-GUCA-P. All green peaks with an 'i' letter on top of them are the internal fragments, which are generated from cleavages of two phosphodiester bonds. They contain neither the 5' nor the 3' end. Internal fragments are more frequently observed in HCD spectra than in CID spectra of RNA oligos.

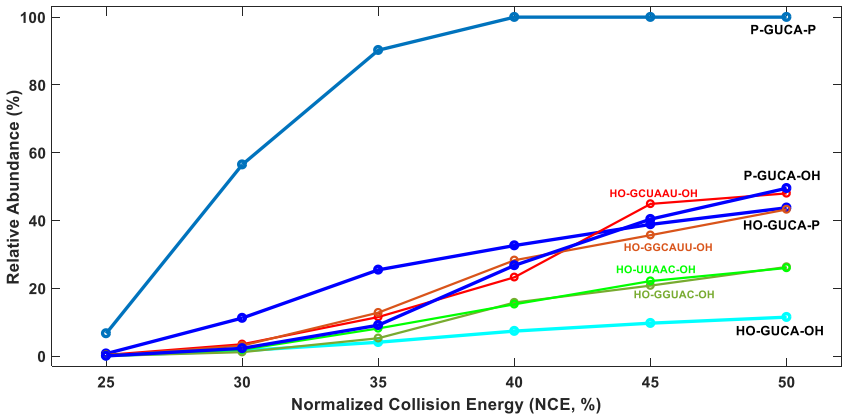


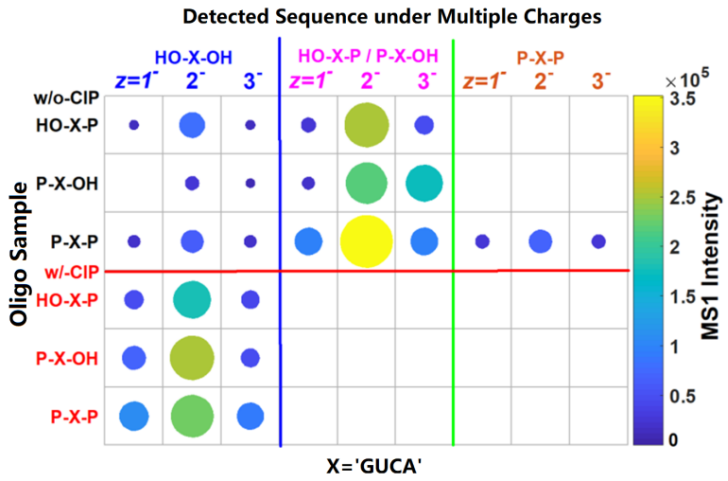
Figure 6. The relative abundance (RA) of the peak 158.925 changes with the increasing NCEs in HCD spectra (NCE=40%). The comparisons with the oligos of the same number of phosphate groups are also plotted.

3.4 CIP treatment can effectively remove terminal phosphate groups of RNA oligonucleotides

Based on our previous finding that the abundant phosphoric/metaphosphoric acid loss from precursor ions caused the low level of sequencing ions in RNA oligos with terminal phosphate group, we think that reducing the charge states or removing terminal phosphate during sample preparation can facilitate RNA oligo sequencing by MS/MS. Charge states can be reduced by manipulation of sample solution conditions[37,38,39], self-modified ion source[40,41] or self-modified quadrupole ion trap[42]. The terminal phosphate group of RNA can be removed by alkaline phosphatase digestion. Krivos et al. used bacterial alkaline phosphatase (BAP) to remove the 3'-terminal phosphate from RNase digestion products to improve RNA sequence coverage[28]. However, BAP is difficult to remove from the dephosphorylation reaction systems. Calf intestinal alkaline phosphatase (CIP) is a better choice because it can be readily digested with proteinase K or inactivated by heating[43].

