Novel Mass Spectrometry-Based Sequencing Method for Modified sgRNA

P # T.22

Brittney Pierson[1], Madeleine Mohrand[2], Timo Schierling[2], Eric Anderson[1], Stephan Seiffert[2], Ingo Roehl[2], Hans-Peter Vornlocher[2], Brent Morse[1], Sadik Kassim[1]

[1]Vor Biopharma, 100 Cambridgepark Dr, Cambridge, MA, USA 02140 [2]Axolabs GmbH, Fritz-Hornschuch-Straße 9, 95326 Kulmbach, Germany

INTRODUCTION

- CRISPR/Cas9 technology has become a primary tool for genome editing and gene regulation, yet the components of this system remain poorly characterized, particularly the single guide RNA (sgRNA).
- The CRIPSR/ Cas9 system utilizes a Cas9 protein which is an RNA-guided DNA endonuclease that complexes with sgRNA to form an active ribonucleoprotein (RNP). The 20-nucleotide variable region of the sgRNA targets a specific gene to enable the Cas9 protein to cleave the DNA resulting in a double-strand break (**Figure 1**).[1] 2'-O-methyl modifications and phosphorothioate linkages are commonly employed modifications on the sgRNA to increase genome editing[2] and increase cellular RNA stability[3,4] (**Figure 2**).
- Sequence confirmation of the sgRNA for clinical and commercial applications is essential to ensure specificity of the RNP complex for the target gene. Although next generation sequencing (NGS) is often used for sequence confirmation, this method indirectly sequences the molecule as RNA must first be reverse transcribed into DNA. Furthermore, NGS is unable to detect modified nucleic acids which enables only partial sequence confirmation for the modified RNA commonly used with CRISPR, base editing, and siRNA. A mass spectrometry-based (MS) sequencing was evaluated to enable direct sequencing of unmodified or modified sgRNA while also detecting modifications.

Figure 1. RNP Configuration

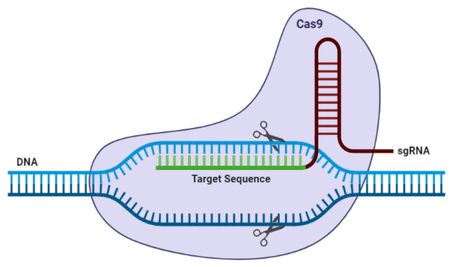
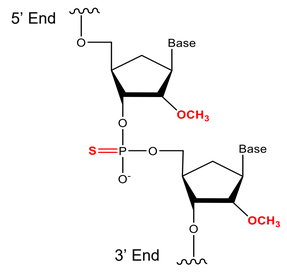


Figure 2. Modifications



MATERIALS AND METHODS

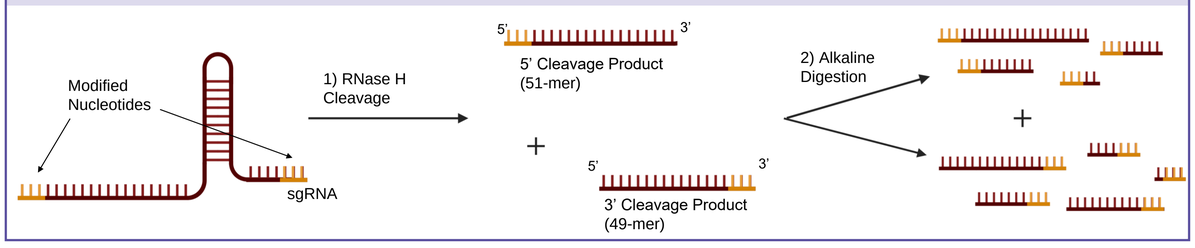
sgRNA Cleavage and Fragment Separation

- An sgRNA was designed and chemically synthesized with three 2'-O-Methyl and three phosphorothioate linkages on both the 5' and 3' ends to mimic common modification strategies used for sgRNA.
- The sgRNA was cleaved into two smaller fragments and separated prior to analysis to simplify the MS deconvolution and fragment detection.
- An antisense oligonucleotide (ASO) was incubated at elevated temperature to denature the sgRNA and hybridize the ASO and sgRNA which enables RNase H mediated cleavage of the sgRNA at position 51 (**Figure 3**).
- Biotinylated capture probes with complementary sequence to either the 5' or 3' end of the sgRNA enables separation of the cleavage products using Streptavidin Dynabeads™ in two separate reactions which magnetically separates the cleavage product from the rest of the mixture.

