

Ultra-performance liquid chromatography/tandem mass spectrometry (UPLC/MS/MS) and UPLC/MS^E analysis of RNA oligonucleotides

Vera B. Ivleva, Ying-Qing Yu and Martin Gilar*

Waters Corporation, 34 Maple Street, Milford, MA 01757, USA

Received 8 February 2010; Revised 23 June 2010; Accepted 25 June 2010

Fast and efficient ultra-performance liquid chromatography/tandem mass spectrometry (UPLC/MS/MS) analysis of short interfering RNA oligonucleotides was used for identity confirmation of the target sequence-related impurities. Multiple truncated oligonucleotides and metabolites were identified based on the accurate mass, and their presumed sequence was confirmed by MS/MS and MS^E (alternating low and elevated collision energy scanning modes) methods. Based on the resulting fragmentation of native and chemically modified oligonucleotides, it was found that the MS^E technique is as efficient as the traditional MS/MS method, yet MS^E is more general, faster, and capable of producing higher signal intensities of fragment ions. Fragmentation patterns of modified oligonucleotides were investigated using RNA 2'-ribose substitutions, phosphorothioate RNA, and LNA modifications. The developed sequence confirmation method that uses the MS^E approach was applied to the analysis of *in vitro* hydrolyzed RNA oligonucleotide. The target RNA and metabolites, including the structural isomers, were resolved by UPLC, and their identity was confirmed by MS^E. Simultaneous RNA truncations from both termini were observed. The UPLC quadrupole time-of-flight (QTOF) MS/MS and MS^E methods were shown to be an effective tool for the analysis and sequence confirmation of complex oligonucleotide mixtures. Copyright © 2010 John Wiley & Sons, Ltd.

RNA interference (RNAi) plays a fundamental role in post-transcriptional gene silencing. The mechanism involves short interfering double-stranded RNA (siRNA) molecules and via a cascade of enzymatic reactions leads to degradation of the homologous mRNA sequences and disruption of protein translation.¹ With the knowledge of the sequence, gene silencing experiments are now routinely performed using ~21 nucleotides (nt) long synthetic RNA probes, which are also being developed as therapeutics.²

Synthetic RNA oligonucleotides need to be purified to avoid off-target silencing of undesirable genes. When intended for use as drugs, RNA oligonucleotides and their duplexes need to be well characterized to satisfy regulatory requirements and minimize possible adverse effects.

Modifications of RNA molecules are often introduced in order to improve their enzymatic stability, cell membrane permeability, or increase the half-life of siRNA duplexes.³ However, the modifications make analysis of modified RNA molecules more challenging.⁴

The ion-pairing reversed-phase (IP RP) high-performance liquid chromatographic (HPLC) method is commonly applied for the quantitation and characterization of oligonucleotides samples.^{5–9} The most common contaminants of RNA samples are shorter length fragments, i.e. failure sequences,

generated during the multi-step synthetic process.¹⁰ Oligonucleotide samples, which are designed as therapeutic candidates, must be purified and the remaining contaminants must be characterized, including the ones coeluting with the target oligonucleotide.

MS-compatible ion-pairing buffers recommended for LC/MS applications are now used for efficient LC separation and sensitive MS detection.^{7,11,12} Recently, an ultra-performance LC (UPLC) method was utilized to improve speed and resolution of siRNA separations.^{13,14}

While LC/MS provides accurate mass and can be used to assign the identity of target peaks and failure sequences, in some cases the characterization of therapeutic oligonucleotides requires detailed sequence information. Sequencing of oligonucleotides is traditionally performed by selective enzymatic or chemical cleavage using the Maxam-Gilbert method or using exonucleases to create ladders that are examined by capillary electrophoresis¹⁵ or mass spectrometry.^{16,17} These approaches are time-consuming and often complicated by chemical modifications of therapeutic oligonucleotides.

The collision-induced dissociation (CID) MS approach is more generic, faster, and informative than traditional sequencing methods and it is well suited for the analysis of relatively short oligonucleotides (<25 nt). It is potentially applicable to the sequencing of the modified oligonucleotides which are typically resistant to enzymatic cleavage.

*Correspondence to: M. Gilar, Waters Corporation, 34 Maple Street, Milford, MA 01757, USA.
E-mail: martin_gilar@waters.com

Tandem mass spectrometry (MS/MS) sequencing data interpretation can be challenging due to the presence of multiply charged species and alkali ion adducts in the spectra. The mass resolution of triple quadrupole mass spectrometers and ion trap instruments may not be sufficient for determinations of the charge state of characteristic fragments, especially for the oligonucleotide molecules exceeding several thousand Daltons in molecular weight. Spectra interpretation is also complicated by oligonucleotide complex fragmentation, such as internal cleavages, base neutral losses, or phosphate rearrangement.¹⁸ Quadrupole time-of-flight (QTOF) MS instruments generate data with high mass accuracy and resolution allowing deconvolution of multiply charged fragment ions to their singly charged m/z values in order to facilitate the assignment of fragment ions.

The CID spectra of oligonucleotides exhibit a complex fragmentation pattern that can be altered considerably by a chemical modification of the molecule, such as thiolation of the phosphodiester backbone or 2'-*O*-sugar modifications. Multiple types of modifications are often present in the same oligonucleotide, and various fragmentation mechanisms may take place simultaneously.

Since the pioneering studies of the principal fragmentation pathways by McLuckey,^{19,20} the fragmentation behavior has been reported of predominantly native DNA or RNA molecules,^{21–23} with a limited knowledge of the mechanism of fragmentation and no conclusive information on the structures of all possible characteristic ion fragments. Less is known about fragmentation of the oligonucleotides with sugar and backbone modifications.^{24,25}

Currently, steadily accumulating knowledge of the gas-phase dissociation mechanisms of oligonucleotides, together with greater availability of software tools such as the SOS program,²⁶ show promise for routine and robust MS/MS analysis.

An alternative to the conventional approach of data-dependent MS/MS fragmentation is the MS^E methodology, which utilizes simultaneous fragmentation of all ions in a single LC/MS run without selecting a specific precursor.^{27,28} According to the MS^E strategy, the energy in a collision cell is alternated between low energy (MS) and elevated energy (MS^E) scans in a single LC/MS experiment. The low-energy data provide information about intact precursor ions, whereas elevated energy data contain ion fragmentation information. The precursor ions and their product ions are easily linked and combined via common retention times.

Since the data-independent acquisition MS^E method utilizes all the charge states of an analyte, it has the potential to improve the signal-to-noise (S/N) of the resulting deconvoluted MS/MS spectrum. The MS^E strategy was successfully applied for the characterization of peptide maps^{29,30} or complex protein digests for qualitative and quantitative purposes,^{27,28} as well as for small molecule analyses.³¹

In this report we utilize UPLC/UV/MS^E methods for the analysis of oligonucleotides. UV and MS data are useful for quantitation, whereas MS^E or conventional MS/MS methods are applied for confirmatory oligonucleotide sequencing. We demonstrated structural characterization of 21 nt long RNA and various modified oligonucleotides. UPLC with the MS^E

approach was used for fast confirmatory sequencing of RNAi, including identification of its unknown metabolites produced by *in vitro* hydrolysis.