Here, we compared MS1 spectra of all six RNA oligos containing terminal phosphate groups with or without CIP treatment to test whether CIP treatment can effectively remove terminal phosphates of RNA oligos. Figure 7 summarized the detected compositions of each oligo sample. Each row is the oligo sample and each column is the detected sequence (including the terminal groups) with its precursor charge state at the top. As shown in Figures 7a and 7b, for GUCA and AUCGAUCG with only one terminal phosphate (The first two rows), a little fraction of them have lost terminal phosphate before CIP treatment (the columns on the left of the blue vertical line). This is probable because the terminal phosphate groups are labile and can be lost during storage and experimentations. After CIP treatment, all terminal phosphate groups were removed and only oligos without terminal phosphates were detected in MS1 spectra. As for GUCA with two terminal phosphates, oligos with two, one and no terminal phosphates were all detected before CIP treatment (Figure 7a). However, only oligos with two terminal phosphate groups were observed in MS1 spectra of freshly synthesized P-AUCGAUCG-P sample (Figure 7b). This difference in sample composition of these two sets of oligos can be attributed to that four GUCA oligos had been stored in -80℃ for about one year and four AUCGAUCG oligos were freshly synthesized. Collectively, after CIP treatment, all terminal phosphate groups had been eliminated, and the CID fragmentation behaviors of RNA oligos with both terminal hydroxyl groups we reported in ref. 15 has also indicated that this sample treatment could improve MS analysis for RNA sequencing and PTxM characterization using tandem mass spectrometry. Therefore, we highly recommend CIP treatment prior to MS analysis when the digested products from RNA samples contain either one or two terminal phosphate groups.

(a)



(b)

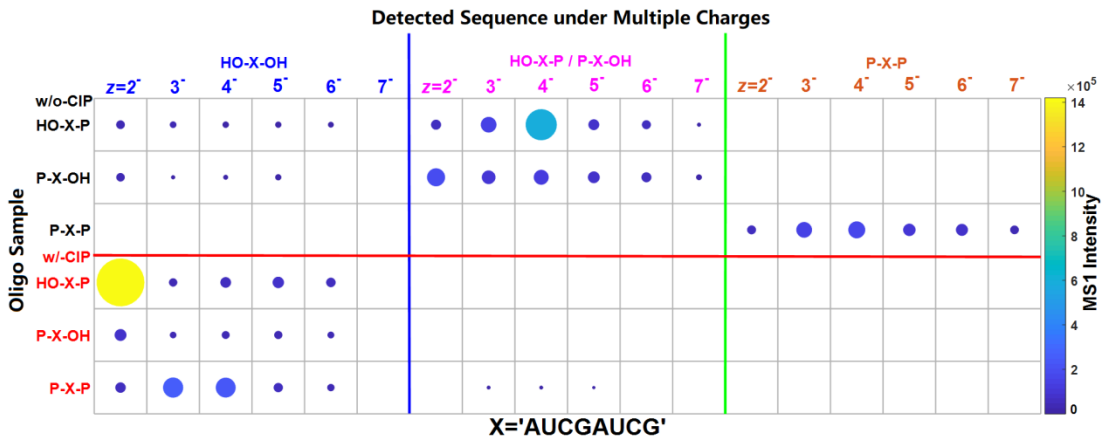


Figure 7. The comparison of RNA oligonucleotide distributions using MS1 intensity before and after CIP (Calf Intestinal Alkaline Phosphatase) treatments. The three rows above the red line correspond to the oligos without CIP treatment and below are treated by CIP.

