International Journal of Mass Spectrometry 312 (2012) 155–162

Contents lists available at ScienceDirect

International Journal of Mass Spectrometry

journal homepage: www.elsevier.com/locate/ijms

Higher-order structure of nucleic acids in the gas phase: Top-down analysis of base-pairing interactions

D. Fabris[a,*], K.A. Kellersberger[b], J.A. Wilhide[c]

[a] The RNA Institute, University at Albany, United States
[b] Bruker Daltonics, United States
[c] University of Maryland Baltimore County, United States

ARTICLE INFO

Article history:
Available online 3 August 2011

Keywords:
Top-down
ECD
EDD
IRMPD
Nucleic acids
Base-pairing
Higher-order structure

ABSTRACT

Non-ergodic as well as ergodic activation methods are capable of maintaining the integrity of base pairs during the top-down analysis of nucleic acids. Here, we investigate the significance of this characteristic in the investigation of higher-order structures of increasing complexity. We show that cognate fragments produced by typical backbone cleavages may not be always detected as separate sequence ions, but rather as individual products that remain associated through mutual pairing contacts. This effect translates into unintended masking of cleavage events that take place in double-stranded regions, thus leading to the preferential detection of fragments originating from unpaired regions. Such effect is determined by the stability of the weak non-covalent association between complementary stretches, which is affected by base composition, length of the double-stranded structure, and charge of the precursor ion selected for analysis. Although such effect may prevent the achievement of full sequence coverage for primary structure determination, it may provide the key to correctly differentiate double- versus single-stranded regions, in what could be defined as gas-phase footprinting experiments. In light of the critical role played by base pairs in defining the higher-order structure of nucleic acids, these approaches will be expected to support an increased utilization of mass spectrometry for the investigation of nucleic acid structure and dynamics.

1. Introduction

The higher-order structure of nucleic acids is largely defined by base pairing interactions. The annealing of contiguous sequences determines elements of secondary structure, whereas that of distal stretches delineates their tertiary structure. For this reason, any technology capable of revealing the identity and correct position of paired nucleotides could in principle provide a platform for solving the 3D structure of these biopolymers. The observation that soft ionization techniques can preserve base pairing during MS analysis [1–4] has opened the door for the possible application of gas-phase methods to the structural elucidation of nucleic acids. Taking advantage of this intrinsic capability, we are investigating whether activation techniques developed to obtain the primary structure (i.e., sequence) of nucleic acids [5–7] could be applied to discriminate paired versus unpaired nucleotides, thus enabling the direct analysis of their higher-order structure.

Non-ergodic activation methods employed for the top-down analysis of biopolymers have been proven capable of inducing fragmentation of the covalent backbone, while simultaneously retaining the weak non-covalent interactions that stabilize structure folding and ligand binding. During electron capture dissociation (ECD) [8] of polypeptides, which involves the recombination of low-energy (<0.2 eV) electrons with multiply protonated precursor ions, the preservation of intramolecular interactions can produce fragmentation patterns consistent with the spatial organization of folded species [9,10]. In similar fashion, intermolecular contacts can be retained upon electron transfer dissociation (ETD) [11], a related activation method that relies on anionic vehicles for delivering electrons to multi-cationic precursors. This property has enabled the identification of interface residues involved in subunit association [12]. The stabilizing effects of intramolecular interactions have been invoked to explain the inability of obtaining full sequence information from oligonucleotide samples by electron

Abbreviations: 3D, three dimensional; MS, mass spectrometry; ECD, electron capture dissociation; ETD, electron transfer dissociation; CID, collision-induced dissociation; BIRD, blackbody infrared radiative dissociation; IRMPD, infrared multiphoton dissociation; DNA, deoxyribonucleic acid; SORI, sustained off-resonance irradiation; RNA, ribonucleic acid; HIV-1, human immunodeficiency virus type 1; FTICR, Fourier transform ion cyclotron resonance; FIV, feline immunodeficiency virus; UV, ultraviolet; SL3, stemloop 3.

* Corresponding author at: The RNA Institute, University at Albany (SUNY), Life Sciences Research Building, Room 1109, 1400 Washington Avenue, Albany, NY 12222, United States. Tel.: +1 518 437 4464; fax: +1 518 442 3462.
E-mail address: fabris@albany.edu (D. Fabris).

1387-3806/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2011.07.014

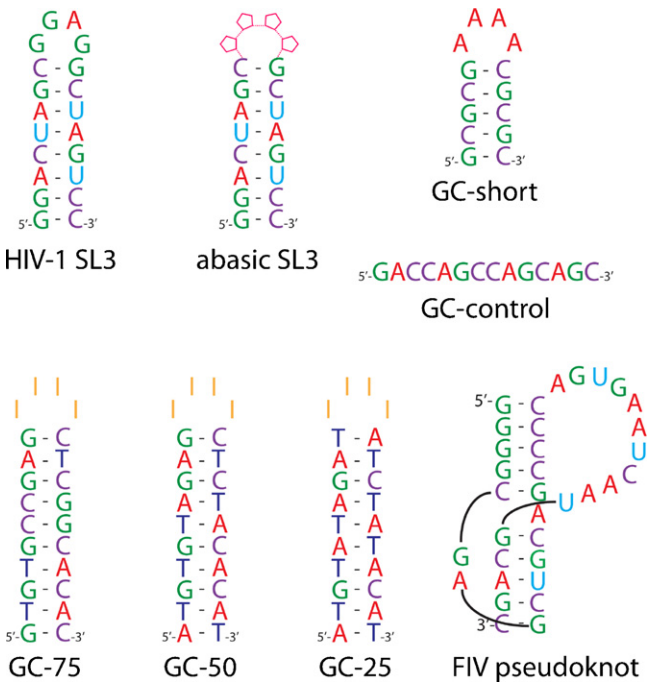
detachment dissociation (EDD) [13,14], a technique in which multi-anionic precursors are irradiated with more energetic electrons (>10 eV) [15]. Although the electronic processes associated with these activation methods are still debated, their common denominator consists of the formation of radical species, followed by immediate cleavage of adjacent covalent bonds. The non-ergodic nature and short time scale ($<10^{-12}$ s) of these events minimize the intramolecular redistribution of vibrational energy, which leaves local non-covalent bonds largely unaffected [9,10].

