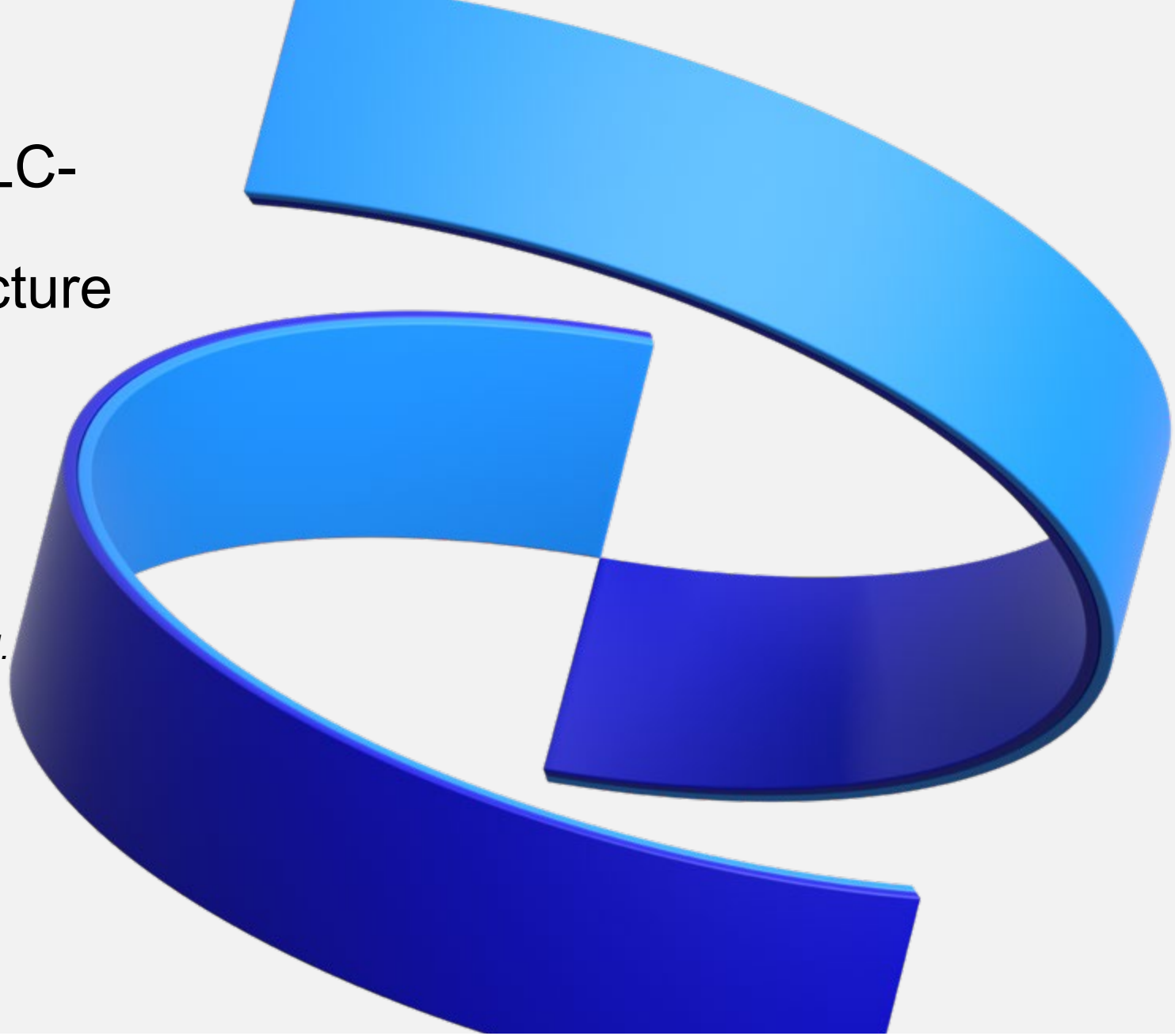
Oligonucleotide Mapping via LC-UV-MS/MS to Enable Comprehensive Primary Structure Characterization of an mRNA Vaccine Against SARS-CoV-2



Andrew W. Dawdy[1], Brian C. Gau[1], Leah Hanliu Wang[1], Carlos H. Castaneda[1], Olga V. Friese[1], Matthew S. Thompson[2], Thomas F. Lerch[1], David J. Cirelli[2], and Jason C. Rouse[2]

[1]Biotherapeutics Pharm. Sci., Pfizer Inc, St Louis, MO, USA and [2]Andover, MA, USA

CASSS MS - Practical Applications of Mass Spectrometry in the Biotechnology Industry