Alkaline Fragmentation of Separated Cleavage Products

- The cleavage products were chemically fragmented via alkaline digestion. Immediately after digestion, the solution was neutralized to prevent further fragmentation.
- The alkaline digestion generated random breaks in the phosphate backbone. The majority of the fragments contain only a single break which results in a mass ladder containing the full cleavage product (either 51 or 49 nucleotides) and all partial fragments (N-1, N-2, N-3, etc.) for each cleavage product **(Figure 3).**

Figure 3. sgRNA Cleavage



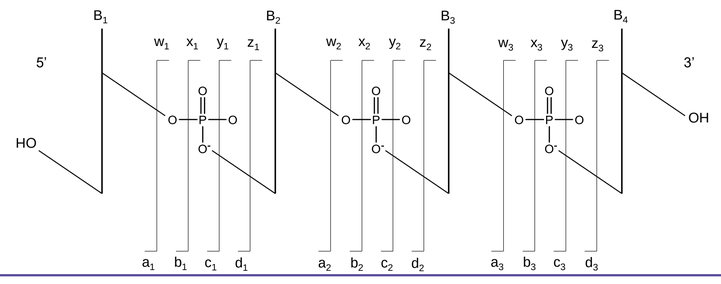
LC-MS Analysis of Chemically Digested Fragments

- Digested fragments were analyzed using denaturing ion-pair reversed-phase chromatography (IP-RP) with UV (Ultimate 3000 RS, Thermo Fisher) and ESI-MS detection (MaXis Q-TOF, Bruker) equipped with an ACQUITY UPLC BEH 130 Å C18 column (Waters). Analytes were separated via a series of gradient elution steps using mobile phase A= 100 mM HFIP, 16.3 mM TEA, 1% MeOH and mobile phase B= 50% ACN and 50% MeOH (v/v) as follows: 0-11% B for 20 minutes, 11-12% B for 0.5 minutes and 12-70% B for 4.5 minutes at 0.25 mL/min and 60°C column temperature.
- The fragments were identified using the MaXis Q-TOF in negative detection mode. UV traces were used to track the progression of the separation at 260 and 280 nm but were not used to generate relevant data.
- The mass spectra were extracted from the total ion chromatogram (TIC) or extracted ion chromatogram (EIC) and deconvoluted using Hystar Data Analysis software (Bruker) and assigned when the observed mass was within 0.45 Da of the theoretical mass, either monoisotopic or most abundant mass depending on the size of the fragment.

Tandem-MS Analysis of Modified 3' and 5' Ends

- The 5' and 3' ends of the sgRNA are chemically modified and thus cannot be fragmented via alkaline digestion. As such, the sequence of the ends is confirmed using tandem-MS.
- The same instrumentation, mobile phases, flow rate, column temperature, and MS detection mode were used for the LC-MS and tandem-MS evaluations, described above. Samples were loaded with 99% mobile phase A and analyzed using gradient elution from 1- 16% B over 10 minutes.
- Observed fragment masses were manually compared to the theoretical masses and identified as outlined by McLuckey[5] for charge states 1 and 2 (**Figure 4**). Both the precursor ion and at least one product ion must be detected in order to verify sequence.

Figure 4. RNA Fragmentation Pattern



RESULTS

- Both the full-length cleavage products (51-mer and 49-mer) and all partial fragments (N-1, N-2, etc.) were successfully identified indicating full sequence confirmation of the sgRNA (Tables 1-4). All results were confirmed by a second analyst on a second day for both the number and assignment of fragments.
- The 2'-O-methyl and phosphorothioate modifications for both ends were confirmed via tandem-MS resulting in complete sequence confirmation of the sgRNA (Table 6).
- Truncated sgRNA can be discriminated from the full-length sgRNA sequence as the full-length intermediate product was not detected for the truncated sgRNA. Other sgRNA sequences were also differentiated as the 5' end (the variable, target region of the sgRNA) cannot be sequenced due to the specificity of the capture probes for the cleavage products.
- Assay robustness was evaluated in triplicate by altering the alkaline digestion duration, alkaline digestion temperature and batch variability of the RNase H enzyme, none of which impacted the assay performance (Table 5).