EXPERIMENTAL

Chemicals and oligonucleotide samples

Complementary RNA strands of 21 nucleotides (nt) length, 'upper' strand 5'-UCG UCA AGC GAU UAC AAG GTT-3' (MW_{monoiso} 6689.954) and 'lower' strand 5'-CCU UGU AAU CGC UUG ACG ATT-3' (MW_{monoiso} 6603.894) were purchased from Integrated DNA Technologies (Coralville, IA, USA). The 10 nt long RNA isobaric sequences 5'-AGU UGU CUTT-3', 5'-UCU GUU GATT-3', and 5'-CUA UGU GUTT-3' were purchased from the same vendor. DNA standard (MassPrep OST Standard) was purchased from Waters (Milford, MA, USA); the sample contained a ladder of oligodeoxythymidines with dominant peaks of 15T, 20T, 25T, 30T, and 35 T oligonucleotides.

The sample concentrations used for the UPLC/MS analysis were 30 pmol/ μ L, and 5 μ L was injected for the routine MS/MS analysis. Digest of the upper strand RNAi was prepared using 1 mU phosphodiesterase II from bovine spleen (Sigma, St. Louis, MO, USA) added to 22.1 nmol of LC-purified RNA sample in 400 μ L of 0.1 M triethylammonium acetate (TEAA) and digested for 1 h at room temperature. Digestion was stopped by heating the sample to 70°C for 10 min and sample was diluted with 0.1 M TEAA to 27 pmol/ μ L concentration. All oligonucleotides were reconstituted in 0.1 M TEAA prior to LC/MS analysis.

LC setup

An ACQUITY UPLC OST C₁₈ column (1.7 μ m, 2.1 \times 50 mm) and an ACQUITY UPLC system (Waters, Milford, MA, USA) were utilized for separation. The column temperature was set to 60°C and the flow rate was 0.2 mL/min. Mobile phase A consisted of an aqueous buffer (pH 7.9) containing 15 mM triethylamine (TEA) and 400 mM hexafluoroisopropanol (HFIP), both purchased from Fluka (St. Louis, MO, USA). Mobile phase B was prepared by adding 50% methanol (v/v) to mobile phase A. For 21 nt RNA separation, the gradient was 20 to 40% B in 10 min. Water was used as weak and strong solvents for needle and injector washes.

Mass spectrometry

The SYNAPT HDMS (Waters, Manchester, UK) Q-TOF mass spectrometer was controlled by MassLynx 4.1 software. Both UV and MS detectors were used in series. The ACQUITY UPLC photo-diode array detector was connected to the electrospray inlet using narrow and short silica capillary tubing (75 μ m \times 70 cm) in order to decrease post-UV peak broadening prior to the MS detection. MS parameters were set as following: capillary 2.7 kV, sampling cone 31 V, extraction cone 3 V, source temperature 120°C, desolvation temperature 300°C, desolvation gas flow 500 L/h, trap collisional energy 6 V, transfer collisional energy 4 V. Mass resolution was \sim 9000 in 'V' mode (FWHM). All data were collected in negative ion mode.

External calibration in negative ion mode was performed with CsI. LockMass calibration was applied using a solution

of 10 mg/mL CsI (water/isopropanol, 1:1) sprayed at a flow rate of 5 μ L/min, 1 s scan time, 30 s frequency, set mass 1685.765 m/z (Cs_6I_7^-). MaxEnt-1 software was applied for mass deconvolution of the MS data; MaxEnt-3 was used for deconvolution of MS/MS and MS^E spectra prior to sequence data analysis.

Collision-induced dissociation of selected m/z ions was performed with a collisional energy ramp from 25 to 55 V. The extent of MS/MS fragmentation was manipulated by selecting an appropriate charge state (−3 to −6) and varying the collisional energy ramp. Product ion spectra were acquired over the range of 500 to 7000 m/z at a rate of 1 scan/s.

MS^E was performed by alternating low (6 V) and elevated collisional energies in the collision cell. The spectral acquisition was 1.0 s with 0.02 s interscan delay. Elevated collisional energy was a function of the oligonucleotide of interest: on average, an energy ramp from 35 to 55 V was applied during the acquisition cycle.

RESULTS AND DISCUSSION

UPLC/MS analysis of RNA

UPLC separation of crude synthetic 21 nt RNA is shown in Fig. 1. The full-length target RNA oligonucleotide was well resolved from its failure sequences. In order to verify the identity of the 5'-truncated 20 nt oligomer (expected N-1 impurity), the sample was spiked with the corresponding 20 nt oligonucleotide (data not shown). This experiment confirmed the correct assignment of the (N−1) peak as a failed sequence (loss of uridine, peak at 8.73 min).

Mobile phases containing ion-pairing buffers TEA, HFIP, and methanol are compatible with ESI.^{6,11} Formation of TEA

adducts (101, 202 Da) in MS spectra along with sodium and potassium adducts is common (Fig. 1, inset). The level of adduction was acceptable and did not obscure mass spectral interpretation. MS desolvation conditions, including desolvation temperature, cone gas flow, and cone voltage, were optimized in order to achieve maximum declustering of TEA adducts. Neutral iron adducts could be practically eliminated by replacing the stainless steel mobile phase filters with titanium ones (data not shown).

MS and MS/MS analysis of RNA

Various synthetic byproducts were observed in the UPLC/MS chromatogram of 21 nt RNA (Fig. 1). The dominant peaks corresponded to 5'-truncated failed sequences and were assigned based on their accurate masses. A large portion of the target oligonucleotide sequence could be elucidated from the array of its 5'-truncated products. Another impurity was tentatively assigned as the (N+1) product containing an additional uridine mononucleotide (22 nt product). The assignments were supported by the oligonucleotide LC retention pattern.

Careful inspection of Fig. 1 revealed the presence of additional minor peaks. The exact masses of these impurities suggested that they are likely failed sequences related to the original RNA with occasional deletions in the sequence. Complete interpretation of these species based exclusively on the exact mass is problematic and time-consuming.