Chicago, September 8th, 2023

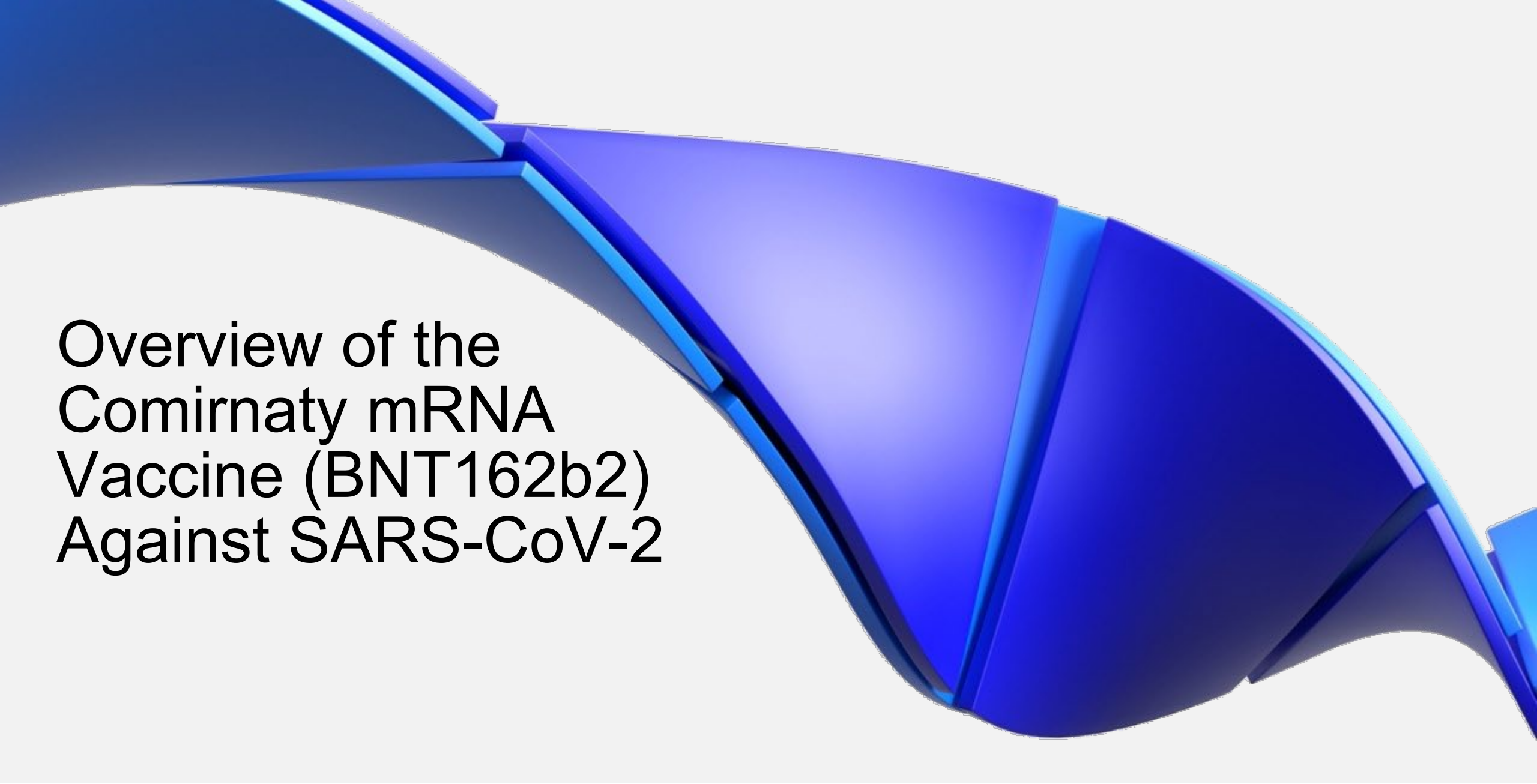
Pfizer

Breakthroughs that change patients' lives

All authors and work were funded by Pfizer.

Outline

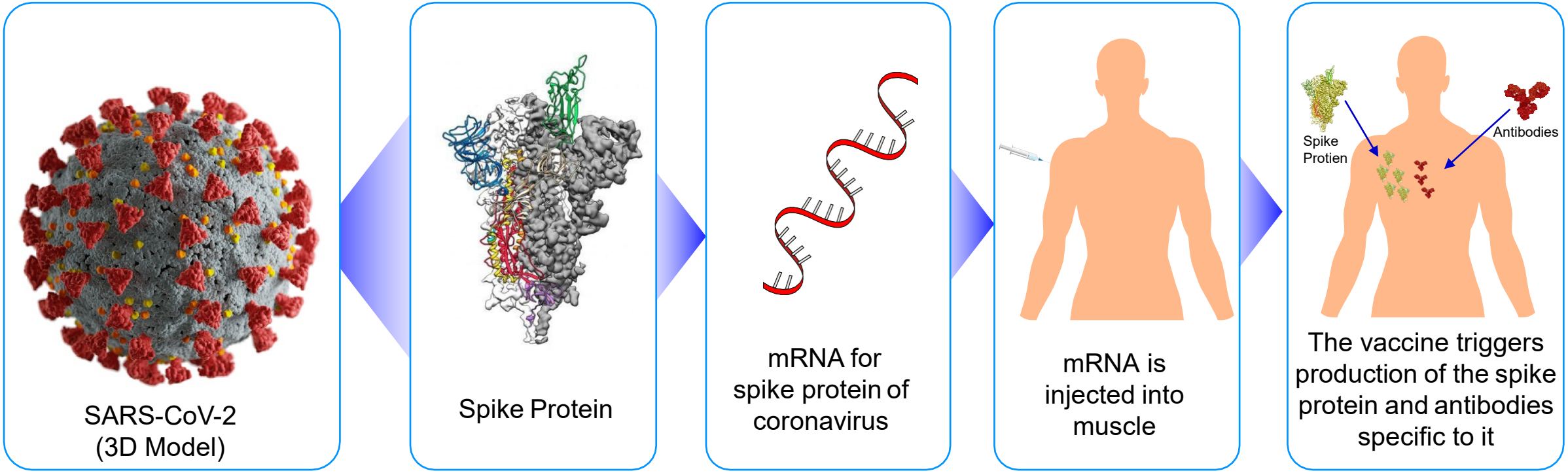
- Overview of the Comirnaty (BNT162b2) mRNA Vaccine Against SARS-CoV-2
- Oligonucleotide Mapping of BNT162b2 mRNA Primary Structure by LC-UV-MS/MS
  - Workflow
  - Core reportables providing direct primary structure understanding
- Oligonucleotide Mapping Method Development Highlights
- Additional Oligonucleotide Mapping Application Highlights
  - Batch Comparability
  - Variant Construct Comparisons



Overview of the Comirnaty mRNA Vaccine (BNT162b2) Against SARS-CoV-2

Basic Design of Pfizer/BioNTech mRNA Vaccine(s) against SARS-CoV-2

- Train patient's immune system to recognize the virus, specifically the spike protein on the surface
- Give the "code" or "recipe" of the spike protein to your cells



(Wrapp et al., 2020, Science)

Analytical Characterization of the Drug Substance (mRNA) is Critical for Development of a High Quality Manufacturing Process & Product

Drug Substance (mRNA)