Ergodic activation methods operate on longer time scales, involve energy redistribution, and tend to disrupt weak non-covalent interactions. Submitted to collision-induced dissociation (CID) [16], double-stranded structures can undergo base pair disruption and strand separation. Duplex dissociation was followed as a function of collision energy to generate gas-phase melting curves that mimicked very closely thermal denaturation curves obtained in solution [17,18]. Based on the ability of trapped ions to absorb infrared photons from their surroundings, blackbody infrared radiative dissociation (BIRD) [19] was employed to study the stability of complementary and non-complementary DNA duplexes in a solvent-free environment [20]. Infrared multiphoton dissociation (IRMPD) [21], which can be accomplished by irradiating precursor ions with a CO_2 laser, was shown capable of inducing strand separation and ligand ejection from selected drug-DNA duplex complexes [22]. For each of these activation methods, dissociation of non-covalent interactions is typically accompanied, or overshadowed, by the concomitant cleavage of covalent bonds, which results in loss of nucleobases and fragmentation of the phosphodiester backbone. In most cases, the presence of relatively stable non-covalent interactions represents an obstacle to the observation of the complete fragment series that provide the sought-after sequence information [5–7]. For this reason, efforts have been dedicated to exploring the interplay between kinetic and thermodynamic factors that determine the prevalence of covalent versus non-covalent dissociation [23].

Capitalizing on the behavior of base pairing and ligand interactions upon gas-phase activation, MS-based technologies will be expected to play an increasing role in the investigation of the structure and dynamics of nucleic acids. In this direction, we have employed sustained off-resonance irradiation (SORI) [24] CID to characterize the unique structural features exhibited by DNA–RNA hybrid duplexes [25,26] and RNA stemloop structures [27,28] involved in viral replication. The observation that backbone fragmentation could be inhibited by the contact of non-covalent ligands [27] prompted us to propose the utilization of classic nucleic acid ligands as non-covalent structural probes [28]. This general strategy relies on firm knowledge of the binding profiles of such ligands and counts on the ability to locate their specific sites by top-down MS. We have also shown that SORI-CID can be effectively employed to interrogate the conformational state of RNA complexes, such as those involved in the dimerization and packaging of the HIV-1 genome [29,30]. In this case, isomeric dimers that could not be recognized from their identical molecular mass were readily distinguished on the basis of the different gas-phase stabilities afforded by their interstrand pairing [30]. We are now investigating the possibility of combining non-ergodic and ergodic techniques available in Fourier transform ion cyclotron resonance (FTICR) [31,32] MS to enable the differentiation of paired versus unpaired nucleotides, which could afford the experimental data necessary to elucidate higher-order structures of nucleic acids.

2. Materials and methods

All nucleic acid samples employed in the study (Scheme 1) were purchased from IDT (Coralville, IA) and desalted by ultrafiltration against 100 mM ammonium acetate (pH adjusted to 7.0). Each construct was heat-refolded to obtain the desired conformation by taking the solution to 95 °C for 5 min in a water-bath, then allowing it to return slowly to room temperature. Stock concentrations were determined by UV absorbance using molar extinction coefficients calculated from the respective sequences. Immediately before analysis, samples were diluted by addition of 100 mM ammonium acetate and a 10% volume of 2-propanol to obtain a final 1 μM concentration. Each analysis typically required 2 μL of sample solution.



Scheme 1. Nucleic acid constructs employed in the study. In abasic SL3, the four loop positions were replaced with deoxy-ribonucleotides devoid of nucleobase.

All experiments were carried out on a Bruker Daltonics (Billerica, MA) Apex Qe FTICR mass spectrometer equipped with a 12 T superconducting magnet and an Apollo II electrospray source that was modified with a resistively heated metal capillary. Analyses were performed in nanospray mode using non-coated quartz emitters that were held at approximately 1100 V relative to the capillary entrance. Selected precursor ions were initially isolated in the mass selective quadrupole situated at the front end of this hybrid instrument. They were then immediately transmitted to the FTICR cell to complete the different top-down experiments. ECD of positively charged precursors was accomplished by using electrons emitted from a hollow dispenser cathode heated by supplying a current of 1.5 A, and then accelerated into the cell by a potential of up to 6.5 V. In similar fashion, EDD was performed in negative ion mode by using electrons from the cathode operating at 1.7 A and up to 23 V acceleration. IRMPD was completed by shining laser pulses of up to 50 ms through a BaF_2 window in the back of the cell. The source was a Synrad (Mukilteo, WA) CO_2 laser producing infrared light of 1057 nm wavelength and up to 25 W power.

Experiments combing consecutive stages of activation were carried out by performing SORI-CID immediately after the initial IRMPD step, without intervening precursor ion isolation. Best results were obtained by applying irradiation frequencies that were 600–2000 Hz below that of the precursor. Mild activation regimes were obtained by applying the off-resonant pulse for about 250 ms using 26–31 dB attenuation of the maximum power output allowed by the hardware. For comparison purposes, typical experiments designed to obtain sequence information rely instead on pulses

generated with 17–20 dB attenuation of maximum power. Argon was used as the collision gas in pulsed bursts of 100–250 ms, which resulted in momentary increases from 1×10^{-11} to 1×10^{-7} mbar of the background pressure measured by the instrument gauge located underneath the ion optics. No attempt was made to determine the actual pressure within the FTICR cell. Typical product ion spectra consisted of the average of 50–100 acquisitions.