UPLC/MS/MS sequencing was first evaluated using 20–35 nt synthetic oligonucleotides. As a general rule, a collisional energy ramp between 25 and 55 V was suitable to generate structurally useful fragments of 20–35 nt oligonucleotides and to obtain a full sequence coverage. In order to achieve reasonable S/N of the MS/MS spectra, at

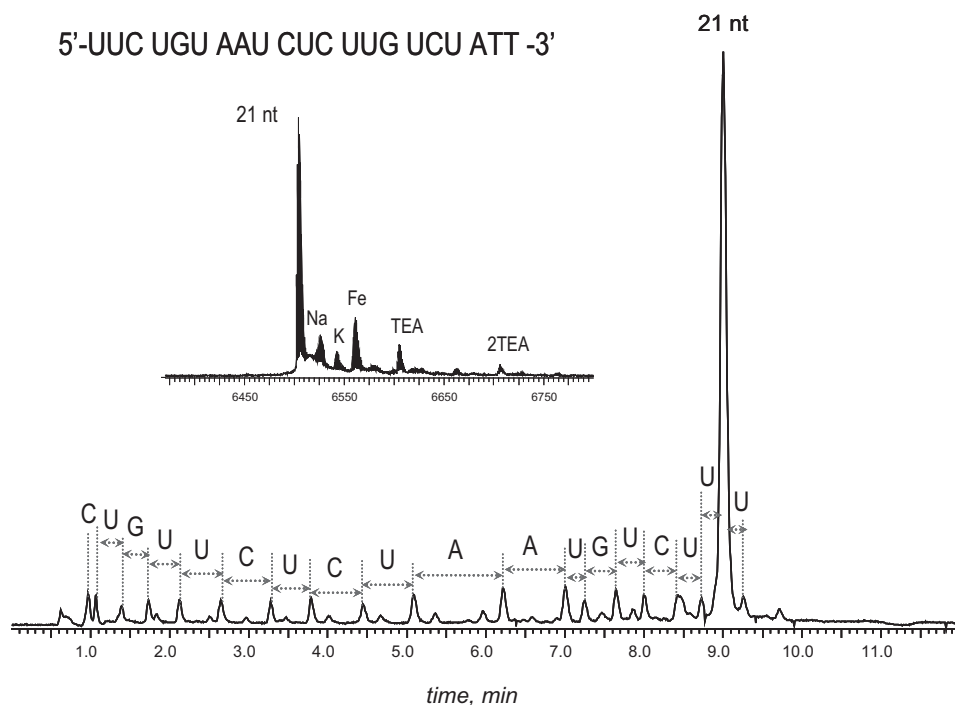


Figure 1. UPLC/MS of 21 nt RNA, demonstrating separation of failed sequences and by-products.

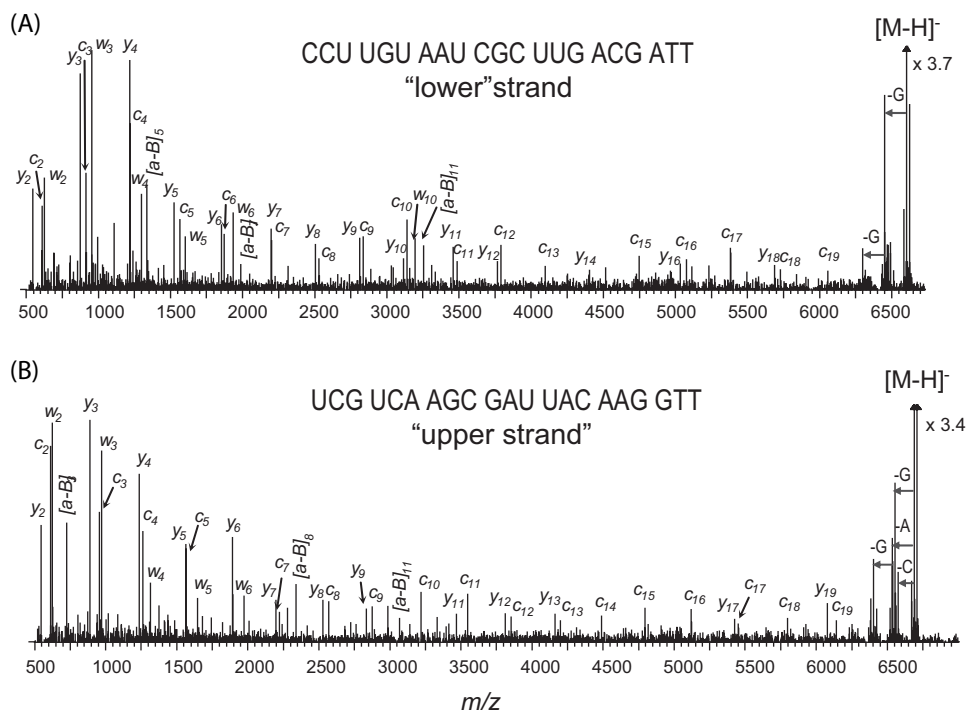


Figure 2. MaxEnt-3 deconvoluted MS/MS spectra of 21 nt RNA (A) 'lower' (E_{col} ramp 25 V to 45 V) and (B) 'upper' (E_{col} ramp 25 V to 40 V) RNA strands.

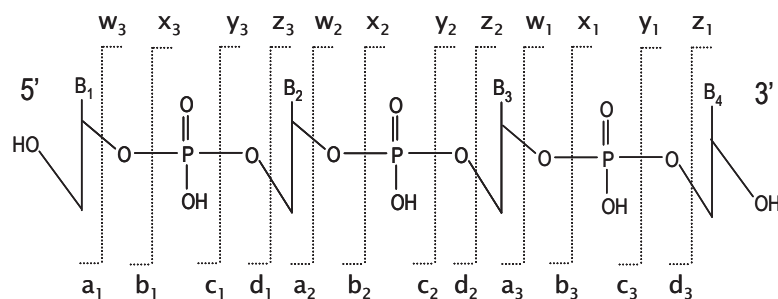
least 10–40 pmol injections of the oligonucleotide on-column was required.

The choice of the precursor ion affected the fragmentation: higher charge states required lower collision energies to yield useful MS/MS spectra. The charge state (-4) was commonly selected for MS/MS fragmentation since it gives dominant signals. Lower charge states reportedly resulted in higher extent of the formation of the neutral base losses that do not contribute to informative ion fragmentation.³² We have evaluated MS/MS fragmentation of the higher charge states of RNA of higher masses up to $[M-14H]^{14-}$ (data not shown). Their fragmentation required a significantly lower collisional energy, yet provided structurally relevant fragment ions. No charge state reduction was observed, as is often the case with ion trap instruments due to charge base losses.³³ While fragmenting higher charge states is a possibility, the selection of less dominant precursors decreased the S/N of the MS/MS spectra. In addition, a rich distribution of the high charge states in the MS/MS spectra complicated data interpretation.

In order to demonstrate sequencing of RNA oligonucleotides of typical length used as therapeutic compounds we

selected the 'lower' and 'upper' strands of the 21 nt RNA sample. In both cases, the $[M-4H]^{4-}$ precursor ion was subjected to MS/MS fragmentation (Fig. 2). The resulting spectra contained multiply charged fragments covering the m/z range up to ~ 2000 Da. In order to simplify data interpretation we have adopted modified MaxEnt-3 software for MS/MS spectra deconvolution. MaxEnt-3 converts all the ions in the spectra into the (-1) charged state and sums all the isotopes into a monoisotopic ion signal. In general, the longer oligonucleotides are more difficult to sequence with full sequence coverage. As shown here by an example of two different 21 nt RNA oligonucleotides, the larger MS/MS fragments are less abundant in the spectrum than the lower molecular weight ones (Fig. 2).

