# **Optimization of a Liquid Chromatography-Tandem Mass Spectrometry mRNA Sequence Mapping Workflow**



## Introduction

Structural characterization is necessary to ensure safety and efficacy of therapeutic mRNA oligonucleotides. LC-MS/MS bottom-up sequence mapping is a powerful solution for assessment and confirmation of mRNA oligonucleotide primary structure and detection of modifications. However, this technique faces several challenges including development of the endoribonuclease digest, separation of oligonucleotide digest products, assignment of MS/MS spectra, distinguishing isomers, and efficient data analysis and reporting. In this study we investigate the primary structure of an eGFP mRNA sequence (RiboPro) using a novel oligonucleotide mapping workflow.

# Methods – Sample Preparation

- $\succ$  hRNase4 (New England Biologics) cleaves at U|G and U|A.
- T4 PNK (New England Biologics) creates uniform hydroxylated 3' ends.
- $\geq$  40-minute digestion at 37°C.
- Sample is injected directly and desalted with a 5-minute hold at 3% B.

# Methods – IPRP-MS/MS

Table. 1: Ion Pairing Reverse Phase (IPRP) Method Summary

Mobile Phas	Phase A		95 mM HFIP, 5.7 mM DIPEA in water			
Mobile Phas	se B	7.1 mM HFIP, 2.1 mM DIPEA in 80% methanol				
Column	F	Phenomenex bioZen 2.6 µm Oligo 150 x 2.1 mm PN: 00F-4790-AN				
Column Temp.		75°C				
Injection Vol.		5 µL				
Flow Rate		300 µL/minute				
Gradient						
Time (min)	%	B	MS Analysis: Thermo Q-Exactive			
0	39	%	Plus using Data Dependent Acquisition in negative mode.			
5	39	%				
50	45	%				

Samples Sequences Processing nodes	
MS/MS Ids (Byonic)	Quant (Byologic)
<ul> <li>&gt; General</li> <li>&gt; Protein database options</li> <li>&gt; Instrument Parameters</li> <li>&gt; Digestion</li> <li>&gt; Modifications</li> <li>&gt; Glycans</li> <li>&gt; Inclusion</li> <li>&gt; MS/MS Filtering</li> <li>&gt; Spectrum Input Options</li> <li>&gt; Peptide Output Options</li> <li>&gt; Protein Output Options</li> <li>&gt; Protein Output Options</li> <li>&gt; Multicore options</li> </ul>	<ul> <li>&gt; General</li> <li>&gt; In-silico options (Theoretical Digest)</li> <li>&gt; In-silico options (CSV import)</li> <li>&gt; MS extract options</li> <li>&gt; Advanced</li> <li>&gt; Report</li> <li>&gt; UI Configurations</li> <li>&gt; Feature Finder</li> <li>&gt; XIC</li> </ul>
> Wildcard > Report	
Fixed and variable modifications are identified and localized	Enter custom modification text in fine control format: Cyclic Phosphate / +61.955 @ CTerm   rare1 Cation:Na / +21.981943 @ A,G,U,C   common1
	Create Project

Fig. 1: Project creation window allows users to define expected sequence and processing parameters for their specific instrument and analysis.

➢ Workflow includes Wildcard Search<sup>™</sup>, Modification Fine Control, OSM scoring, and Automatic/Custom Reporting.



separated with Phenomenex bioZen 2.6 um Oligo 150 x 2.1 mm column.

Steven Broome<sup>1</sup>, Roxana Eggleston-Rangel<sup>2</sup>, Maria Basanta-Sanchez<sup>1</sup>, and Marshall Bern<sup>1</sup> <sup>1</sup>Protein Metrics, Cupertino, CA and <sup>2</sup>Phenomenex, Torrance, CA

# Methods – Data Processing

> The data processing workflow used in this study is a Protein Metrics beta product still under development.

➤ Data analysis was performed using Byos<sup>TM</sup> and a repurposed version of Byonic® for MS/MS oligonucleotide spectrum matching (OSM).

> The solution accepts all major MS vendor data formats.

# **Results – Distinguishing Isomers**



Fig. 3: XIC of two resolved isomeric oligonucleotides. In Byos red dots represent OSMs assigned to the currently selected oligonucleotide while grey dots represent OSMs assigned to unique isomers.



Fig. 4: Select screenshots of the Byos project investigation view used for manual inspection of assignments. Oligonucleotide spectra for each isomer are annotated with assigned fragment ions (A, B). Some columns from the oligonucleotide table are shown (C).

- and d4-18 are observed.
- Delta Score: OSM score difference between the best (assigned) and next best assignment.
- Intelligent OSM scoring and IPRP separation are critical.

