Optimized Fragmentation of Oligonucleotides Suppresses Undesired Fragmentation Products and Enables Confident Sequence Assignment

Joshua D Hinkle,[1] William McGee,[2] Joshua P Salem,[3] Chris Mullen,[1] Scott R. Kronwitter,[2] , Ping Yip,[1] John E. P. Syka,[1] James L. Stephenson Jr.,[2]
[1]Thermo Fisher Scientific, 355 River Oaks Pkwy, San Jose, CA, USA 95134, [2]Thermo Fisher Scientific, 250 Binney St, Cambridge, MA, USA 02142, [3]Department of Chemistry, University of Michigan, 930 N University Avenue, Ann Arbor, MI 48109-1055, USA

INTRODUCTION

Multiply-charged RNA anions have a very low critical energy of dissociation and can readily undergo consecutive dissociation events. There are multiple pathways for dissociation that require similar activation energies. Conventionally applied activation methods yield more than one sequence ion type as well as an abundance second generation ions - internal fragment ions and base losses from sequence ions. These redundant and interfering products greatly complicate spectral interpretation. Here we summarize the results from broad investigation of oligonucleotide dissociation methods. We have discovered that performing ion trap type resonant collisional activation (CID) using a reduced precursor ion activation Mathieu *q* value (ion resonant frequency) reduces the upper limit of the kinetic energy (KE) of activating collisions not available through adjustment of the auxiliary field amplitude (NCE) and suppresses the yield of the unwanted internal and base loss fragment ion species. The resulting product ion spectra are dominantly composed of first-generation products (sequence ions and base losses from the precursor). The reduction in product ion spectral complexity in favor of sequence informative species simplifies and increases the confidence of m/z peak annotation and therefore confirmation or elucidation of the primary RNA sequence. The low *q* CID methodology (lqCID) described here builds on previous studies involving optimization of the auxiliary field amplitude[1] to limit activation energy and effect such spectral simplification. The lqCID approach is applicable to both even electron RNA precursor ions obtained from electrospray ionization or odd electron intact species derived from negative ion ETD reactions or activated electron photo-detachment.

MATERIALS AND METHODS

Sample Preparation Oligonucleotide standards (obtained from IDT) were reconstituted in pure MS grade water at 1 ug/uL. Aliquots were further diluted in either pure MS grade water or 50 mM ammonium acetate to 10 uM for analysis.

Mass Spectrometry. Oligonucleotide samples (~4 uL) were loaded into gold coated pulled silica emitters and ionized via negative mode static ESI on either a Thermo Scientific™ Orbitrap™ Ascend Tribrid™ (with UV laser at 213 nm) or an experimentally modified Thermo Scientific™ Orbitrap Eclipse™ Tribrid™ adapted to enable infrared (IR,10.6 µm) activation in the linear quadrupole ion trap analyzer (LIT). Precursor ions were m/z selected (2 Th window) with the quadrupole. Conventional collision cell type collision induced dissociation (HCD, 8 mTorr N_2, 17% NCE) was performed in an ion routing multipole device. All other studied activation methods were performed in the LIT: Conventional trap CID (*q*= 0.25, 30% NCE, 10ms), low *q* CID (lqCID; *q* = 0.15, NCE 20%, 30ms), IRMPD (3W, 5ms, q = 0.05), NETD (Flouranthene reagent cation, 80ms) and AI-NETD (80ms ms, 1.5 watts). For the MS^3 experiments, NET-lqCID (Flouranthene cation, 80ms, q = 0.15) and EPD-lqCID (8 ms, 6 mTorr), the intact charge reduced product ions (NETnoD or EPD products) were m/z isolated (2 Th window) in the LIT and dissociated using low frequency (q) CID (*q* = 0.15, 15% and 10% NCE, respectively). Tandem MS spectra were acquired in the Orbitrap (RNA 20 mers 120,000 resolution, 100 avg. transients; RNA 50 mers 240k resolution, 300 avg. transients).

Data Analysis 20mer HCD, CID, lqCID and IRMPD spectra were analyzed via Thermo Scientific™ BioPharma Finder™ Version 5.0 (10 ppm mass accuracy, 0.5% relative intensity threshold). 20mer NETD - lqCID and EPD Spectra were assigned manually. 50mer spectra were deconvoluted with Zscape.[6] Assignments within 10 ppm were accepted. Ions <2 kDa and those identified with off-by-integer mass errors were validated manually.