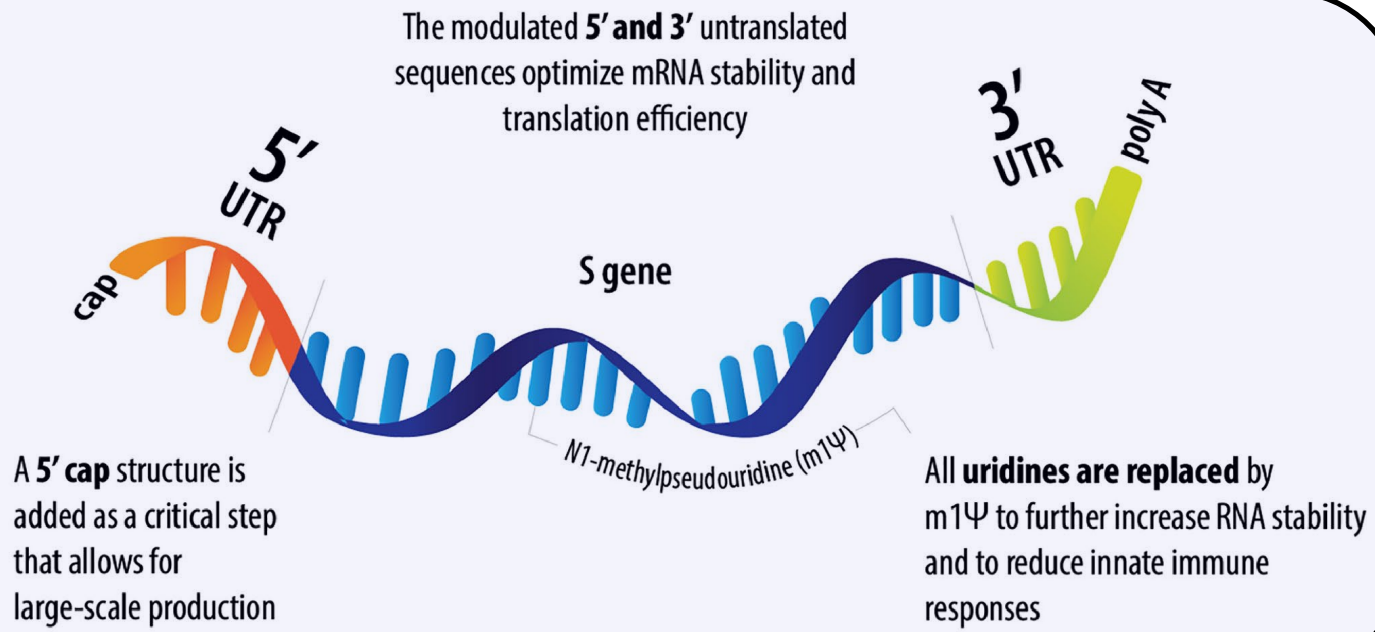


Figure adapted from Lewis LM, Badkar AV, Cirelli D, Combs R, Lerch TF. J Pharm Sci. 2023 Mar;112(3):640-647.

Platform QC Assays

- **Compendial methods**
- **Purity** by **Capillary Gel Electrophoresis**
- **Purity** by **Immunoblot**
- **Concentration** by **UV spectroscopy**
- **Identity, Impurities** by **PCR-based methods**

Heightened Characterization

Primary Structure

- **Oligonucleotide mapping (LC-UV-MS/MS)**
- **Nucleoside Analysis (LC-UV-MS)**
- **NextGen Sequencing (NGS)**

Higher Order Structure

- **Circular Dichroism (CD)**

Protein Expression

- **FACS**
- **Western Blot**

Oligonucleotide Mapping of mRNA Primary Structure by LC-UV-MS/MS is Applied in Three Ways to Support mRNA Vaccine Development

Direct Primary Structure Understanding
- 5´ terminus (capping)
- 3´ terminus (poly(A) – tail)
- Full-length mRNA

Batch Comparability Assessment
- Process changes
- Scale-up
- Scale-out

Orthogonal Identity
- BNT162b2 (Original)
- Variant constructs (Delta, Omicron)

Oligonucleotide Mapping of mRNA Primary Structure by LC-UV-MS/MS has Supported Regulatory Filings And Launches in 180+ Markets Globally