Isomers with distinct mRNA sequences are common for oligonucleotide mapping. Confident assignment of isomers is critical and

challenging.

An example isomer pair is separated by IPRP and automatically assigned using our data processing solution (Figures 3 and 4).

Isomer distinguishing fragment ions y2, w2, y4, b2, d2-18,

# Results – Automatic Reporting

ji mRNA eGFP									
	2 <u>0</u>	30	4 <u>0</u>	5 <u>0</u>	6 <u>0</u>	7 <u>0</u>	8 <u>0</u>	9 <u>0</u>	
GGAGACGCCGCC	ACCAUGGUGA	GCAAGGGGGGA	GGAGCUGUUC	ACCOGGGGGGG	UGCCCAUCCU			UAAACGGCCA	
140	150	160	170	180	190	200	210	220	230
CCUACGGCAAGC	UGACCCUGAA	GUUCAUCUGC	ACCACCGGCA	AGCUGCCCGU	GCCCNGGCCC	ACCCUCGUGA	CCACCCUGAC	CUACGGCGUG	CAGUGCUU
270	280	290	300	31.0	320	330	340	350	360
CGACUUCUUCAA	GUCCGCCAUG	CCCGAAGGCU	ACGUCCAGGA	GCGCACCAUC	UUCUUCAAGG	ACGACGGCAA	CUACAAGACC	CGCGCCGAGG	UGAAGUUC
	_								
40 <u>0</u>	41 <u>0</u>	42 <u>0</u>	43 <u>0</u>	<u>440</u>	45 <u>0</u>	46 <u>0</u>	47 <u>0</u>	<u>480</u>	<u>490</u>
CUGAAGGGCAUC	GACUUCAAGG				AGUACAACUA	CAACAGCCAC	AACGUCUAUA	UCAUGGCCGA	
500	- 10		5.60	5.5.0	500			61.0	60.0
530 GCCACAACAUCG	540 AGGACGGCAG	550 CGUGCAGCUC	560 GCCGACCACU	570 ACCAGCAGAA	58 <u>0</u> CACCCCCAUC	590 GGCACGGCCC		610 CCCGACAACC	62 <u>0</u> ACUACCUG
									=
66 <u>0</u>	67 <u>0</u>	<u>680</u>	69 <u>0</u>	70 <u>0</u>	71 <u>0</u>	72 <u>0</u>	73 <u>0</u>	740	75 <u>0</u>
AACGAGAAGCGC	GAUCACAUGG			GCCGCCGGGA	UCACUCUCGG	CAUGGACGAG	CUGUACAAGU	AAGGAGAAGA	GAAGGAAG
790	800	810	820	830	840	850	860	870	880

Fig. 5: eGFP sequence coverage map with black lines representing oligonucleotide IDs. 80% sequence coverage excluding the PolyA tail.

### **OSM Assignment Table: Score with Delta Score Heatmap**

					MS Alias n
Start AA $\uparrow$	End AA $\uparrow$	Sequence (unformatted) $\uparrow$	z↑	Calc. m/z ↑	
	72	GGUCGAGCU	-2	1433.7	
64	84	GGUCGAGCUGGACGGCGA		756.655	
		CGU	-8	851.363	
67	72	CGAGCU	-2	935.642	
73	101	GGACGGCGACGUAAACGG		784.435	
		CCACAAGUUCA	-11	855.839	
	02		2	1264.2	

Fig. 6: A report table with numerical OSM Scores and heatmap color code illustrating Delta Scores, valuable for assessing reliability of assignment.

### Variable Modification: 2',3'-Cyclic Phosphate

			MS Alias n
Sequence (unformatted) ↑	Apex Time (Posit) ↑	Mod. Summary ↓	
ACAACAGCCACAACGUCU	27.5698		
	28.0611	CTerm(Cyclic Phosphate/61.95	50)
ACAACU	16.8051		
	19.3696	CTerm(Cyclic Phosphate/61.95	50)
ACAAGACCCGCGCCGAGGU	J <b>27.6605</b>		
	28.1019	CTerm(Cyclic Phosphate/61.95	50)

Fig. 7: A report table with assigned variable modifications. Relative abundance of modified vs. unmodified is calculated automatically.

## Conclusions

- $\geq$  80% eGFP seq. coverage (-PolyA) with a single digest.
- Beta OSM tool (repurposed Byonic) assigns OSMs with intelligent scoring and detection of modifications.
- Workflow separates and automatically detects isomers.

#### Acknowledgments

We gratefully acknowledge Roxana Eggleston-Rangel and Phenomenex for preparing samples and acquiring data.