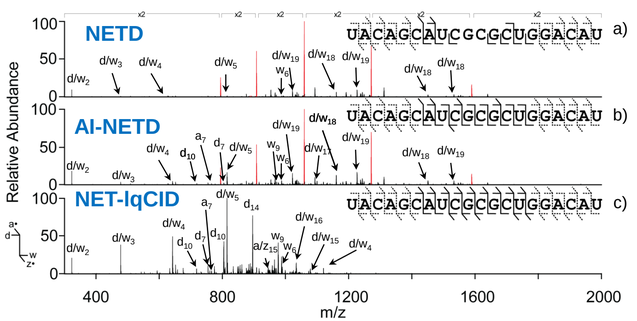

Figure 2: MS^2 spectra of the [M-8H] precursor (794 m/z) using a) NETD, b) AI-NETD, and c) NET-lqCID. Sequence coverage maps denote d/w and a•/z• ions. The dashed cleavages denote isobaric sequence fragment masses due to the palindromic sequence termini. The m/z peaks indicated in red are from intact charge-reduced products representing the -8 to -4 charge states of the 20mer.

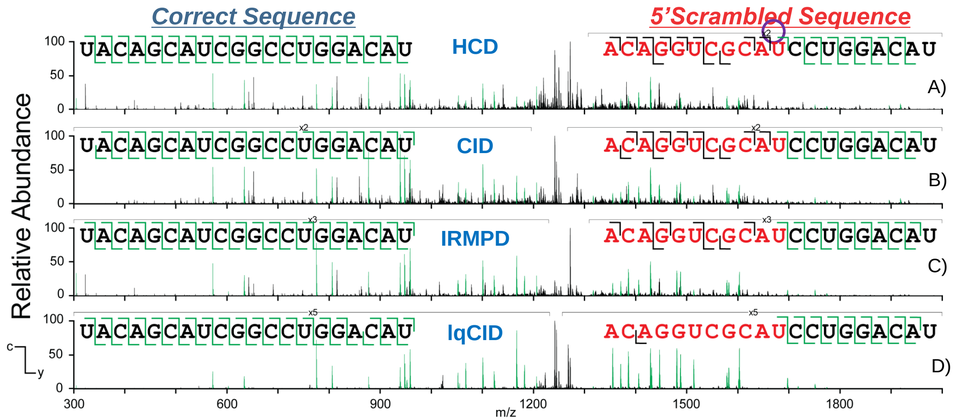

Figure 1: MS^2 spectra of RNA 20 mer, $[M-5H]^{5-}$ (1272 m/z), under well optimized/standard conditions using a) HCD , b) CID c) IRMPD, and d) lqCID. c/y fragment m/z peaks are colored green. Fragment maps for the correct sequence (left) and a 5' sequence scramble (right, scrambled bases in red) are above the corresponding spectra.

(only 4 bases) means many possible non-sequence fragment ions are

RESULTS AND DISCUSSION

Performance Comparison of Vibrational Activation-Based Dissociation Methods: Figure 1 shows the product ion spectra of the -5 charge state of an RNA 20mer using HCD, CID, IRMPD and lqCID. While all these ion activation techniques generated sequence information consistent with nearly all theoretical c/y fragments of the correct RNA 20mer (left), the degree of sequential dissociation in all methods except for lqCID also resulted in many "identified" c/y fragment ions for the scrambled sequence on the right. Both HCD and standard CID produced 11 false c/y ion pairs and IRMPD yielded 8 in total, while for lqCID only one 3' false positive y-ion was observed.

These results illustrate the limits to which any sequence of a particular region of an oligonucleotide molecule can be "confirmed" by simple observation of a few masses consistent with predicted bond cleavages. The limited diversity of elemental composition isobaric with c/y fragments of a candidate sequence. For all ion activation methods evaluated except for lqCID, which suppresses the production of such interfering product species, increasing uncertainty in the assigned fragment m/z peaks occurs when oligonucleotide length is extended.

Supplemental activation of electron deficient radical ePD or NEnoD products. An evaluation of lqCID as supplemental activation technique for odd electron intact charge reduced products was also studied. In Figure 2, ion activation of the RNA 20mer precursor, $[M-8H]^{8-}$ using NETD, AI-NETD, and NET-lqCID is shown. Negative ion ETD yielded predominantly intact electron products (NETnoD) with some low intensity d/w fragment products observed. AI-NETD improved the number of sequence fragment observed, but still yielded mostly NETnoD product ions. The NET-lqCID MS^3 analysis effected complete dissociation of the intermediate precursor, $[M-8H-e]^{7-}$, and yielded a more abundant and complete series of complementary d/w fragment ions.