In agreement with the published reports,³⁴ predominant observed characteristic fragments of RNA molecules were complementary c - and y -ions and less intense $[a-B]$ and w -ions (Scheme 1 and Fig. 2). The spectrum also showed the internal fragments produced by double cleavage of a backbone, as well as a neutral loss of cytosine from c -ions (c_3 , c_4 , c_5 of the lower strand and c_5 of the upper strand). These fragments do not provide structurally useful information.



Scheme 1. Fragmentation nomenclature of oligonucleotides adopted from McLuckey *et al.*¹⁹

MS/MS analysis of the 'upper' 21 nt RNA strand via ramping of the collisional energy from 25 V to 40 V resulted in efficient fragmentation coverage allowing confirmation of the sequence of the investigated oligonucleotide (Fig. 2). A few peaks corresponding to A, G, and C gas-phase nucleobase losses of the deprotonated molecular ion were also detected.

It is worth repeating that the fragment ions of the MS/MS spectra of $[M-4H]^{4-}$ 21 nt RNA were present in several charge states. Deconvolution of multiply charged ions to singly charged ions using MaxEnt-3 reduced the MS/MS spectral complexity and simplified the data interpretation. Both 'upper' and 'lower' RNA strands could be deciphered using the combination of the most intense 3'-sequence *y*-ions and 5'-sequence *c*-ions. Both strands showed a representation of sequential nucleobase losses from *c*₂ to *c*₁₉ (Fig. 2).

The MS/MS approach offers a more detailed knowledge of the nature of oligonucleotide samples than simple LC or LC/MS analysis, allowing for confirmatory and possible *de novo* sequencing and characterization of unknown modifications. LockMass calibration resulted in a mass accuracy of 12.4 ppm for the lower strand and 10.6 ppm for the upper strand (Table 1), which is highly useful for fragment ion assignment.

MS^E analysis of native oligonucleotides

To compare the results of the MS/MS and MS^E approaches, a simple DNA sample consisting of polythymine oligonucleotides of various lengths (15T, 20T, 25T, 30T, 35T) was subjected to both fragmentation methods using an identical collisional energy profile (Fig. 3). Such a homogeneous sample was chosen for clearer illustration of differences

Table 1. Observed and theoretical fragments of 'lower' (E_{col} ramp 25 V to 45 V) and 'upper' strands (E_{col} ramp 25 to 40 V) of 21 nt RNA generated by UPLC/MS/MS (see Fig. 1)

Lower strand			Upper strand		
ion fragment	observed <i>m/z</i>	theoretical <i>m/z</i>	ion fragment	observed <i>m/z</i>	theoretical <i>m/z</i>
y ₂	545.13	545.13	y ₂	545.13	545.13
c ₂	609.07	609.08	c ₂	610.06	610.06
w ₂	625.09	625.10	w ₂	625.10	625.10
y ₃	874.18	874.18	a ₃ -B	724.10	724.09
c ₃	915.10	915.10	y ₃	890.19	890.18
w ₃	954.15	954.15	c ₃	955.12	955.11
y ₄	1219.22	1219.23	w ₃	970.16	970.14
c ₄	1221.12	1221.13	y ₄	1235.24	1235.22
w ₄	1299.19	1299.20	c ₄	1261.15	1261.13
a ₅ -B	1335.16	1335.16	w ₄	1315.21	1315.19
y ₅	1524.26	1524.27	a ₅ -B	1375.18	1375.16
c ₅	1566.17	1566.17	y ₅	1564.30	1564.28
w ₅	1604.23	1604.24	c ₅	1566.19	1566.17
y ₆	1853.31	1853.32	w ₅	1644.26	1644.24
c ₆	1872.16	1872.20	a ₆ -B	1680.22	1680.21
w ₆	1933.28	1933.29	y ₆	1893.34	1893.33
a ₇ -B	1986.20	1986.23	w ₆	1973.30	1973.30
y ₇	2198.33	2198.37	y ₇	2198.38	2198.37
c ₇	2201.22	2201.25	c ₇	2224.28	2224.28
a ₈ -B	2315.24	2315.28	w ₇	2278.35	2278.34
y ₈	2504.35	2504.40	a ₈ -B	2338.31	2338.31
c ₈	2530.26	2530.30	y ₈	2527.44	2527.42
y ₉	2810.38	2810.42	c ₈	2569.36	2569.33
c ₉	2836.29	2836.33	y ₉	2833.47	2833.45
w ₉	2890.34	2890.39	c ₉	2874.38	2874.37
y ₁₀	3115.42	3115.46	a ₁₁ -B	3333.49	3333.45
c ₁₀	3141.32	3141.37	c ₁₀	3219.44	3219.41
w ₁₀	3195.37	3195.43	y ₁₁	3468.56	3468.53
a ₁₁ -B	3255.35	3255.40	c ₁₁	3548.51	3548.47
y ₁₁	3460.47	3460.51	y ₁₂	3813.59	3813.57
c ₁₁	3486.39	3486.42	c ₁₂	3854.53	3854.49
y ₁₂	3765.52	3765.55	y ₁₃	4118.71	4118.61
c ₁₂	3791.45	3791.46	c ₁₃	4160.58	4160.52
c ₁₃	4097.46	4097.48	c ₁₄	4489.61	4489.57
y ₁₄	4400.60	4400.63	c ₁₅	4794.56	4794.61
c ₁₅	4748.45	4748.56	c ₁₆	5123.71	5123.66
y ₁₆	5035.65	5035.71	y ₁₇	5426.83	5426.81
c ₁₆	5077.50	5077.61	c ₁₇	5452.77	5452.72
w ₁₆	5115.55	5115.67	y ₁₉	6077.89	6077.88
y ₁₇	5380.77	5380.75	c ₁₉	6142.88	6142.81
y ₁₈	5686.77	5686.78			
y ₁₉	5992.71	5992.80			
c ₁₉	6056.66	6056.75			
RMS = 12.4 ppm			RMS = 10.6 ppm		