Data processing and interpretation were respectively carried out by using DataAnalysis v4.0 (Bruker Daltonics) and Mongo Oligo Mass Calculator v2.06 [33], which is hosted at http://rna-mdb.cas.albany.edu/RNAmods/masspec/mongo.htm by the RNA Institute of the University at Albany. For the evaluation of loop versus stem fragmentation, signals corresponding to $\mathbf{I}^-$, $\mathbf{II}^-$ and $\mathbf{III}^-$ (491.0, 805.1, and 1119.1 *m*/*z*, respectively) were selected to represent the incidence of loop cleavages, whereas those corresponding to $\boldsymbol{[a_3\text{-}B]}^-$, $\boldsymbol{w_3}^-$, $\boldsymbol{[a_5\text{-}B]}^-$, and $\boldsymbol{w_5}^-$ (714.1, 923.2, 1347.2 or 1363.2, and 1510.3 or 1525.3 or 1540.3 *m*/*z*, respectively) were employed for stem events. The relative intensities of all the isotopic species detected for each fragment were summed together, then the results for loop or stem species were pooled separately. Owing to the fact that all the representative fragments exhibited the same charge, it was not necessary to normalize the relative intensities by the respective charge state. It has been demonstrated that utilizing relative intensities of fully resolved isotopic signals is equivalent to employing the corresponding peak areas for most quantification purposes in FTICR mass spectrometry [34]. The value obtained from each pool was divided by their sum to express the separate contributions in percent form.

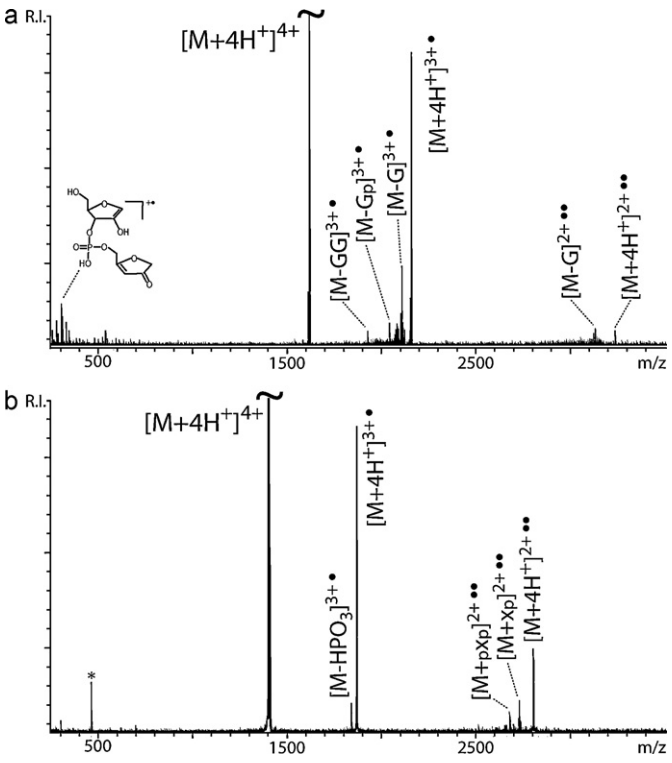


Fig. 1. ECD product ion spectra of (a) SL3 RNA and (b) abasic SL3. The letter p indicates a phosphate group; ×, an abasic nucleotide; *, noise produced by frequency harmonics.

3. Results and discussion

3.1. Gas-phase dissociation of structured constructs

Non-ergodic activation methods were initially employed for the top-down analysis of RNA structures involved in the lifecycle of retroviruses responsible for infectious diseases (Scheme 1). Representative results are provided in Fig. 1a, which displays the spectrum obtained by ECD activation of the 4+ charge state of SL3 hairpin. Its examination immediately reveals that the predominant products originated from the neutralization of up to two positive charges, which was induced by recombination of two electrons with the cationic precursor. Only a few fragments were detected in the low range and near the charge-neutralized products. Their masses correspond to structures produced by loss of G nucleobase and backbone cleavages resulting in mono- and dinucleotide species. The fact that G nucleotides are located both at the end of the double-stranded stem and in the single-stranded loop made it impossible to unambiguously assign the origin of such losses. For this reason, an analogue was analyzed in which the loop bases were replaced with abasic nucleotides (abasic SL3, Scheme 1). Under the same experimental conditions, the analogue exhibited the loss of abasic fragments, which clearly traced the cleavage events to the loop region of the construct (Fig. 1b). Similar outcomes were observed when these RNA samples were analyzed in negative ion mode and submitted to EDD activation. Although nucleic acid analytes can be detected in positive ion mode under appropriate conditions, their analysis is typically completed in negative ion mode to take advantage of their highly charged phosphate groups. In this case, electron detachment from the anionic precursors provided charge-reduced species with very limited fragmentation (not shown). These results were consistent with those reported by Mo and Håkannson for DNA oligonucleotides of similar length and charging [14].

Ergodic activation methods are generally expected to possess greater ability to generate characteristic backbone fragments that are consistent with the primary structure of the analyte under investigation. Fulfilling this prediction, IRMPD activation of the 4-charge state of SL3 produced typical ***c*** and ***y*** fragments by cleavage of the 5′ P–O bond (Fig. 2, see Ref. [6] for systematic nomenclature). Annotating the fragmentation events on the hairpin structure helped appreciate the fact that these ion series are not complete and do not cover the full extent of the RNA sequence (Fig. 2 inset and Table 1). A careful examination of the experimental data revealed the presence of numerous internal fragments corresponding to losses from the GGAG loop of the hairpin structure. At the same time, the observed sequence ions (i.e., ***c*** and ***y***) did not appear to be distributed over the entire length of the stem, but were limited to its terminal stretches. Taken together, these results suggest a greater susceptibility to fragmentation by single-stranded nucleotides located either in loop regions, or at the end of helical structures that may be momentarily destabilized by characteristic base-pair dynamics commonly described as "fraying" or "breathing."

3.2. Significance of structural context

The apparent structure-specific effects suggested by both the ECD/EDD and IRMPD data were further investigated by utilizing relatively short DNA oligonucleotides designed *ad hoc* (Scheme 1). In these constructs, different types of nucleotides were segregated in specific regions to facilitate tracing the origin of cleaved products and to enable direct correlations with the respective structural contexts. When the GC-short hairpin was activated by IRMPD, the predominant formation of ***a-B*** and ***w*** fragments was observed, corresponding to the cleavage of the 3′ C–O bond (Fig. 3a, see Ref. [6] for systematic nomenclature). The fragments carried charges ranging from 1- to 4-, the precursor's charge state. In this case, the characteristic ion series covered the entire extent of the hairpin structure, thus providing full sequence coverage. No evidence