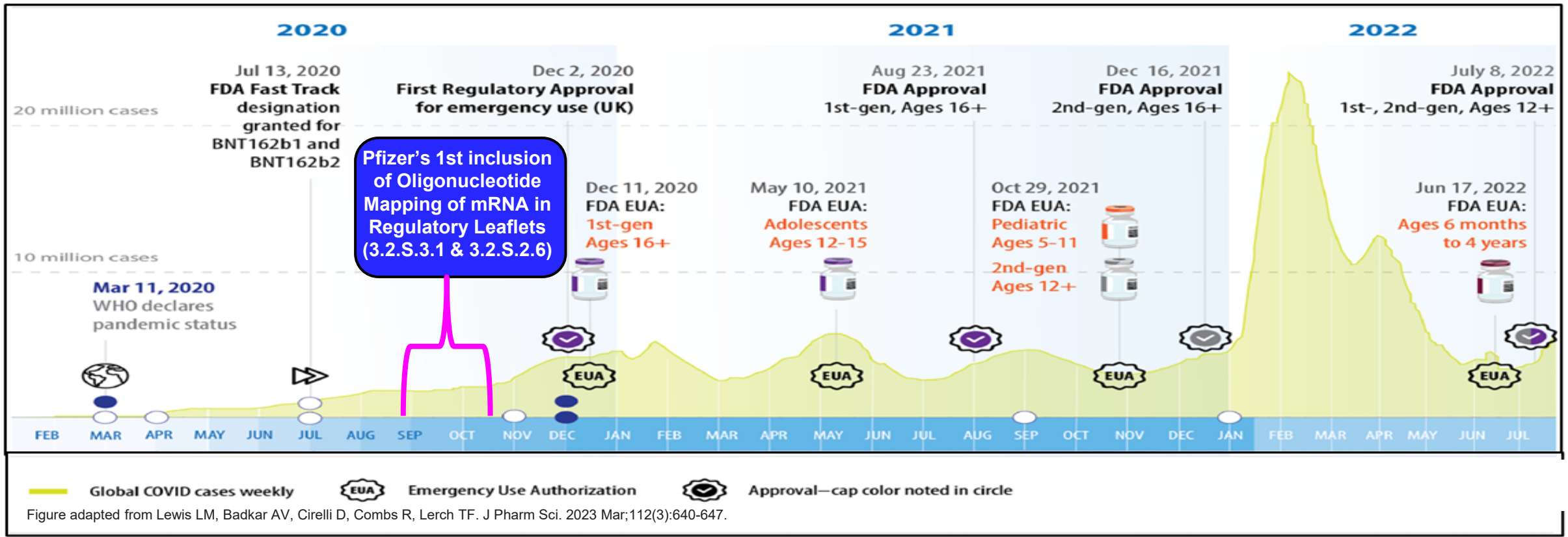
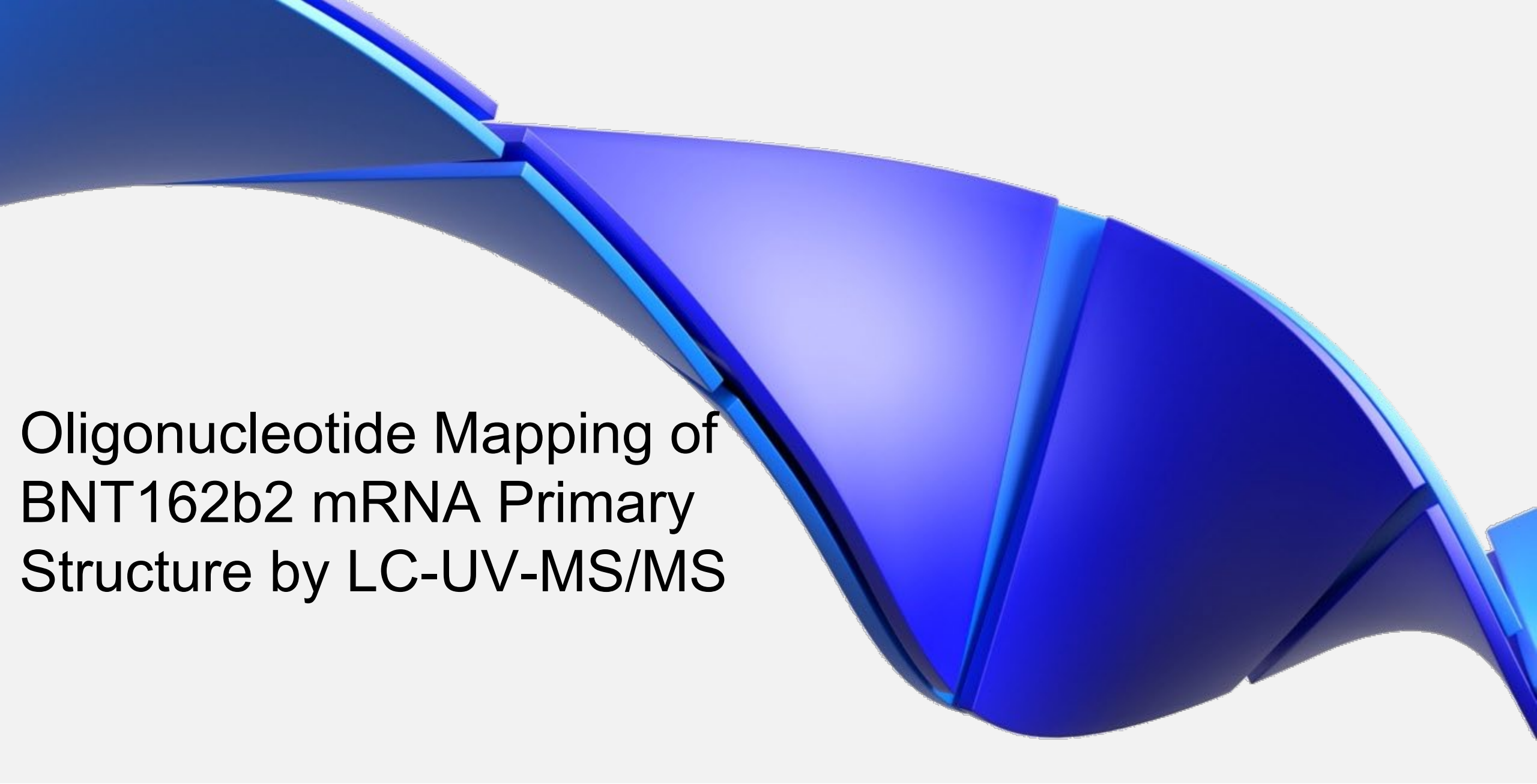


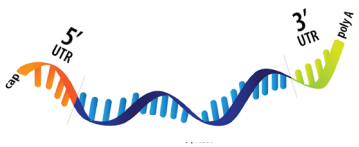
Figure adapted from Lewis LM, Badkar AV, Cirelli D, Combs R, Lerch TF. J Pharm Sci. 2023 Mar;112(3):640-647.



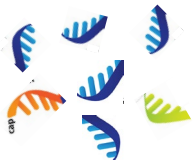
Oligonucleotide Mapping of BNT162b2 mRNA Primary Structure by LC-UV-MS/MS

Simple, Robust, Semi-Automated Workflow

Rapid One-Pot, One-Enzyme Sample Preparation



Ribonuclease T1
Cleave 3′ to G



+

IP-RP-HPLC-UV

MS-MS/MS

+

Semi-Automated Data Analysis

Commercial Software + In-House VBA Tools

Semi-Automated Data Analysis Workflow

1. Automated Search

- Mass table by retention time
- Identifications (~70% Coverage)

2. Automated LC-UV Annotation

- Match Peak IDs to Chromatogram
- Reformatted Mass Table

3. Supplement LC-UV Annotation

- Data mining & MS/MS Analysis Tools

4. Supplement Missing Coverage

- Data mining & MS/MS Analysis Tools

5. Add 5′ & 3′ Termini Characterization

Final Reportables

- Fully-Annotated Chromatographic Map
- Sequence Coverage Calculation & Map
- Curated Mass Table
- 5′ & 3′ terminus characterization