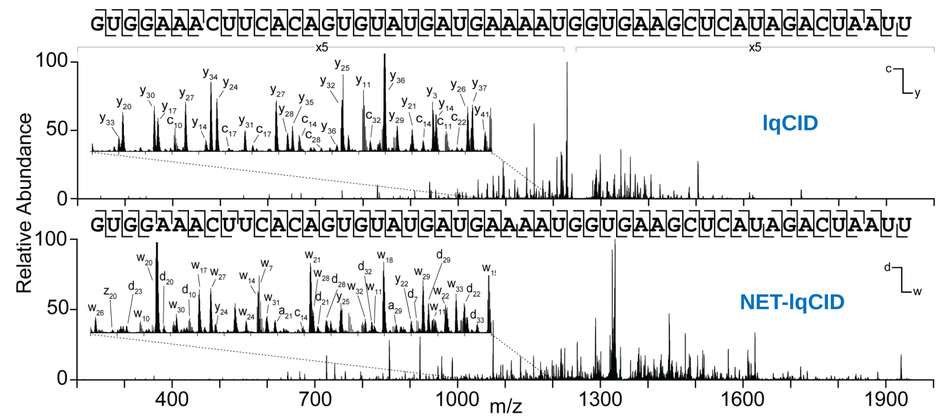

Figure 4: Annotated lqCID MS^2 spectrum (top) and NET-lqCID MS^3 spectrum (bottom) of RNA 50 mer (1239 m/z, -13). Fragmentation map denotes c/y (lqCID) or d/w (NET-lqCID) cleavage products.

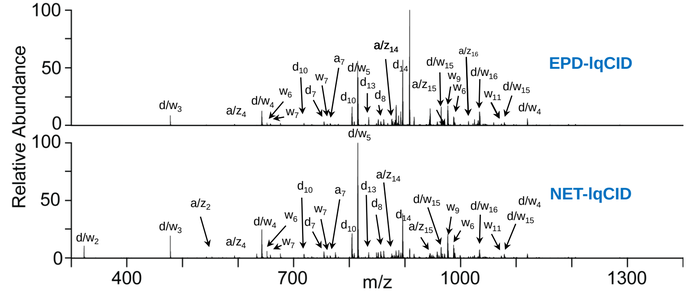

Figure 3: Fragmentation Spectra of the -8 UACAGCAUCGCGCUGGACAU precursor using EPD-lqCID and NET-lqCID

The highly labile a•/z• radical fragment products [2,3] were also observed, albeit at a lower intensity than d/w sequence ion paris. The NETD and AI-NETD yielded only a few m/z peaks corresponding to such a•/z• fragments.

NET-lqCID was also compared to EPD-lqCID (Figure 3). Resulting spectra were highly similar, containing nearly identical product ions m/z peaks with only marginal differences in abundance. Interestingly, the only notable difference was the auxiliary amplitude required to induce fragmentation, with EPD-lqCID requiring slightly less (10% NCE) than NET-lqCID (15% NCE). This indicates that performing EPD-lqCID using 213 nm UV lasers (commercially offered) provides comparable results to any NETD-based (not presently commercially available) method.

Fragmentation of Larger Oligonuceotides The lqCID method was applied to characterization of an RNA 50mer. Figure 4 shows the lqCID MS^2 and NET-lqCID MS^3 spectra of a $[M-13\ H]^{13-}$ precursor ion along with associated sequence fragment maps. Both lqCID and NET-lqCID succeeded in generating nearly complete c/y or d/w ion series, respectively. Notably, the spectral overlap of isotopic m/z peak envelopes was surprisingly low given the size of the RNA molecule, likely due to the limiting of the number of dissociation channels and the relatively high m/z (low charge density) of the precursor ion. Other results indicate that lqCID based methods may be successfully applied to substantially larger RNA molecules with similar performance enabling the extensive characterization of up to RNA 100mers. We believe that lqCID approaches will likely enable sequence analysis of even larger oligonucleotides.

CONCLUSIONS

Resonantly exciting oligonucleotides at lower Mathieu *q* (0.15) than the conventionally used Mathieu *q* of 0.25 promotes single dissociation events along the backbone of any given oligonucleotide. The observed spectra are less complex and more informative, maximizing the interpretability of tandem MS spectra for sequencing any oligonucleotide. It should be noted that in other work, the lqCID method has been applied successfully to single stranded DNA, modified RNA and to longer RNA oligomers (greater than 100 base pairs in length).

REFERENCES

1. Huang T et al. JASMS. **2008.** 19, 1832-1840.
2. Gao Y & McLuckey SA. Rapid Comm in Mass Spectrometry **2013,** 249-257.
3. Taucher M & Breuker K. Angewandte Chemie, **2012,** 124(45), 11451-11454.
4. Peters-Clarke TM.et al (2020). Anal. Chem., **2020,** 92(6), 4436-4444.
5. Gabelica V et al. Anal chem., **2006,** 78(18), 6564-6572.
6. Yip et al. Methods for Data Dependent Mass Spectrometry of Mixed Protein Analysis. US 10,217,619 B2, **2019.**

ACKNOWLEDGEMENTS

Many thanks to Corinne Lutomski and Tarick El-baba for providing the static spray emitters used for this analysis and Mike Senko for insightful discussion.

TRADEMARKS/LICENSING

PO2023-56EN