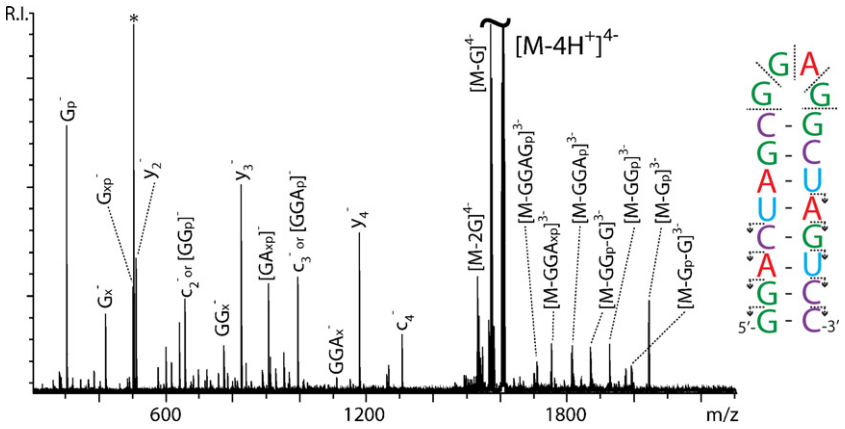


Fig. 2. IRMPD product ion spectrum provided by SL3 RNA. Refer to citation 6 for the systematic nomenclature of sequence ion series. The letter p indicates a phosphate group; ×, a base-loss product; *, noise. In the inset, the direction of the arrows points toward the end of the construct contained by a given fragment. Absence of arrow indicates internal fragments that do not contain construct ends.

Table 1
Summary of product ions obtained by submitting SL3 RNA to IRMPD and IRMPD/SORI-CID. The list includes typical series ions, as well as internal fragments. Expected mass over charge ratios (m/z) were calculated from the corresponding monoisotopic masses.

Product ion	Calculated m/z	Observed by IRMPD (Fig. 2)	Observed by IRMPD/SORI-CID (Fig. 5)
c_1^-	344.039	344.046	344.040
y_2^-	547.118	547.129	547.119
c_2^-	689.086	689.104	689.087
y_3^-	853.143	853.161	853.144
c_3^-	1018.138	1018.155	1018.139
y_4^-	1198.190	1198.207	1198.191
c_4^-	1323.179	1323.197	1323.179
y_5^-	1527.242	1527.261	1527.243
c_5^{2-}	814.098	–	814.099
y_6^{2-}	916.129	–	916.130
c_6^{2-}	978.624	–	978.625
y_7^{2-}	1068.650	–	1068.651
c_7^{2-}	1151.148	–	1151.149
y_8^{2-}	1241.174	–	1241.176
c_8^{2-}	1303.668	–	1303.669
y_9^{2-}	1413.697	–	1413.697
c_9^{2-}	1476.192	–	1476.193
y_{10}^{2-}	1578.223	–	1578.227
c_{10}^{3-}	1098.808	–	1098.808
y_{11}^{3-}	1166.828	–	1166.830
c_{11}^{3-}	1208.492	–	1208.494
y_{12}^{3-}	1281.844	–	1281.845
c_{12}^{3-}	1323.507	–	1323.510
y_{13}^{3-}	1383.525	–	1383.521
c_{13}^{3-}	1438.523	–	1438.522
y_{14}^{3-}	1498.540	–	1498.543
c_{14}^{3-}	1540.203	–	1540.206
y_{15}^{3-}	1608.224	–	1608.228
y_{16}^{4-}	1282.422	–	1282.512
Gp$^-$	344.039	344.046	344.040
Gx$^-$	458.071	458.082	458.082
Gxp$^-$	538.038	538.052	538.039
GGp$^-$	689.086	689.104	689.088
GGx$^-$	803.119	803.147	803.121
GAxp$^-$	867.090	867.122	–
GGAp$^-$	1018.138	1018.155	1018.139
GGAx$^-$	1132.172	1132.223	–
[M-GGAG$_p$]$^{3-}$	1704.359	1704.374	–
[M-GGAxp]$^{3-}$	1759.329	1759.339	–
[M-GGAp]$^{3-}$	1819.375	1819.383	–
[M-GAxp]$^{3-}$	1874.364	1874.381	–
[M-GGp-G]$^{3-}$	1878.701	1878.716	–
[M-GGp]$^{3-}$	1929.059	1929.064	–
[M-Gp-G]$^{3-}$	1993.717	1993.725	–
[M-Gp]$^{3-}$	2044.067	2044.092	–