Result: Fully Annotated BNT162b2 Oligonucleotide Map

Complete Mass Table:
1062 consensus oligonucleotides

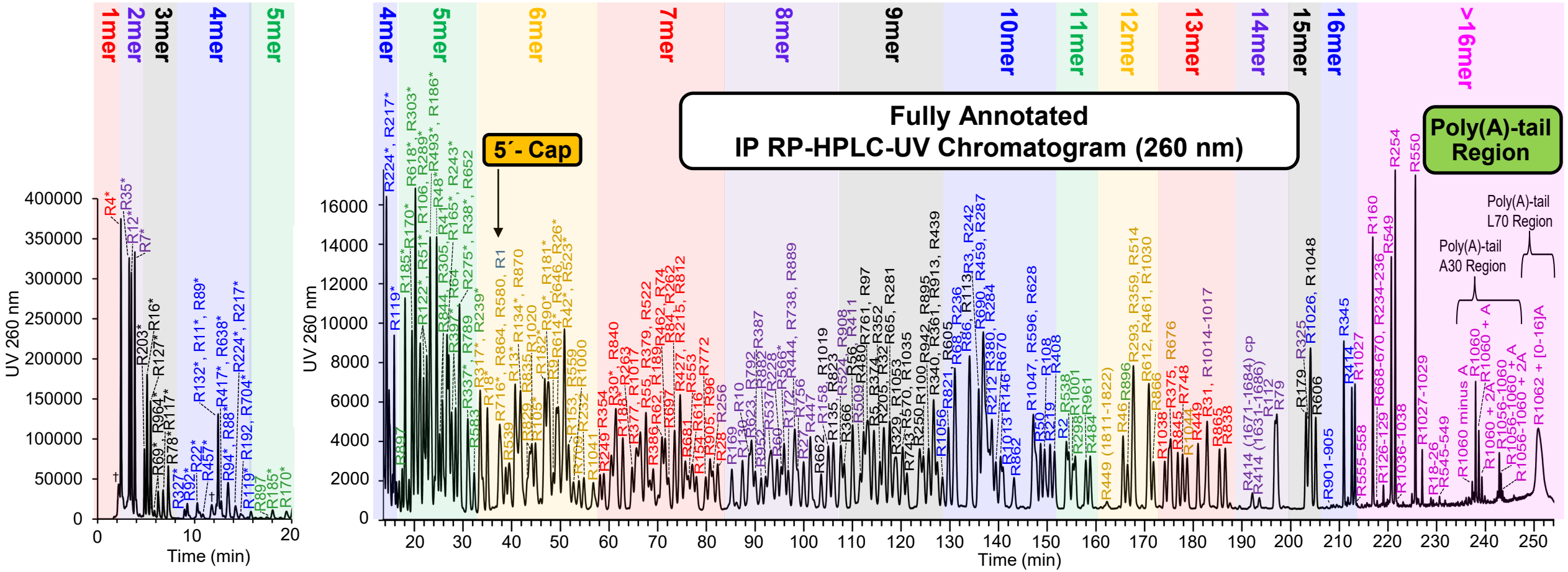
→

Group to 302 unique oligonucleotides

→

232 oligonucleotides
(48.8% Coverage)
One Locus

70 oligonucleotides
(51.2% Coverage)
Multiple Loci



Pfizer

Result: 100% BNT162b2 Sequence Coverage Observed

Approximately half of consensus RNAseT1 cleavages map to one locus and half contain multiple loci

232 oligonucleotides (48.8% Coverage) — **One Locus**

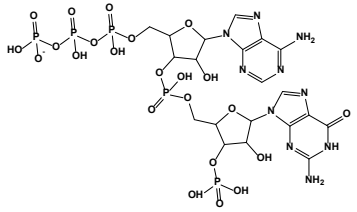
70 oligonucleotides (51.2% Coverage) — **Multiple Loci**

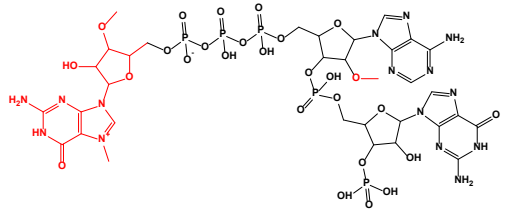
46 oligonucleotides — **14 oligonucleotides** — **Missed-Cleavages & Non-Concensus Cleavages**