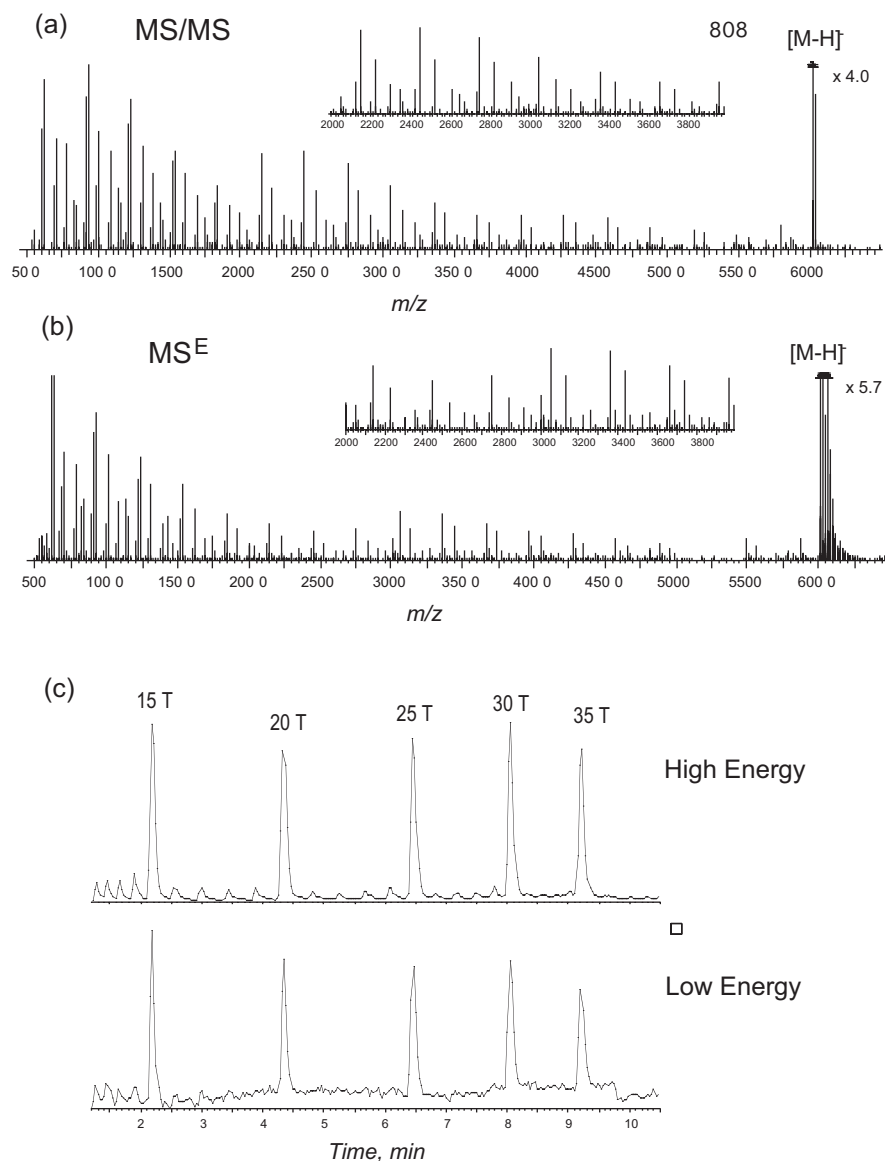


Figure 3. (A) MaxEnt-3 deconvoluted MS/MS spectra of (-4) charge state of 20 nt oligodeoxythymidine vs. (B) UPLC/MS^E of DNA mixture (E_{col} ramp 30 to 55 V). (C) High/low-energy UPLC/TIC of poly-T mixture.

between MS vs. MS^E spectra. MS/MS fragmentation was applied to the (-4) charge state of the 20 T oligomer. The MS^E method allowed collisional fragmentation of the ions of all charge states and the whole range of the oligonucleotide sizes to be induced during a single UPLC/MS analysis. A full set of the characteristic $[a-B]$ and w -ions, typical for DNA fragmentation, was observed in both MS/MS and MS^E modes, reflecting consecutive thymine losses from T₂₀ to T₁. Due to the advantage of the high chromatographic resolution of UPLC, all peaks were resolved, which eased the interpretation of MS^E data by aligning precursor and fragment ion signals to each peak.

When using otherwise identical conditions, the signal intensity of the ion fragments produced by MS^E was 2-fold greater compared to the signal resulting from MS/MS fragmentation (see insets in Fig. 3). Low-energy MS^E spectra provided the masses of the precursor ions, while high-energy spectra were used for sequence verification.

To demonstrate the applicability of the UPLC/MS^E approach, three 10 nt RNA isomers – AGU UGU CUTT, 'reversed' sequence UCU GUU GATT, and 'scrambled' sequence CUA UGU GUTT – were mixed in a vial in approximate equimolar ratios. The LC conditions were adjusted in order to separate the RNA isomers using a shallow gradient (Fig. 4(A)). The elution order of chromatographic peaks could not be verified by MS detection, because all three RNA oligonucleotides have the same mass. To confirm their identity we performed an LC/MS^E sequencing experiment. Deconvoluted MS^E mass spectra showed consecutive losses of each nucleotide based on y -fragment ions, providing verification of the sequence of closely eluting RNA isomers (Fig. 4(B)).

To further demonstrate the capabilities of the UPLC/MS^E method, the developed method was applied for the analysis of a more complex RNA sample. Chromatographic separation of a 21 nt RNA phosphodiesterase II digest revealed