4. Conclusions

To investigate the effect of terminal phosphate groups on RNA fragmentation, CID spectra of NCE 35 from eight synthetic RNA oligonucleotides of two sequences (GUCA and AUCGAUCG) and four types of terminal group combinations (5'-OH and 3'-OH, 5'-OH and 3'-P, 5'-P and 3'-OH, 5'-P and 3'-P) were characterized. We found that the terminal phosphate groups had a crucial influence on RNA oligonucleotide fragmentation by CID/HCD in a way that depended on both the precursor charge state and the oligo

length. At MS1 level, higher negative precursor charges were detected for RNA oligonucleotides with terminal phosphate groups. At MS/MS level, the loss of phosphoric acid or metaphosphoric acid anion from precursor ions was observed with a high-abundance peak, especially for the oligos with medium charge states. This type of loss, neutral or charged as an anion, caused the lower abundance of sequencing ions in RNA oligonucleotides with terminal phosphate groups. Based on these findings, we attempted to remove the terminal phosphate groups during sample preparation for more effective RNA analysis using CIP. Our data sets verified that CIP treatment can effectively remove terminal phosphate groups of RNA oligonucleotides. In summary, this investigation on the effect of the terminal phosphate groups on CID/HCD fragmentation behaviors will be helpful to sample preparation of RNA MS analysis and simultaneously facilitate the search engine development for RNA sequencing by high-throughput mass spectrometry.

Acknowledgments

This study was supported by the National Key Research and Development Program of China (Grant 2020YFF01014505), the Ministry of Science and Technology of China, and Beijing Municipal Science and Technology Commission. We also thank Yan Ma in Metabolomics Center, National Institute of Biological Sciences, Beijing, China (NIBS) and Ji-Shuai Zhang in the Department of Chemistry, Peking University, China for their constructive advice during the identification of the peak 158.925 in HCD spectra.

Associated Content

Supplemental Table S1, the detailed information of eight RNA oligonucleotides synthesized in this manuscript

REFERENCES

(1) Mortimer, S. A., Kidwell, M. A. & Doudna, J. A. Insights into RNA structure and function from genome-wide studies. Nature Reviews Genetics 15, 469-479 (2014).

(2) Jiang, T. et al. Oligonucleotide Sequence Mapping of Large Therapeutic mRNAs via Parallel Ribonuclease Digestions and LC-MS/MS. Anal Chem 91, 8500-8506, (2019).

(3) Thomas, B. & Akoulitchev, A. V. Mass spectrometry of RNA. Trends Biochem Sci 31, 173-181, (2006).

(4) Krahmer, M. T. et al. MS for identification of single nucleotide polymorphisms and MS/MS for discrimination of isomeric PCR products. Anal Chem 72, 4033-4040, (2000).

(5) Gao, Y., Yang, J., Cancilla, M. T., Meng, F. & McLuckey, S. A. Top-down interrogation of chemically modified oligonucleotides by negative electron transfer and collision induced dissociation. Anal Chem 85, 4713-4720, (2013).

(6) Thuring, K., Schmid, K., Keller, P. & Helm, M. Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. Methods 107, 48-56, (2016).

(7) Goyon, A. & Zhang, K. Characterization of Antisense Oligonucleotide Impurities by Ion-Pairing Reversed-Phase and Anion Exchange Chromatography Coupled to Hydrophilic Interaction Liquid Chromatography/Mass Spectrometry Using a Versatile Two-Dimensional Liquid Chromatography Setup. Anal Chem 92, 5944-5951, (2020).

(8) Andersen, T. E.; Kirpekar, F.; Haselmann, K. F. RNA Fragmentation in MALDI Mass Spectrometry Studied by H/D-exchange: Mechanisms of General Applicability to Nucleic Acids. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (10), 1353-1368.

(9) Wu, J. & McLuckey, S. A. Gas-phase fragmentation of oligonucleotide ions. International Journal of Mass Spectrometry 237, 197-241, (2004).

(10) Nordhoff, E., Kirpekar, F. & Roepstorff, P. Mass spectrometry of nucleic acids. Mass spectrometry reviews 15, 67-138 (1996).

(11) Tromp, J. M. & Schurch, S. Gas-phase dissociation of oligoribonucleotides and their analogs studied by electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 16, 1262-1268, (2005).