1	AGAAVAAACVAGVAVVCVVCVGGVCCCCACAGACVCAGAGAGAACCCGCCACCAVGVVCGVGVVCCVGGVGCVGCVGCCVCVGGVGVCCAGCCAGVGVGVGAACCVGACCACCAGAACAC	120
121	AGCVGCCVCCAGCCVACACCAACAGCVVVACCAGAGGCGVGVACVACCCCGACAAGGVGVVCAGAVCCAGCGVGCVGCACVCVACCCAGGACCVGVVCCVGCCVVVCVVCAGCAACGVGA	240
241	CCVGGVVCCACGCCAVCCACGVGVCCGGCACCAAVGGCACCAAGAGAVVCGACAACCCCGVGCVGCCCVVCAACGACGGGGVGVACVVVGCCAGCACCGAGAAGVCCAACAVCAVCAGAG	360
361	GCVGGAVCVVCGGCACCACACVGGACAGCAAGACCCAGAGCCVGCVGAVCGVGAACAACGCCACCAACGVGGVCAVCAAAGVGVGCGAGVVCCAGVVCVGCAACGACCCCVVCCVGGGCG	480
481	VCVACVACCACAAGAACAACAAGAGCVGGAVGGAAAGCGAGVVCCGGGVGVACAGCAGCGCCAACAACVGCACCVVCGAGVACGVGVCCCAGCCVVVCCVGAVGGACCVGGAAGGCAAGC	600
601	AGGGCAACVVCAAGAACCVGCGCGAGVVCGVGVVVAAGAACAVCGACGGCVACVVCAAGAVCVACAGCAAGCACACCCCVAVCAACCVCGVGCGGGAVCVGCCVCAGGGCVVCVCVGCVC	720
721	VGGAACCCCVGGVGGAVCVGCCCAVCGGCAVCAACAVCACCCGGVVVCAGACACVGCVGGCCCVGCACAGAAGCVACCVGACACCVGGCGAVAGCAGCAGCGGAVGGACAGCVGGVGCCG	840
841	CCGCVVACVAVGVGGGCVACCVGCAGCCVAGAACCVVCCVGCVGAAGVACAACGAGAACGGCACCAVCACCGACGCCGVGGAVVGVGCVCVGGAVCCVCVGAGCGAGACAAAGVGCACCC	960
961	VGAAGVCCVVCACCGVGGAAAAGGGCAVCVACCAGACCAGCAACVVCCGGGVGCAGCCCACCGAAVCCAVCGVGCGGVVCCCCAAVAVCACCAAVCVGVGCCCCVVCGGCGAGGVGVVCA	1080
1081	AVGCCACCAGAVVCGCCVCVGVGVACGCCVGGAACCGGAAGCGGAVCAGCAAVVGCGVGGCCGACVACVCCGVGCVGVACAACVCCGCCAGCVVCAGCACCVVCAAGVGCVACGGCGVGV	1200
1201	CCCCVACCAAGCVGAACGACCVGVGCVVCACAAACGVGVACGCCGACAGCVVCGVGAVCCGGGGAGAVGAAGVGCGGCAGAVVGCCCCVGGACAGACAGGCAAGAVCGCCGACVACAACV	1320
1321	ACAAGCVGCCCGACGACVVCACCGGCVGVGVGAVVGCCVGGAACAGCAACAACCVGGACVCCAAAGVCGGCGGCAACVACAAVVACCVGVACCGGCVGVVCCGGAAGVCCAAVCVGAAGC	1440
1441	CCVVCGAGCGGGACAVCVCCACCGAGAVCVAVCAGGCCGGCAGCACCCCVVGVAACGGCGVGGAAGGCVVCAACVGCVACVVCCCACVGCAGVCCVACGGCVVVCAGCCCACAAAVGGCG	1560
1561	VGGGCVAVCAGCCCVACAGAGVGGVGGVGCVGAGCVVCGAACVGCVGCAVGCCCCVGCCACAGVGVGCGGCCCVAAGAAAAGCACCAAVCVCGVGAAGAACAAAVGCGVGAACVVCAACV	1680
1681	VCAACGGCCVGACCGGCACCGGCGVGCVGACAGAGAGCAACAAGAAGVVCCVGCCAVVCCAGCAGVVVGGCCGGGAVAVCGCCGAVACCACAGACGCCGVVAGAGAVCCCCAGACACVGG	1800
1801	AAAVCCVGGACAVCACCCCVVGCAGCVVCGGCGGAGVGVCVGVGAVCACCCCVGGCACCAACACCAGCAAVCAGGVGGCAGVGCVGVACCAGGACGVGAACVGVACCGAAGVGCCCGVGG	1920
1921	CCAVVCACGCCGAVCAGCVGACACCVACAVGGCGGGVGVACVCCACCGGCAGCAAVGVGVVVCAGACCAGAGCCGGCVGVCVGAVCGGAGCCGAGCACGVGAACAAVAGCVACGAGVGCG	2040
2041	ACAVCCCCAVCGGCGCVGGAAVCVGCGCCAGCVACCAGACACAGACAAACAGCCCVCGGAGAGCCAGAAGCGVGGCCAGCCAGAGCAVCAVVGCCVACACAAVGVCVCVGGGCGCCGAGA	2160
2161	ACAGCGVGGCCVACVCCAACAACVCVAVCGCVAVCCCCACCAACVVCACCAVCAGCGVGACCACAGAGAVCCVGCCVGVGVCCAVGACCAAGACCAGCGVGGACVGCACCAVGVACAVCV	2280
2281	GCGGCGAVVCCACCGAGVGCVCCAACCVGCVGCVGCAGVACGGCAGCVVCVGCACCCAGCVGAAVAGAGCCCVGACAGGGAVCGCCGVGGAACAGGACAAGAACACCCAAGAGGVGVVCG	2400
2401	CCCAAGVGAAGCAGAVCVACAAGACCCCVCCVAVCAAGGACVVCGGCGGCVVCAAVVVCAGCCAGAVVCVGCCCGAVCCVAGCAAGCCCAGCAAGCGGAGCVVCAVCGAGGACCVGCVGV	2520
2521	VCAACAAAGVGACACVGGCCGACGCCGGCVVCAVCAAGCAGVAVGGCGAVVGVCVGGGCGACAVVGCCGCCAGGGAVCVGAVVVGCGCCCAGAAGVVVAACGGACVGACAGVGCVGCCVC	2640
2641	CVCVGCVGACCGAVGAGAVGAVCGCCCAGVACACAVCVGCCCVGCVGGCCGGCACAAVCACAAGCGGCVGGACAVVVGGAGCAGGCGCCGCVCVGCAGAVCCCCVVVGCVAVGCAGAVGG	2760
2761	CCVACCGGVVCAACGGCAVCGGAGVGACCCAGAAVGVGCVGVACGAGAACCAGAAGCVGAVCGCCAACCAGVVCAACAGCGCCAVCGGCAAGAVCCAGGACAGCCVGAGCAGCACAGCAA	2880
2881	GCGCCCVGGGAAAGCVGCAGGACGVGGVCAACCAGAAVGCCCAGGCACVGAACACCCVGGVCAAGCAGCVGVCCVCCAACVVCGGCGCCAVCAGCVCVGVGCVGAACGAVAVCCVGAGCA	3000
3001	GACVGGACCCVCCVGAGGCCGAGGVGCAGAVCGACAGACVGAVCACAGGCAGACVGCAGAGCCVCCAGACAVACGVGACCCAGCAGCVGAVCAGAGCCGCCGAGAVVAGAGCCVCVGCCA	3120
3121	AVCVGGCCGCCACCAAGAVGVCVGAGVGVGVGCVGGGCCAGAGCAAGAGAGVGGACVVVVGCGGCAAGGGCVACCACCVGAVGAGCVVCCCVCAGVCVGCCCCVCACGGCGVGGVGVVVC	3240
3241	VGCACGVGACAVAVGVGCCCGCVCAAGAGAAGAAVVVCACCACCGCVCCAGCCAVCVGCCACGACGGCAAAGCCCACVVVCCVAGAGAAGGCGVGVVCGVGVCCAACGGCACCCAVVGGV	3360
3361	VCGVGACACAGCGGAACVVCVACGAGCCCCAGAVCAVCACCACCGACAACACCVVCGVGVCVGGCAACVGCGACGVCGVGAVCGGCAVVGVGAACAAVACCGVGVACGACCCVCVGCAGC	3480
3481	CCGAGCVGGACAGCVVCAAAGAGGAACVGGACAAGVACVVVAAGAACCACACAAGCCCCGACGVGGACCVGGGCGAVAVCAGCGGAAVCAAVGCCAGCGVCGVGAACAVCCAGAAAGAGA	3600
3601	VCGACCGGCVGAACGAGGVGGCCAAGAAVCVGAACGAGAGCCVGAVCGACCVGCAAGAACVGGGGAAGVACGAGCAGVACAVCAAGVGGCCCVGGVACAVCVGGCVGGGCVVVAVCGCCG	3720
3721	GACVGAVVGCCAVCGVGAVGGVCACAAVCAVGCVGVGVVGCAVGACCAGCVGCVGVAGCVGCCVGAAGGGCVGVVGVAGCVGVGGCAGCVGCVGCAAGVVCGACGAGGACGAVVCVGAGC	3840
3841	CCGVGCVGAAGGGCGVGAAACVGCACVACACAVGAVGACVCGAGCVGGVACVGCAVGCACGCAAVGCVAGCVGCCCCVVVCCCGVCCVGGGVACCCCGAGVCVCCCCCGACCVCGGGVCC	3960
3961	CAGGVAVGCVCCCACCVCCACCVGCCCCACVCACCACCVCVGCVAGVVCCAGACACCVCCCAAGCACGCAGCAAVGCAGCVCAAAACGCVVAGCCVAGCCACACCCCCACGGGAAACAGC	4080
4081	AGVGAVVAACCVVVAGCAAVAAACGAAAGVVVAACVAAGCVAVACVAACCCCAGGGVVGGVCAAVVVCGVGCCAGCCACACCCVGGAGCVAGCAAAAAAAAAAAAAAAAAAAAAAAAAAA	4200
4201	AAAGCAVAVGACVAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	4283