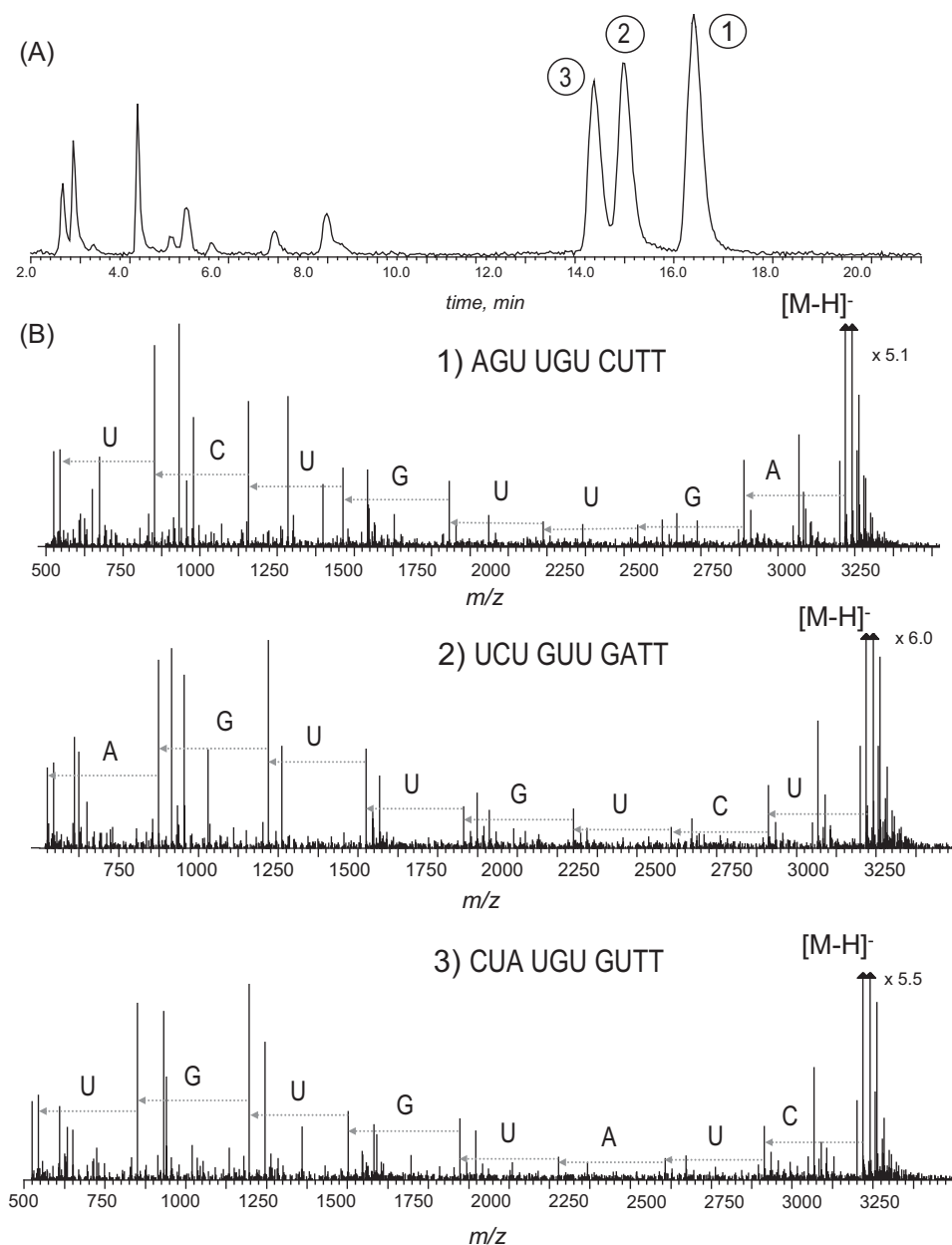


Figure 4. (A) TIC of UPLC/MS^E of 10 nt RNA isomers and (B) MaxEnt-3 deconvoluted MS^E spectra of the separated isomers (only y-ion series were annotated).

numerous peaks of different intensities. Judging by accurate MS data, the cleavage was more complex than just sequential losses of the nucleotides from both oligonucleotide termini (Fig. 5). Interpretation of the MS^E spectra of each peak revealed a highly heterogeneous mixture. Some peaks represented truncations from either the 3'- or the 5'-end, or from both. Both termini carried a hydroxyl or phosphate group and some species had a 3'-terminus with a cyclic phosphate as well. Two closely eluting oligonucleotides were found to be the compositional isomer that differ by the position of uridine in the sequence (Fig. 5(B)).

MS^E analysis of modified oligonucleotides

Tandem mass spectrometry methods were further applied for structural characterization of chemically modified oligonucleotides using a set of 8 nt RNA oligonucleotides

with the same general sequence: native CAGCGCAC, 2-O-methylated CAG(mC)G(mC)AC, 2-fluorinated (CAG(fC)G(fC)AG, phosphorothioate CpApGpCpGpCpApC, and locked-nucleic acid (LNA) CAGCGCAC (Fig. 6). The high S/N of MS^E spectra helped to reveal low intensity fragments and to assess the influence of each type of modification on the extent of fragmentation.

The effect of the presence of a ribose 2'-OH group on the dissociation mechanism has been studied and the differences in fragmentation stability of RNA vs. DNA molecules were reported.^{22,33–35} Other modifications of the ribose sugar affecting the N-glycosidic bond were shown to drastically alter the traditional fragmentation pattern and relative intensity of the characteristic ions.²⁴

The acquired data (Fig. 6(B)) demonstrated that substitution of the ribose hydroxyl group with fluorine significantly reduced the extent of fragmentation at the fluorinated

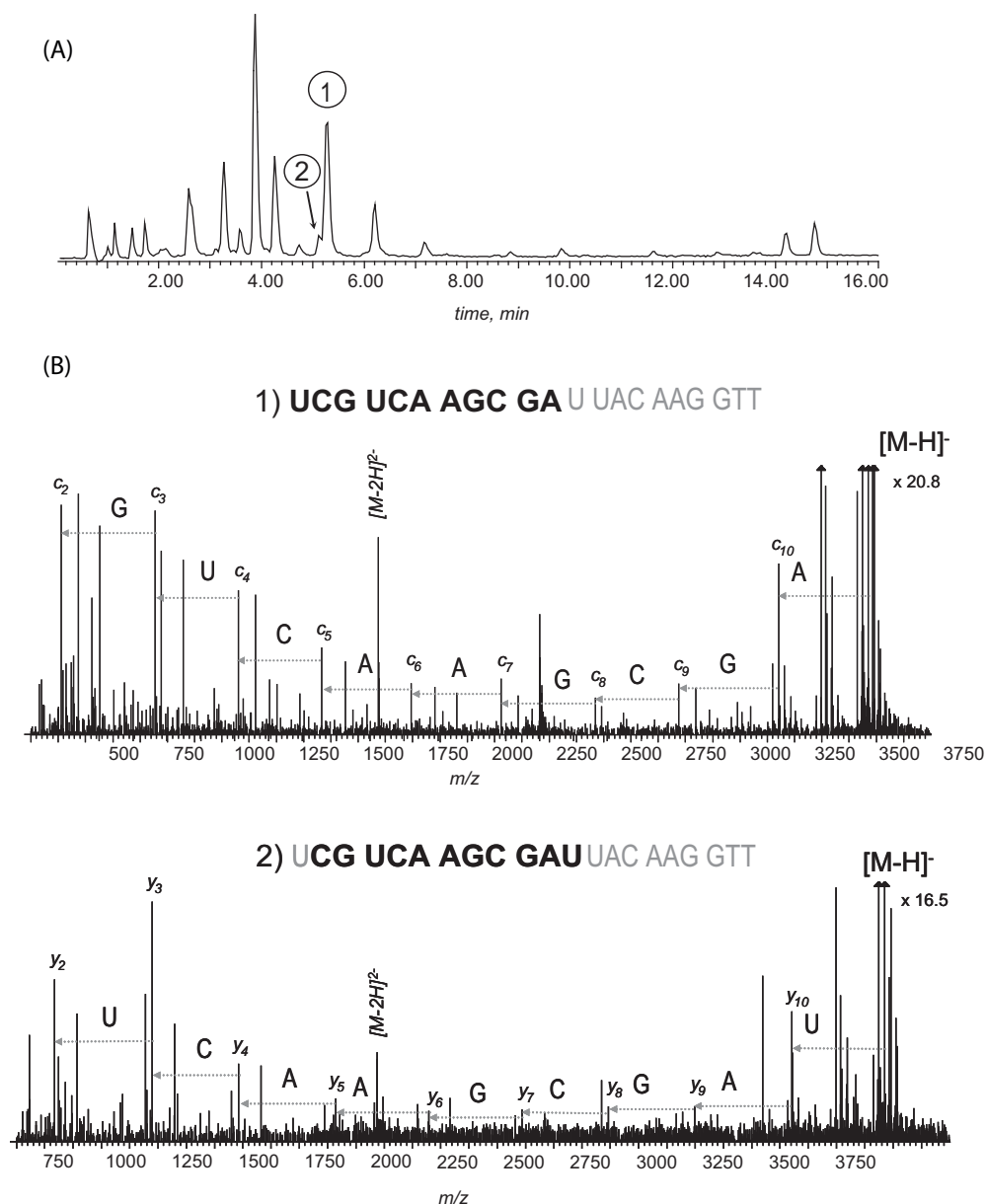


Figure 5. (A) TIC of UPLC/MS^E of RNA phosphodiesterase II digest and (B) MaxEnt-3 deconvoluted MS^E spectra of separated isomers (E_{col} ramp 25 to 50 V).