(12) Huang, T. Y.; Kharlamova, A.; Liu, J.; McLuckey, S. A. Ion Trap Collision-induced Dissociation of Multiply Deprotonated RNA: c/y-ions versus (a-B)/w-ions. *J. Am. Soc. Mass Spectrom.* 2008, *19* (12), 1832-1840.

(13) Nyakas, A., Stucki, S. R. & Schurch, S. Tandem mass spectrometry of modified and platinated oligoribonucleotides. J Am Soc Mass Spectrom 22, 875-887, (2011).

(14) Krause, J., Scalf, M. & Smith, L. M. High resolution characterization of DNA fragment ions produced by ultraviolet matrix-assisted laser desorption/ionization using linear and reflecting time-of-flight mass spectrometry. Journal of the American Society for Mass Spectrometry 10, 423-429 (1999).

(15) Sun, R.X.; Zuo, M.Q.; Zhang, J.S.; Dong, M.Q. Charge-State-Dependent Collision-Induced Dissociation Behaviors of RNA Oligonucleotides via High-Resolution Mass Spectrometry. *J. Am. Soc. Mass Spectrom*. 2023, *34* (8), 1598−1608.

(16) Grotjahn, L., Frank, R. & Blöcker, H. Ultrafast sequencing of oligodeoxyribonucleo tides by FAB-mass spectrometry. Nucleic acids research 10, 4671-4678 (1982).

(17) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* 1989, *246* (4926), 64-71.

(18) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. Protein and Polymer Analyses up to m/z 100,000 by Laser Ionization Time-of-flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* 1988, *2* (8), 151–153.

(19) Jora, M.; Burns, A. P.; Ross, R. L.; Lobue, P. A.; Zhao, R.; Palumbo, C. M.; Beal, P. A.; Addepalli, B.; Limbach, P. A. Differentiating Positional Isomers of Nucleoside Modifications by Higher-energy Collisional Dissociation Mass Spectrometry (HCD MS). *J. Am. Soc. Mass Spectrom.* 2018, *29* (8), 1745-1756.

(20) Yang, J. & Hakansson, K. Fragmentation of oligoribonucleotides from gas-phase ion-electron reactions. J Am Soc Mass Spectrom 17, 1369-1375, (2006).

(21) Peters-Clarke, T. M. et al. Ribonucleic acid sequence characterization by negative electron transfer dissociation mass spectrometry. Analytical chemistry 92, 4436-4444 (2020).

(22) Brodbelt, J. S., Morrison, L. J. & Santos, I. Ultraviolet photodissociation mass spectrometry for analysis of biological molecules. Chemical reviews 120, 3328-3380 (2019).

(23) McLuckey, S. A.; Berker, G. V.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* 1992, *3* (1), 60-70.

(24) Keller, K. M. & Brodbelt, J. S. Charge state-dependent fragmentation of oligonucleotide/metal complexes. J Am Soc Mass Spectrom 16, 28-37, (2005).

(25) Nyakas, A., Eymann, M. & Schurch, S. The influence of Cisplatin on the gas-phase dissociation of oligonucleotides studied by electrospray ionization tandem mass spectrometry. J Am Soc Mass Spectrom 20, 792-804, (2009).

(26) Monn, S. T. & Schurch, S. New aspects of the fragmentation mechanisms of unmodified and methylphosphonate-modified oligonucleotides. J Am Soc Mass Spectrom 18, 984-990, (2007).

(27) Shigematsu, M., Kawamura, T. & Kirino, Y. Generation of 2',3'-Cyclic Phosphate-Containing RNAs as a Hidden Layer of the Transcriptome. Front. Genet., 9, 562 (2018).

(28) Krivos, K. L., Addepalli, B. & Limbach, P. A. Removal of 3'-phosphate group by bacterial alkaline phosphatase improves oligonucleotide sequence coverage of RNase digestion products analyzed by collision-induced dissociation mass spectrometry. Rapid Commun Mass Spectrom 25, 3609-3616, (2011).