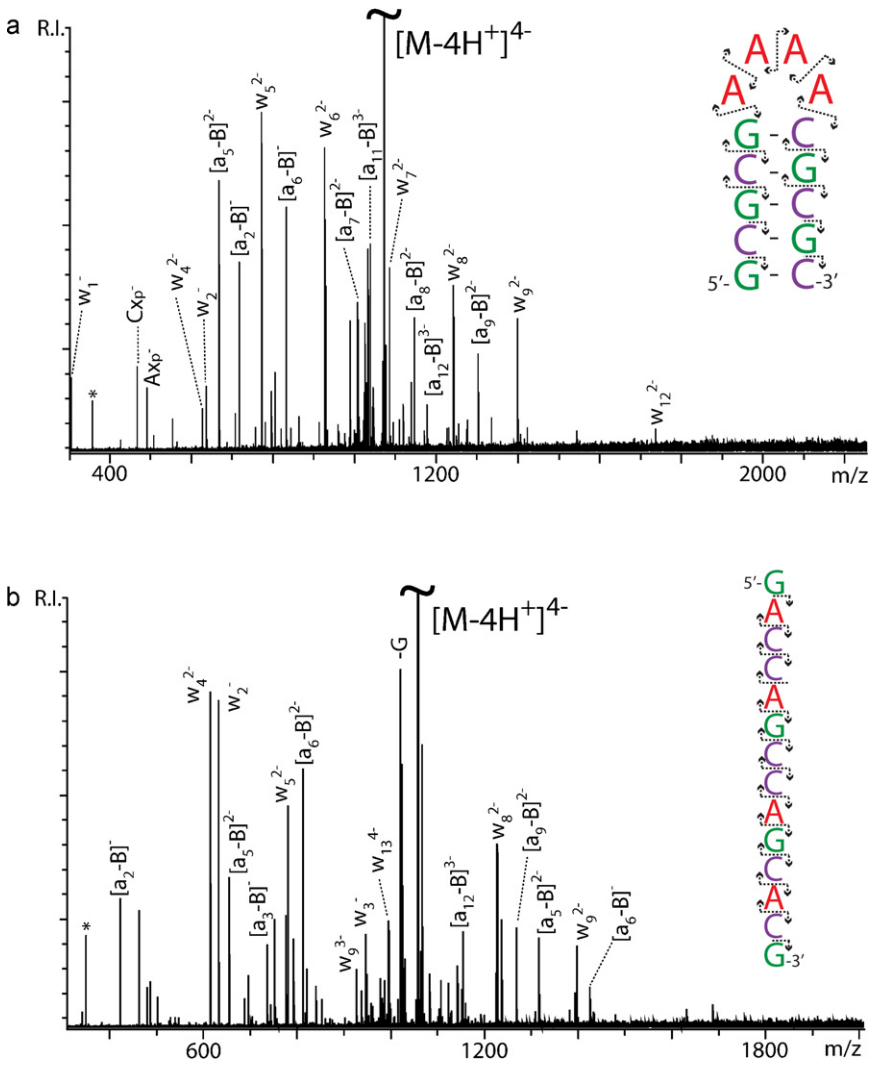


Fig. 3. IRMPD product ion spectra obtained from the 4- charge state of (a) GC-short and (b) GC-control. Refer to citation 6 for the systematic nomenclature of sequence ion series. The asterisk * indicates noise. In the insets, the direction of the arrows points toward the end of the construct contained by a given fragment.

was noted for preferential fragmentation patterns. These results matched those obtained from a construct sharing the same number of nucleotides, but devoid of any putative secondary structure (GC-control, Fig. 3b). However, a different outcome was observed when progressively lower charge states were selected for activation. For instance, the 3- charge state of GC-short displayed an appreciable decrease of sequence ions (i.e., ***a-B*** and ***w***) and a corresponding increase of internal fragments, the majority of which included adenine nucleotides (Fig. 4a). This trend was even more accentuated for the 2- charge state, which displayed almost exclusively internal fragments under the same experimental conditions (Fig. 4b). In contrast, the lower charge states of GC-control provided no remarkable variations from the results obtained from the 4- precursor (Fig. 3b), with well-distributed detection of sequence ions across the board (not shown). In these experiments, the prominent detection of A-containing internal fragments, which place the cleavage events in the loop region, should be considered in the backdrop of the simultaneous reduction of G- and C-containing fragments from the stem of the hairpin. The excellent correlation between the precursor charge state and the decreased incidence of stem fragmentation is consistent with the expected effects of Coulombic repulsion on possible strand separation (or "unzipping") dynamics of double-stranded structures, which increases with the charge state [17]. The loss of G and C nucleobases, which persisted also at lower charge states, could be attributed to terminal base-pairs affected by helix "fraying."

The influence of stem stability on the balance between loop and stem events was investigated by analyzing analogues with different base composition. A series of constructs of identical length was designed, which shared a conserved loop consisting of four inosines (I), but included stems of different GC content (Scheme 1). The comparison was carried out by monitoring three representative fragments from either the loop or the stem region, which were obtained by activating the 4- precursor ions under the same experimental conditions. The relative intensities for each set were pooled and then divided by their sum to express the different contributions in percentage terms (see Section 2). Plotted against the respective GC content, the contributions showed an increasing imbalance in favor of loop cleavage as the number of GC base-pairs increased (Fig. 5). Not surprisingly, this trend tracked very closely the increase of melting temperature (t_m) exhibited by the stems in the series. Analogous results were obtained by analyzing series in which t_m was increased by extending the length of the double-stranded region, thus further supporting the observation that fragmentation patterns are greatly influenced by the stability of underlying base-pairing interactions.

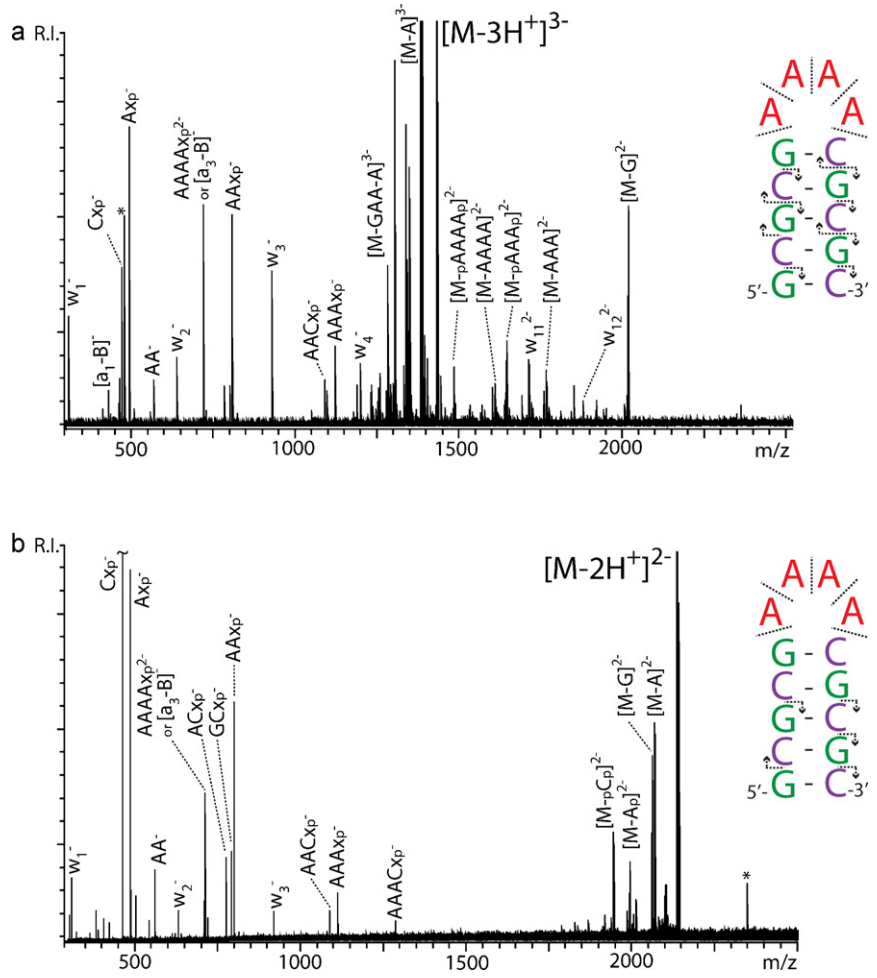


Fig. 4. IRMPD product ion spectra obtained from (a) the 3- charge state and (b) the 2- of GC-short. The asterisk * indicates noise. In the insets, the direction of the arrows points toward the end of the construct contained by the fragment. Absence of arrow indicates internal fragments that do not contain construct ends.