ESI-MS Analysis of Cleavage Products

Figure 5. TIC of 3' Intermediate Product

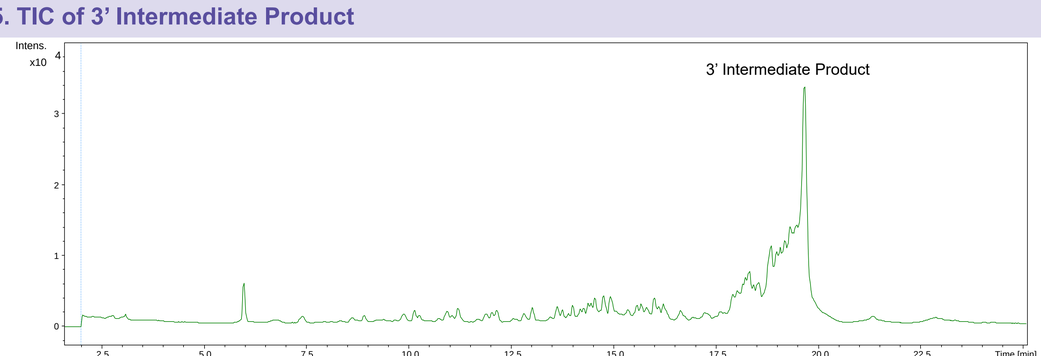
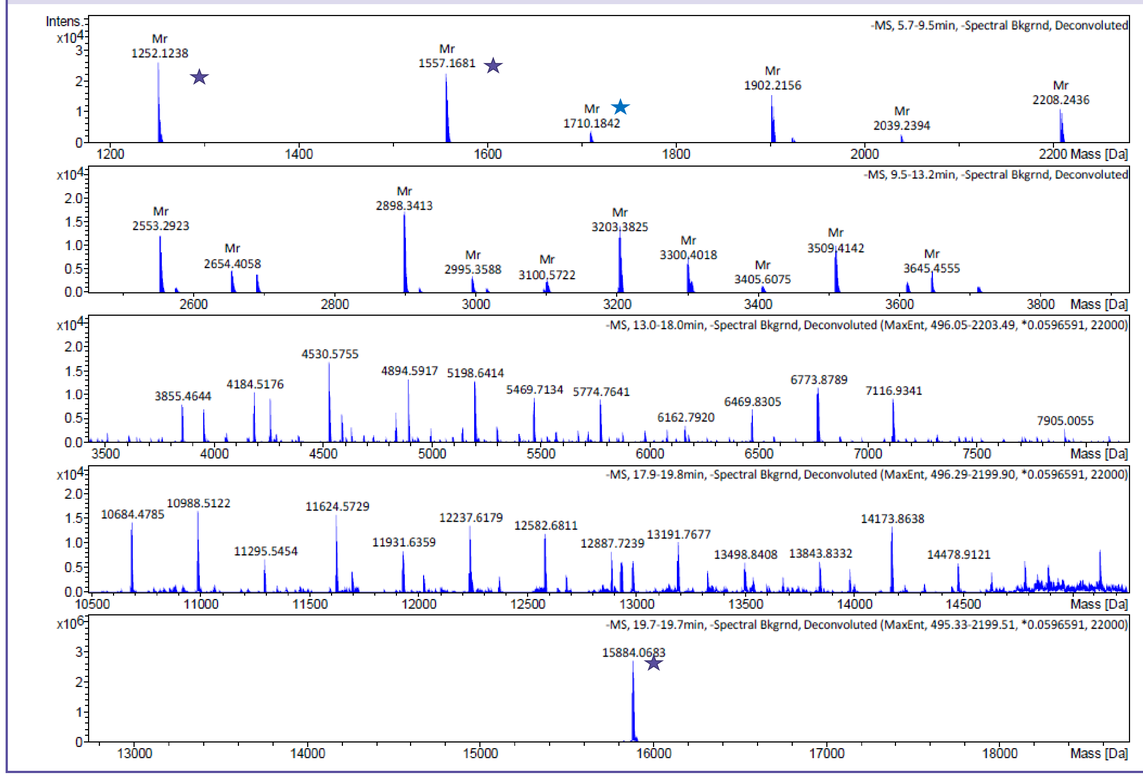


Figure 6. Deconvoluted Mass Spectra of 3' Cleavage Product (49-mer)



- TIC and representative mass spectra for the 3' cleavage product only. 5' cleavage product resulted in similar chromatographic and spectral profiles. Starred fragments are shown in Table 3.

Table 1. Fragment Detection for 3' Cleavage Product

Theoretical Fragments	46
5' Fragments Detected	46
3' Fragments Detected	43
Undetected Fragments	3' N-1, 3' N-2, and 3' N-3

Table 2. Fragment Detection for 5' Cleavage Product

Theoretical Fragments	48
5' Fragments Detected	46
3' Fragments Detected	48
Undetected Fragments	5' N-2 and 5' N-3

- Not all the theoretical fragments were detected for either cleavage product. However, there was no missing sequence coverage as the same fragments were detected from the opposite end. Further, the intermediate product was evaluated from both ends resulting in sequence overlap. The three nucleotides with modifications on each end were detected via tandem-MS.