nucleotide sites. With similar quantities injected on-column (~150 pmol), no *w*₄-ions were present at the fluorinated site. Substitution of the 2'-hydroxyl proton with a methyl group with a 2'-hydroxyl oxygen still present led to a decrease in the abundance of *c*-, *y*-, and [*a*-B]₆ ions at the methylation sites to a lesser extent (Fig. 6(C)).

Alteration of the backbone structure greatly affected the traditional RNA fragmentation pathway, as was demonstrated on examples of phosphorothioate RNA and LNA molecules (Figs. 6(D) and 6(E)). Both oligonucleotides readily fragmented with no predominant ions being formed. Both *b*- and *x*-ion series become prominent in the spectra of the phosphorothioate. Unlike the rest of the modified samples fragmented using a E_{col} ramp from 30 to 55 V, the collisional energy for the phosphorothioate was lowered to the 25 to 45 V ramp to avoid major nucleobase neutral losses. However, guanine depurination and cytosine depyrimidination of the *c*- and *x*-ion series could not be entirely eliminated, whereas

the other samples exhibited a nucleobase loss to a much lesser extent.

MS^E of 8 nt LNA generated all the theoretically possible fragment ions without exhibiting the preferred fragmentation pathway (Fig. 6(E)). The only characteristic fragment series that were not observed were [*a*-B] ions, due to the structural restriction of LNA ribose. A molecular weight shift of 56 Da compared to the theoretical MW of this synthetic sample was detected in the UPLC/MS chromatogram, yet it did not provide sufficient information on the type of modification or its location within a sequence. This unknown sequence alteration was deciphered by means of MS^E data. The delta mass of gas-phase depyrimidination of cytosine was found to be (+14 Da), and the ion series in the CAGCGCAC sequence were shifted at each cytosine linkage. Therefore, the mass difference was assumed to be the methyl group on the cytosine nucleobase apparently introduced during oligonucleotide synthesis.

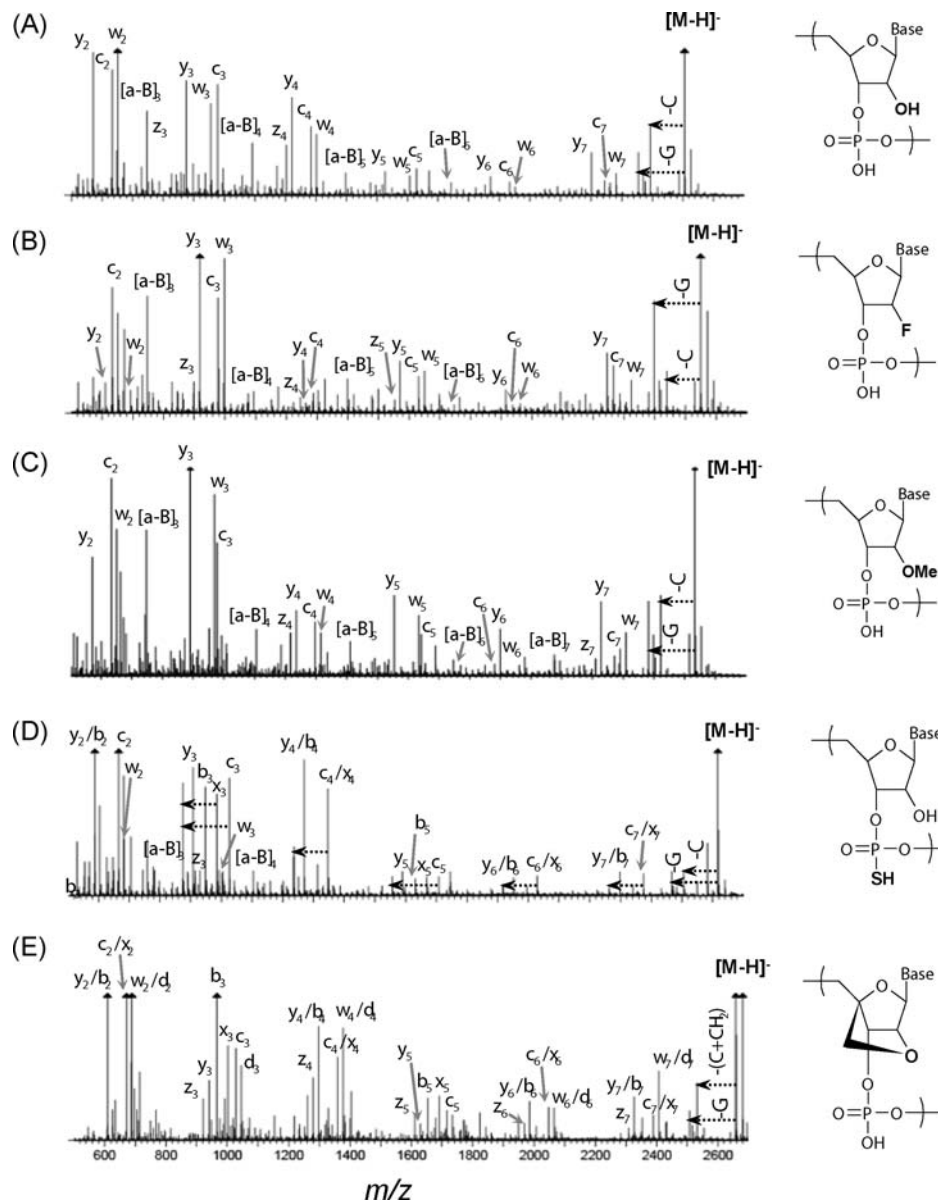


Figure 6. MaxEnt-3 deconvoluted MS^E spectra (E_{col} ramp 30 to 55 V) of (A) native RNA (CAGCGCAC), (B) 2-fluorinated RNA (CAG(fC)G(fC)AG), (C) 2-O-methylated RNA (CAG(mC)G(mC)AC), and (D) phosphorothioate RNA (CpApGpCpGpCpApC) (E_{col} ramp 25 to 45 V), (E) LNA (CAGCGCAC). The dotted arrows represent neutral base losses.