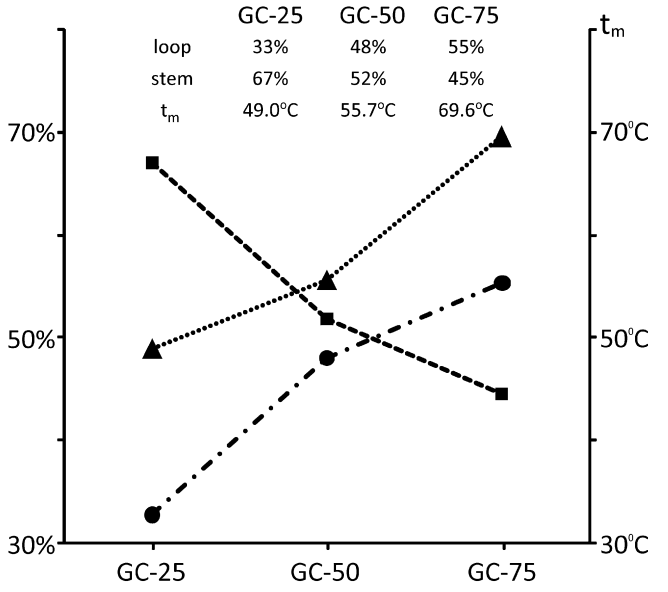


Fig. 5. Effects of GC content on the balance between loop and stem fragmentation. The percent contribution of representative loop species are marked with squares, those of stem species with circles. Triangles mark the melting temperatures of the respective constructs.

3.3. *Effects of base pairing on dissociation patterns*

Explaining the apparent inhibitory effects induced by base pairing must account for the fact that hydrogen bonding and π-stacking interactions involve functional groups that are distal to cleaved phosphodiester bonds. From a mechanistic point of view, their situation discourages a direct cause-effect rationalization. The initial loss of nucleobase by dissociation of the N-glycosidic bond is considered a necessary prelude to backbone cleavage, at least in DNA, which leads to the formation of ***a-B*** fragments [35–37]. Participation of the nitrogen base in hydrogen bonding and π-stacking interactions could possibly stabilize the N-glycosidic bond, thus preventing the loss that triggers the subsequent chain of events. However, the inhibitory effects are prominent also for RNA analytes, the dissociation of which is decoupled from the loss of nucleobase to produce typical ***c*** and ***y*** fragments [38–40]. It is possible that, in indirect fashion, base pairing may not inhibit regular cleavage processes at all, but may only mask their detection by keeping the ensuing products associated in individual complexes that are observed as whole. The increased stem fragmentation obtained from constructs of lower t_m and precursors of higher charge state is consistent with the ability of such species to dissociate during the analysis.

We tested this hypothesis by implementing a modified experimental design in which a supplemental activation step was inserted immediately after the main activation, with the intent of promoting the dissociation of non-covalent interactions that may still hold

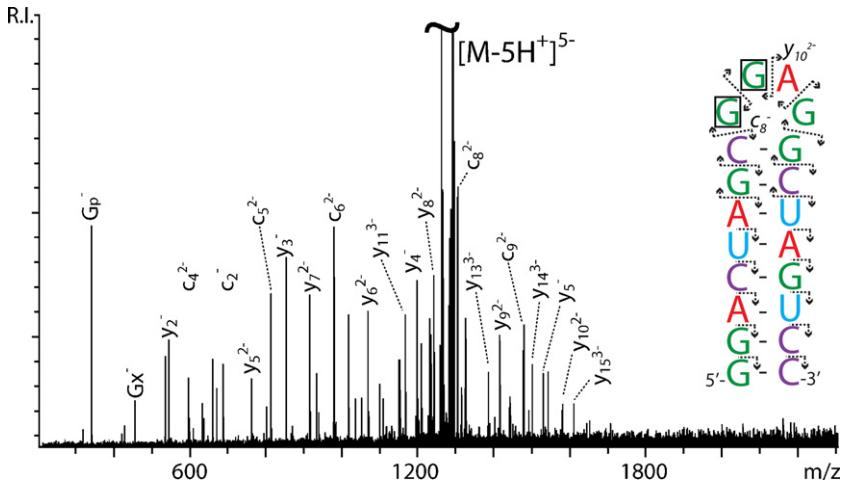


Fig. 6. Product ion spectrum obtained from SL3 RNA by IRMPD activation, followed by gentle SORI-CID dissociation. In the inset, the direction of the arrows points toward the end contained by a fragment. The loss of the boxed nucleotides could provide either the c_8^- and y_{10}^{2-} fragments detected here, or the $[M\text{-}GGp]^{3-}$ species observed in Fig. 2.

together the initial cleavage products. For example, the data in Fig. 6 were obtained from the 5- charge state of SL3, which was initially activated by IRMPD, then submitted to gentle SORI-CID with no intervening isolation step (see Section 2). Abundant sequence ions were readily recognizable, whereas no significant internal fragments were detected (Table 1). In contrast, no product ions were obtained when IRMPD was turned off to verify whether the selected SORI-CID conditions would lead by themselves to similar fragmentation. As shown by the secondary structure annotation, the cleavage events were not concentrated in a particular region, but