Table 3. 3' Cleavage Product Fragment Identification Scheme

Fragment	MW (Most Abundant Mass)	MW (Monoisotopic)	Detected MW	Detection MW
FLP 3' End	15884.02	15877.00	15884.07	Most Abundant
5'-p-(N-1)	15475.00	15467.98	15475.06	Most Abundant
5'-p-(N-2)	15129.95	15122.94	15129.99	Most Abundant
5'-p-(N-3)	14783.90	14777.89	14783.98	Most Abundant
…				
5'-p-(N-43)	1902.21	1902.21	1902.21	Monoisotopic
5'-p-(N-44)	1557.16	1557.17	1557.17	Monoisotopic
5'-p-(N-45)	1252.12	1252.12	1252.13	Monoisotopic
3' (N-1)	15639.95	15632.93	Not Detected	Most Abundant
3' (N-2)	15303.93	15296.92	Not Detected	Most Abundant
3' (N-3)	14966.91	14960.90	Not Detected	Most Abundant
…				
3' (N-43)	2039.23	2039.23	2039.24	Monoisotopic
3' (N-44)	1710.18	1710.18	1710.17	Monoisotopic
3' (N-45)	1404.15	1404.15	1404.16	Monoisotopic

- Although not all fragments were detected, there was no missing sequence coverage as the same partial sequence was detected from the opposite end of the cleavage product (i.e., 3'(N-1) was not detected, but 5'-p-(n-1) was detected, thus the 48-mer for the 3' cleavage product was detected).

Table 4. 5' Cleavage Product Fragmentation Identification Scheme

Fragment	MW (Most Abundant Mass)	MW (Monoisotopic)	Detected MW	Detection MW
FLP 5' End	16381.2	16374.18	16381.29	Most Abundant
5' (N-1)	16046.17	16039.15	16046.20	Most Abundant
5' (N-2)	15711.13	15704.12	Not Detected	Most Abundant
5' (N-3)	15374.11	15368.10	Not Detected	Most Abundant
…				
5' (N-45)	1889.33	1889.33	1889.34	Monoisotopic
5' (N-46)	1560.28	1560.28	1560.28	Monoisotopic
5' (N-47)	1231.23	1231.23	1231.23	Monoisotopic
3' (N-1)	16114.1	16107.09	16114.15	Most Abundant
3' (N-2)	15808.08	15801.06	15808.12	Most Abundant
3' (N-3)	15479.03	15472.01	15479.02	Most Abundant
…				
3' (N-45)	1946.23	1945.22	1945.22	Monoisotopic
3' (N-46)	1641.19	1640.18	1640.18	Monoisotopic
3' (N-47)	1312.13	1311.13	1311.13	Monoisotopic

Table 5. Assay Robustness

Extended Digestion Time	*Sequence Confirmed*
Shorted Digestion Time	*Sequence Confirmed*
Increased Digestion Duration	*Sequence Confirmed*
Decreased Digestion Duration	*Sequence Confirmed*
RNase H Batch Variability	*Sequence Confirmed*

Tandem-MS Analysis of Modified sgRNA Ends

Table 6. Sequence Confirmation of Modified Nucleotides

Identified Fragment	Position	MW (monoisotopic)	*Detected MW*
[a4]-1	5' Cleavage Product	1230.15	1230.21
[b3]-1	5' Cleavage Product	927.15	927.14
[b2]-1	5' Cleavage Product	591.13	591.13
[c1]-1	5' Cleavage Product	334.03	334.03
[Z3]-1	3' Cleavage Product	897.09	897.09
[y2]-1	3' Cleavage Product	579.08	579.08
[w2]-1	3' Cleavage Product	339.01	339.01

- All modified nucleotides were detected confirming entire sgRNA sequence. All fragments were identified with a charge state of 1.

CONCLUSIONS

- A novel method employing ESI-MS and Tandem-MS was established to directly sequence modified or unmodified sgRNA. This method was demonstrated to be capable of both sequence and modification confirmation.
- The full-length fragments and all partial fragments for both cleavage products were detected resulting in full sequence confirmation.
- The assay is reproducible across multiple analysts and days, sequence specific, and robust.
- This method enables sequence confirmation across small RNA utilizing technologies, modified or unmodified, such as siRNA, base editing, CRISPR, or antisense RNA.

References

1. Jinek M., et al. *Science.* 2012; 337(6096): 816-821. **2**. Hendel A, et al. *Nat Biotech.* 2015; 33(9): 985-989. **3.** Allerson R, et al. *J Med Chem*. 2005; 48 (4): 901-904. **4.** Eckstein, F. *Nucleic Acid Ther.* 2014. 24 (6): 374-387. **5.** McLuckey SA, et al. *J Am Soc Mass Spectrom.* 1992; 3 (1): 60-70.

Acknowledgments

We thank Andreas Heider, Tobias Brambrink, Amy Mendel, Ankit Mehta, Robert Pietrusko, Jessica Higgins, Suzanne Pavluk, and Bruno Haas.

Certain illustrations created with BioRender.com.

Presented at

TIDES USA, Sept. 21-23, 2021

NOTE: Underlined authors contributed equally