(29) McAlister, G. C. et al. Analysis of the acidic proteome with negative electron-transfer dissociation mass spectrometry. Anal Chem 84, 2875-2882, (2012).

(30) Nyakas, A.; Eberle, R. P.; Stucki, S. R.; Schürch, S. More than Charged Base Loss—Revisiting the Fragmentation of Highly Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* 2014, *25* (7), 1155-1166.

(31) Pan, S., Verhoeven, K. & Lee, J. K. Investigation of the initial fragmentation of oligodeoxynucleotides in a quadrupole ion trap: charge level-related base loss. Journal of the American Society for Mass Spectrometry 16, 1853-1865 (2005).

(32)Jones, J.W., Shaffer, S.A., Ernst, R.K., Goodlett, D.R. and Tureček, F., 2008. Determination of pyrophosphorylated forms of lipid A in Gram-negative bacteria using a multivaried mass spectrometric approach. Proceedings of the National Academy of Sciences, 105(35), 12742-12747.

(33)Thomas, M.C., Mitchell, T.W. and Blanksby, S.J., 2005. A comparison of the gas phase acidities of phospholipid headgroups: experimental and computational studies. Journal of the American Society for Mass Spectrometry, 16(6), 926-939.

(34)Hintz, M., Reichenberg, A., Altincicek, B., Bahr, U., Gschwind, R.M., Kollas, A.K., Beck, E., Wiesner, J., Eberl, M. and Jomaa, H., 2001. Identification of (E) - 4 - hydroxy - 3 - methyl - but - 2 - enyl pyrophosphate as a major activator for human γδ T cells in Escherichia coli. FEBS letters, 509(2), 317-322.

(35)Ward, J.L. and Beale, M.H., 2002. Synthesis of (2E)-4-hydroxy-3-methylbut-2-enyl diphosphate, a key intermediate in the biosynthesis of isoprenoids. Journal of the Chemical Society, Perkin Transactions 1(6), 710-712.

(36)Strittmatter, E. F.; Schnier, P. D.; Klassen, J. S.; Williams, E. R. Dissociation Energies of Deoxyribose Nucleotide Dimer Anions Measured Using Blackbody Infrared Radiative Dissociation. J. Am. Soc. Mass. Spectrom. 1999, 10, 1095 – 1104.

(37)Muddiman, D. C., Cheng, X., Udseth, H. R. & Smith, R. D. Charge-state reduction with improved signal intensity of oligonucleotides in electrospray ionization mass spectrometry. Journal of the American Society for Mass Spectrometry 7, 697-706 (1996).

(38)Griffeycor, R. H., Sasmor, H. & Greig, M. J. Oligonucleotide charge states in negative ionization electrospray-mass spectrometry are a function of solution ammonium ion concentration. Journal of the American Society for Mass Spectrometry 8, 155-160 (1997).

(39)Cheng, X., Gale, D. C., Udseth, H. R. & Smith, R. D. Charge state reduction of oligonucleotide negative ions from electrospray ionization. Analytical Chemistry 67, 586-593 (1995).

(40)Scalf, M., Westphall, M. S., Krause, J., Kaufman, S. L. & Smith, L. M. Controlling charge states of large ions. Science 283, 194-197 (1999).

(41)Scalf, M., Westphall, M. S. & Smith, L. M. Charge reduction electrospray mass spectrometry. Analytical chemistry 72, 52-60 (2000).

(42)McLuckey, S. A., Wu, J., Bundy, J. L., Stephenson, J. L. & Hurst, G. B. Oligonucleotide mixture analysis via electrospray and ion/ion reactions in a quadrupole ion trap. Analytical chemistry 74, 976-984 (2002).

(43)Green, M.R. & Sambrook, J. Alkaline phosphatase. Cold Spring Harbor Protocols 2020, pdb. Top100768 (2020)