CONCLUSIONS

It was demonstrated that UPLC/MS/MS and UPLC/MS^E methods are well suited for a purity check of oligonucleotides and monitoring of the efficiency of the chemical synthesis. A high degree of separation of the full-size RNA molecules from their truncated products by the fast UPLC method fulfills the needs of RNA metabolism studies. The advanced UPLC resolution enabled the fast MS^E approach to be applied for the whole range of species separated on the chromatogram, since their resulting ion fragments are time-aligned with well-resolved corresponding precursor ions. UPLC/MS^E analysis of very complex oligonucleotide samples became possible, such as the products of spontaneous oligonucleotide hydrolysis. Achieved levels of MS/MS and MS^E fragmentations were sufficient for the

confirmatory sequencing of up to 21 nt RNA molecules. The high mass accuracy and resolution of QTOF MS is necessary for unambiguous assignments of the fragment ions. A relative increase in spectral S/N from MS^E spectra of DNA ions was observed compared to the MS/MS spectra, as higher charge state ions can be over-fragmented. The MS/MS method requires fine optimization of the collisional energy for each precursor ion in order to obtain the entire range of characteristic ions of reasonable intensities. The MS^E approach is universal, fast, and suitable for preliminary scanning of highly complex LC/MS chromatograms.

The MS^E method was applied for sequence conformations of LNA and modified RNA. Results showed that different fragmentation mechanisms took place affecting the pattern

of the characteristic fragments for the modified oligonucleotide species.

The UPLC/MS analysis and sequencing of the complex mixture, such as enzymatically digested RNA, revealed the termini modifications and failed sequences truncated from either the 3'- or the 5'-end, or both. The UPLC/MS method fulfills the FDA requirements on oligonucleotide therapeutics sequence conformation. Performing *de novo* sequencing and profiling of impurities for modified RNA sequences becomes possible using this approach. Unlike the conventional time-consuming and labor-intensive solution-based methods, MS/MS and MS^E approaches are more amenable to the analysis of chemically modified oligonucleotides that are difficult to sequence by enzymatic methods.

Acknowledgements

We thank Dr. Sean M. McCarthy for providing the samples of RNA digest.

REFERENCES

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. *Nature* 1998; **391**: 806.
2. Agami R. *Curr. Opin. Chem. Biol.* 2002; **6**: 829.
3. Manoharan M. *Curr. Opin. Chem. Biol.* 2004; **8**: 570.
4. Gao H, Liu Y, Rumley M, Yuan H, Mao B. *Rapid Commun. Mass Spectrom.* 2009; **23**: 3423.
5. Beverly M, Hartsough K, Machemer L. *Rapid Commun. Mass Spectrom.* 2005; **19**: 1675.
6. Gilar M, Fountain KJ, Budman Y, Holyoke JL, Davoudi H, Gebler JC. *Oligonucleotides* 2003; **13**: 229.
7. Fountain KJ, Gilar M, Gebler JC. *Rapid Commun. Mass Spectrom.* 2003; **17**: 646.
8. Huber CG, Krajete A. *Anal. Chem.* 1999; **71**: 3730.
9. Lin ZJ, Li W, Dai G. *J. Pharm. Biomed. Anal.* 2007; **44**: 330.
10. Gilar M, Bouvier ESP. *J. Chromatogr. A* 2000; **890**: 167.
11. Appfel A, Chakel JA, Fischer S, Lichtenwalter K, Hancock WS. *Anal. Chem.* 1997; **69**: 1320.
12. Gilar M. *Anal. Biochem.* 2001; **298**: 196.
13. McCarthy SM, Gilar M, Gebler J. *Anal. Biochem.* 2009; **390**: 181.
14. Farand J, Gosselin F. *Anal. Chem.* 2009; **81**: 3723.
15. Froim D, Hopkins CE, Belenky A, Cohen AS. *Nucleic Acids Res.* 1997; **25**: 4219.
16. Schuette JM, Piele U, Maleknia SD, Srivatsa GS, Cole DL, Moser HE, Afeyan NB. *J. Pharm. Biomed. Anal.* 1995; **13**: 1195.
17. Wu H, Chan C, Aboleneen H. *Anal. Biochem.* 1998; **263**: 129.
18. Hettich RL, Stemmler EA. *Rapid Commun. Mass Spectrom.* 1996; **10**: 321.
19. McLuckey SA, Van Berkel GJ, Glish GL. *J. Am. Soc. Mass Spectrom.* 1992; **3**: 60.
20. McLuckey SA, Habibi-Goudarzi S. *J. Am. Chem. Soc.* 1993; **115**: 12085.
21. Kirpekar F, Krogh TN. *Rapid Commun. Mass Spectrom.* 2001; **15**: 8.
22. Schurch S, Bernal-Mendez E, Leumann CJ. *J. Am. Soc. Mass Spectrom.* 2002; **13**: 936.
23. Andersen TE, Kirpekar F, Haselmann KF. *J. Am. Soc. Mass Spectrom.* 2006; **17**: 1353.
24. Tromp JM, Schurch S. *J. Am. Soc. Mass Spectrom.* 2005; **16**: 1262.
25. Flora JW, Muddiman DC. *J. Am. Soc. Mass Spectrom.* 2001; **12**: 805.
26. Rozenski J, McCloskey JA. *J. Am. Soc. Mass Spectrom.* 2002; **13**: 200.
27. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li GZ, McKenna T, Nold MJ, Richardson K, Young P, Geronimos S. *Anal. Chem.* 2005; **77**: 2187.
28. Silva JC, Denny R, Dorschel C, Gorenstein MV, Li GZ, Richardson K, Wall D, Geromanos SJ. *Mol. Cell. Proteomics* 2006; **5**: 589.
29. Xie H, Gilar M, Gebler JC. *Anal. Chem.* 2009; **81**: 5699.
30. Chakraborty AB, Berger SJ, Gebler JC. *Rapid Commun. Mass Spectrom.* 2007; **21**: 730.
31. Plumb RS, Johnson KA, Rainville P, Smith BW, Wilson ID, Castro-Perez JM, Nicholson JK. *Rapid Commun. Mass Spectrom.* 2006; **20**: 1989.
32. Nordhoff E, Kirpekar F, Roepstorff P. *Mass Spectrom. Rev.* 1996; **15**: 67.
33. Huang TY, Kharlamova A, Liu J, McLuckey SA. *J. Am. Soc. Mass Spectrom.* 2008; **19**: 1832.
34. Tang W, Zhu L, Smith LM. *Anal. Chem.* 1997; **69**: 302.
35. Nordhoff E, Cramer R, Karas M, Hillenkamp F, Kirpekar F, Kristiansen K, Roepstorff P. *Nucleic Acids Res.* 1993; **21**: 3347.