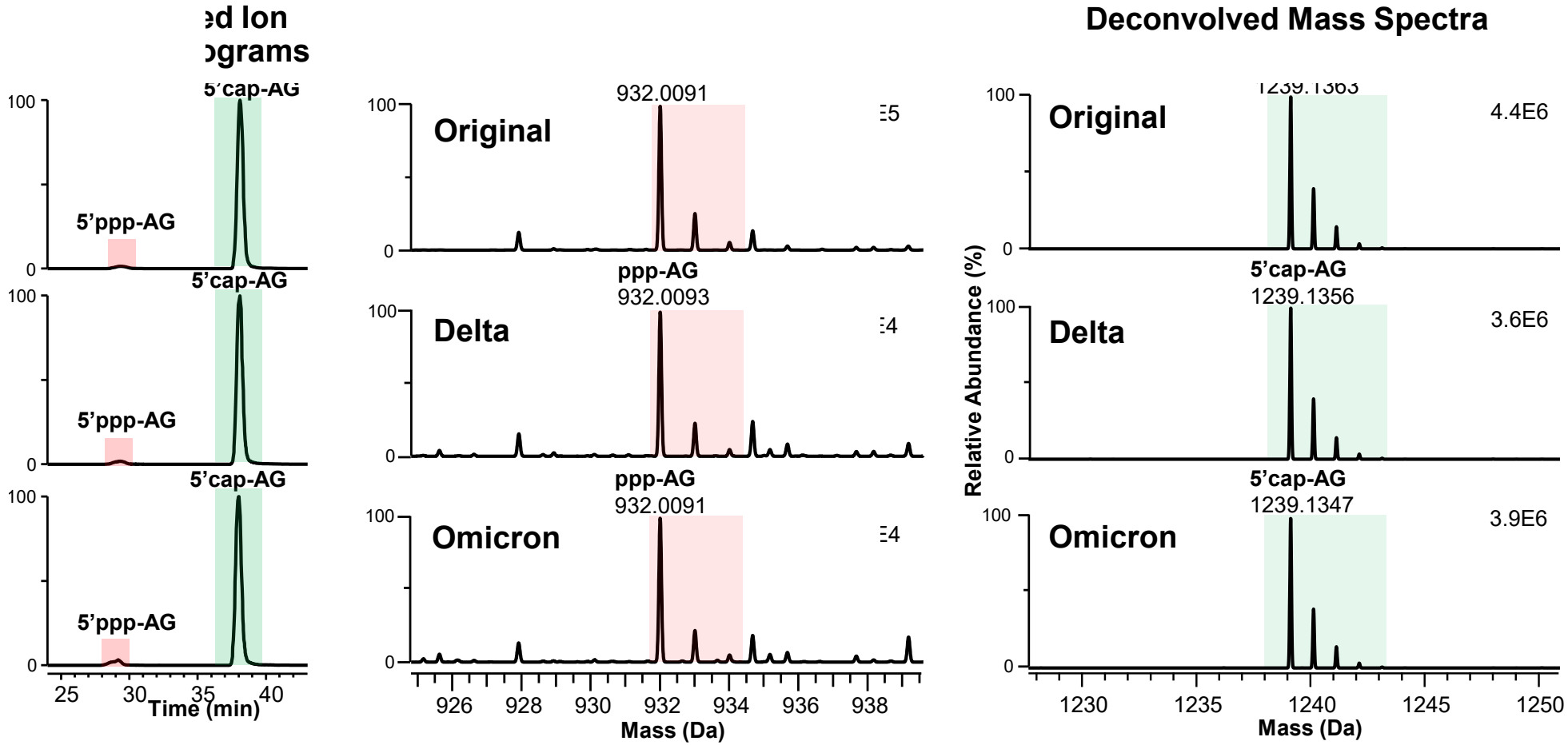
V = N1-methyl pseudouridine

Oligonucleotide Mapping Enables Simultaneous Characterization of the 5′ Terminus Without Affinity Purification