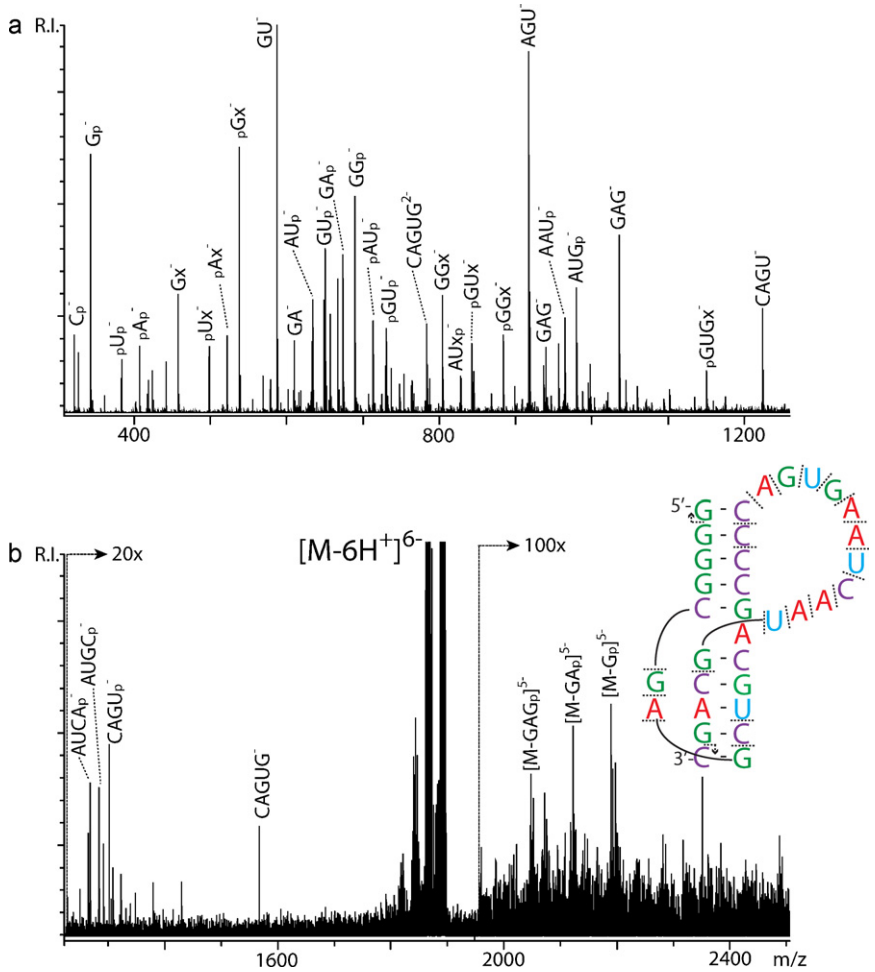


Fig. 7. IRMPD product ion spectra obtained from FIV pseudoknot. Panel (a) and (b) show different regions of the same spectrum with increased magnification. The letter p indicates a phosphate group; ×, a base-loss product; *, noise. In the inset, the direction of the arrows points toward the end contained by a fragment. Absence of arrow indicates internal fragments that do not contain construct ends.

were instead distributed over the entire construct to provide complete sequence coverage (Fig. 6 inset). In particular, it is important to note the absence of the prominent species corresponding to the loss of sizeable sections of the loop, which were observed instead in the experiments devoid of the additional activation step (Table 1). This absence can be explained with the detection in their place of sequence ions corresponding to the complementary strands that had been held together in annealed products. For example, if the cleaved section that formed $[M\text{-}GGp]^{3-}$ by IRMPD (Fig. 3) corresponded to guanine 9 and 10 of the construct, then the remaining strands would produce the c_8^- and $y_{10}{}^{2-}$ fragments upon dissociation of their mutual base pairing (Fig. 6 inset). These results indicate that the non-covalent association between cleaved strands can be effectively disrupted by a combination of additional activation and increased Coulombic repulsion.

The masking effects introduced by base pairing are expected to become more prominent as a function of the complexity of the structures under examination. This prediction is substantiated in the results afforded by the top-down analysis of the ribosomal frameshifting pseudoknot of feline immunodeficiency virus (FIV). Pseudoknots are tertiary structures stabilized by the intramolecular annealing of complementary sequences, which defines at least two distinct paired regions delimited by two or more unpaired stretches (Scheme 1). Submitted to IRMPD, the 6- charge state of FIV pseudoknot produced abundant fragmentation over a broad range (Fig. 7). However, with only a few exceptions, the signals could not be readily assigned to sequence ion series that, by definition, must include either the 5′- or the 3′-end of the RNA sequence. Instead, the vast majority corresponded to internal fragments originating from the cleavage of single-stranded loops, somewhat extending into adjacent regions affected by partial "fraying" effects (Fig. 7 inset). The fragments detected in the high m/z range (e.g., $[M\text{-}Gp]^{5-}$, $[M\text{-}GAp]^{5-}$, etc.) are consistent with the cleavage of the phosphodiester backbone in at least two points of a loop region, which would leave the resulting products associated through their paired contacts. Based on the apparent conservation of base-paring interactions, these product ions could favorably compare with the type II fragments described by Gabelica and De Pauw for DNA duplexes, which resulted from cleavage of the terminal ends of their partially "unzipped" strands [23].

4. Conclusions

The data presented here clearly support the existence of masking effects induced by base pairing interactions, which may prevent the observation of the full extent of fragmentation produced by complex nucleic acid structures. The fact that qualitatively similar results were obtained from both non-ergodic and ergodic methods indicates that such effects are independent of the type of activation employed, which is expected instead to affect their degree. Under appropriate conditions, the overall outcome can be the preferential observation of fragments originating from the unpaired regions of the substrate of interest. This apparent selectivity is influenced by the stability of the intervening double-stranded regions, which is a function of the number of interstrand hydrogen bonds (greater for GC than for AT/U pairs) and π-stacking interactions (greater for longer uninterrupted helices).

The ability to positively identify single-stranded regions could open the door for the utilization of these approaches to perform gas-phase footprinting of nucleic acids. The limited availability of structural information represents a major barrier to completing the functional annotation of the human genome, which consists mostly of non-protein coding sequences and, thus, cannot be classified directly on the basis of corresponding products. However, complex nucleic acid substrates are characterized by structural motifs involving a variety of interactions that differ substantially from the canonic Watson-Crick pairing considered here [41]. The practical application of top-down MS to elucidate higher-order structures will require establishing the rules for the formation of informative internal fragments from such motifs. In order to achieve this goal, further studies will be necessary to reach a greater understanding of the factors that determine the fate of paring interactions during gas-phase activation and fragmentation.

Acknowledgements

Support from The RNA Institute and the College of Arts and Sciences of University at Albany (SUNY) and from the National Institutes of Health (GM064328-11) is gratefully acknowledged.

References

[1] B. Ganem, Y.-T. Li, J.D. Henion, Tetrahedron Lett. 34 (1993) 1445–1448.
[2] K.J. Light-Wahl, D.L. Springer, B.E. Winger, C.G. Edmonds, D.G. Camp II., B.D. Thrall, R.D. Smith, J. Am. Chem. Soc. 115 (1993) 803–804.
[3] D. Fabris, Z. Wu, C. Fenselau, J. Mass Spectrom. 30 (1995) 140–143.
[4] P. Lecchi, L.K. Pannell, J. Am. Soc. Mass Spectrom. 6 (1995) 972–975.
[5] R.L. Cerny, K.B. Tomer, M.L. Gross, L. Grotjahn, Anal. Biochem. 165 (1987) 175–182.
[6] S.A. McLuckey, S. Habibi-Goudarzi, J. Am. Chem. Soc. 115 (1993) 12085–12095.
[7] E. Nordhoff, F. Kirpekar, P. Roepstorff, Mass Spectrom. Rev. 15 (1996) 67–138.
[8] R.A. Zubarev, N.L. Kelleher, F.W. McLafferty, J. Am. Chem. Soc. 120 (1998) 3265–3266.
[9] D.M. Horn, K. Breuker, A.J. Frank, F.W. McLafferty, J. Am. Chem. Soc. 123 (2001) 9792–9799.
[10] K. Breuker, F.W. McLafferty, Angew. Chem. Int. Ed. Engl. 42 (2003) 4900–4904.
[11] J.E.P. Syka, J.J. Coon, M.J. Schroeder, J. Shabanowitz, D.F. Hunt, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 9528–9533.
[12] S.N. Jackson, S. Dutta, A.S. Woods, J. Am. Soc. Mass Spectrom. 20 (2009) 176–179.
[13] J. Yang, J. Mo, J.T. Adamson, K. Hakansson, Anal. Chem. 77 (2005) 1876–1882.
[14] J. Mo, K. Hakansson, Anal. Bioanal. Chem. 386 (2006) 675–681.
[15] B.A. Budnik, K.F. Haselmann, R. Zubarev, Chem. Phys. Lett. 342 (2001) 299–302.
[16] S.A. McLuckey, J. Am. Soc. Mass Spectrom. 3 (1992) 599–614.
[17] K.X. Wan, M.L. Gross, T. Shibue, J. Am. Soc. Mass Spectrom. 11 (2000) 450–457.
[18] V. Gabelica, E. De Pauw, J. Mass Spectrom. 36 (2001) 397–402.
[19] R.C. Dunbar, T.B. McMahon, Science 279 (1998) 194–197.
[20] P.D. Schnier, J.S. Klassen, E.F. Strittmatter, E.R. Williams, J. Am. Chem. Soc. 120 (1998) 9605–9613.
[21] R.L. Woodin, D.S. Bomse, J.L. Beauchamp, J. Am. Chem. Soc. 100 (1978) 3248–3250.
[22] J.J. Wilson, J.S. Brodbelt, Anal. Chem. 79 (2007) 2067–2077.
[23] V. Gabelica, E. De Pauw, J. Am. Soc. Mass Spectrom. 13 (2002) 91–98.
[24] J.W. Gauthier, T.R. Trautman, D.B. Jacobson, Anal. Chim. Acta 246 (1991) 211–225.
[25] K.B. Turner, R.G. Brinson, H.Y. Yi-Brunozzi, J.W. Rausch, J.T. Miller, S.F.J. Le Grice, J.P. Marino, D. Fabris, Nucleic Acids Res. 36 (2008) 2799–2810.
[26] R.G. Brinson, K.B. Turner, H.Y. Yi-Brunozzi, S.F. Le Grice, D. Fabris, J.P. Marino, Biochemistry 48 (2009) 6988–6997.
[27] K.B. Turner, N.A. Hagan, A. Kohlway, D. Fabris, J. Am. Soc. Mass Spectrom. 17 (2006) 1401–1411.
[28] K.B. Turner, A.S. Kohlway, N.A. Hagan, D. Fabris, Biopolymers 91 (2009) 283–296.
[29] N.A. Hagan, D. Fabris, J. Mol. Biol. 365 (2007) 396–410.
[30] K.B. Turner, N.A. Hagan, D. Fabris, J. Mol. Biol. 369 (2007) 812–828.
[31] M.B. Comisarow, A.G. Marshall, Chem. Phys. Lett. 25 (1974) 282–283.
[32] C.L. Hendrickson, M.R. Emmett, A.G. Marshall, Annu. Rev. Phys. Chem. 50 (1999) 517–536.
[33] Rozenski, J. 1999.
[34] K.L. Goodner, K.E. Milgram, K.R. Williams, C.H. Watson, J.R. Eyler, J. Am. Soc. Mass Spectrom. 9 (1998) 1204–1212.
[35] S.A. McLuckey, G.J. Van Berkel, G.L. Glish, J. Am. Soc. Mass Spectrom. 3 (1992) 60–70.
[36] Z. Wang, K.X. Wan, R. Ramanathan, J.S. Taylor, M.L. Gross, J. Am. Soc. Mass Spectrom. 9 (1998) 683–691.
[37] S.T. Monn, S. Schurch, J. Am. Soc. Mass Spectrom. 18 (2007) 984–990.
[38] F. Kirpekar, T.N. Krogh, Rapid Commun. Mass Spectrom. 15 (2001) 8–14.
[39] S. Schurch, E. Bernal-Mendez, C.J. Leumann, J. Am. Soc. Mass Spectrom. 13 (2002) 936–945.
[40] T.Y. Huang, A. Kharlamova, J. Liu, S.A. McLuckey, J. Am. Soc. Mass Spectrom. 19 (2008) 1832–1840.
[41] N.B. Leontis, A. Lescoute, E. Westhof, Curr. Opin. Struct. Biol. 16 (2006) 279–287.