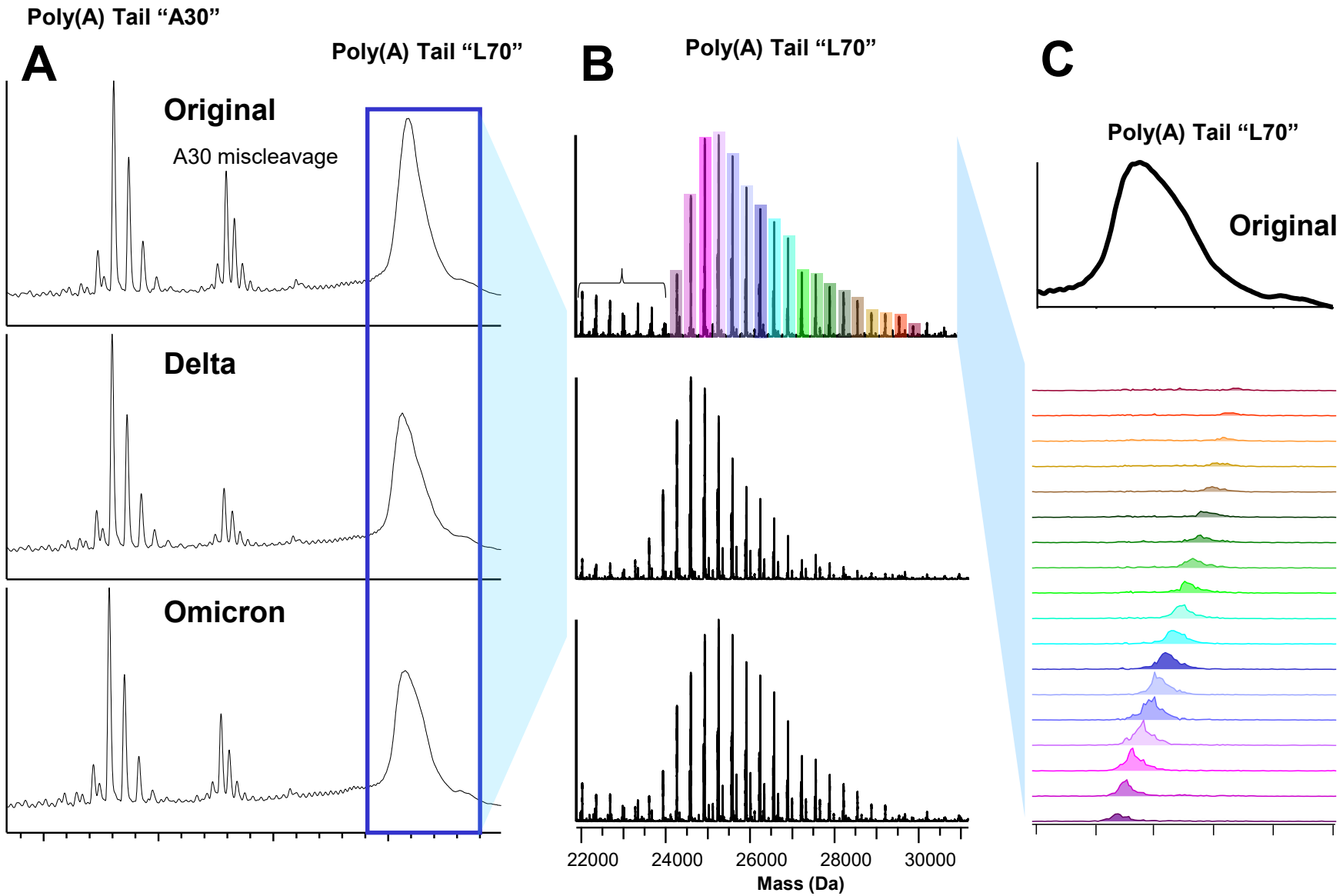


C Capped 5′ terminus



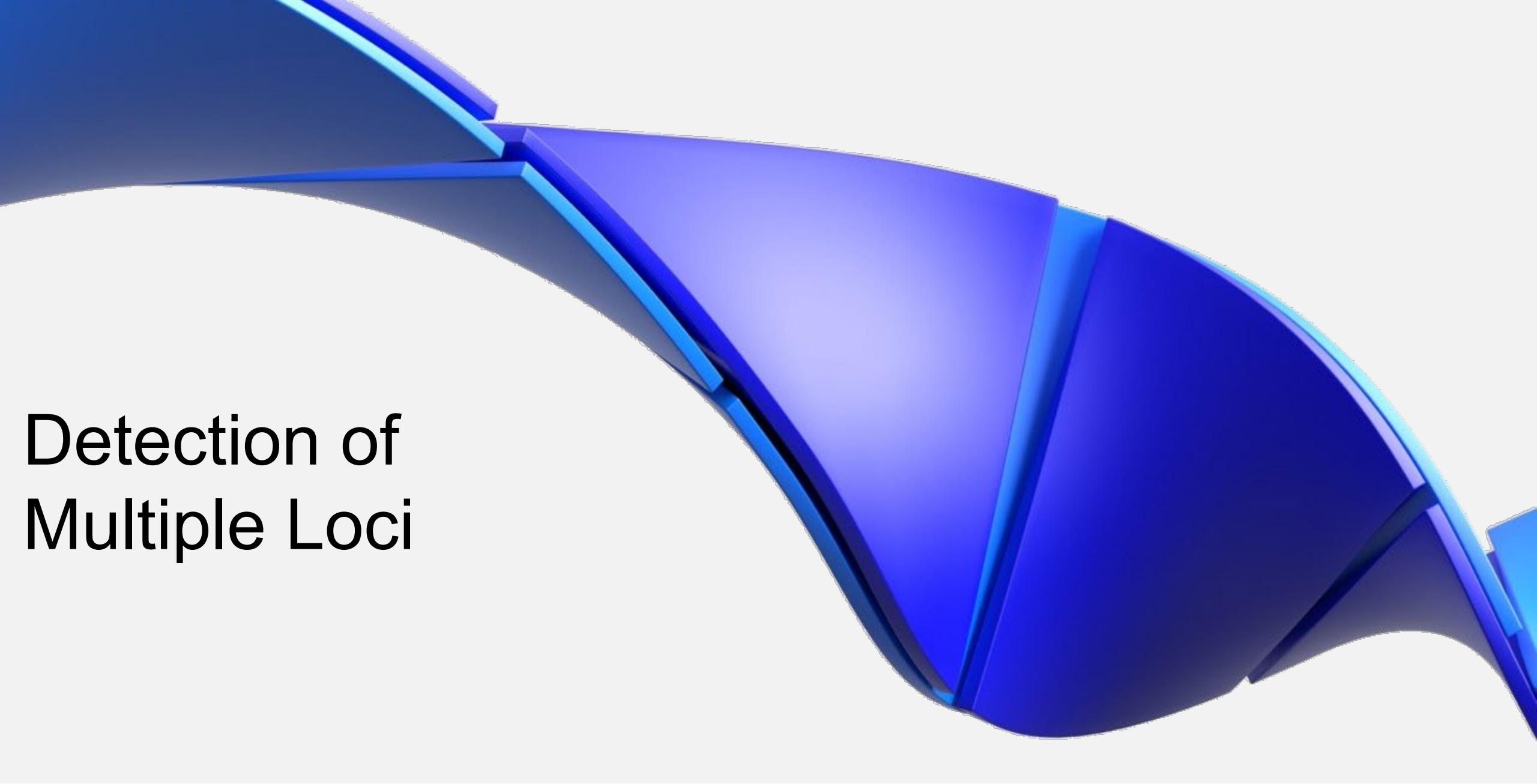
Majority of 5′ terminus is capped

Oligonucleotide Mapping of mRNA Enables Simultaneous Characterization of the 3´ Terminus Without Affinity Purification



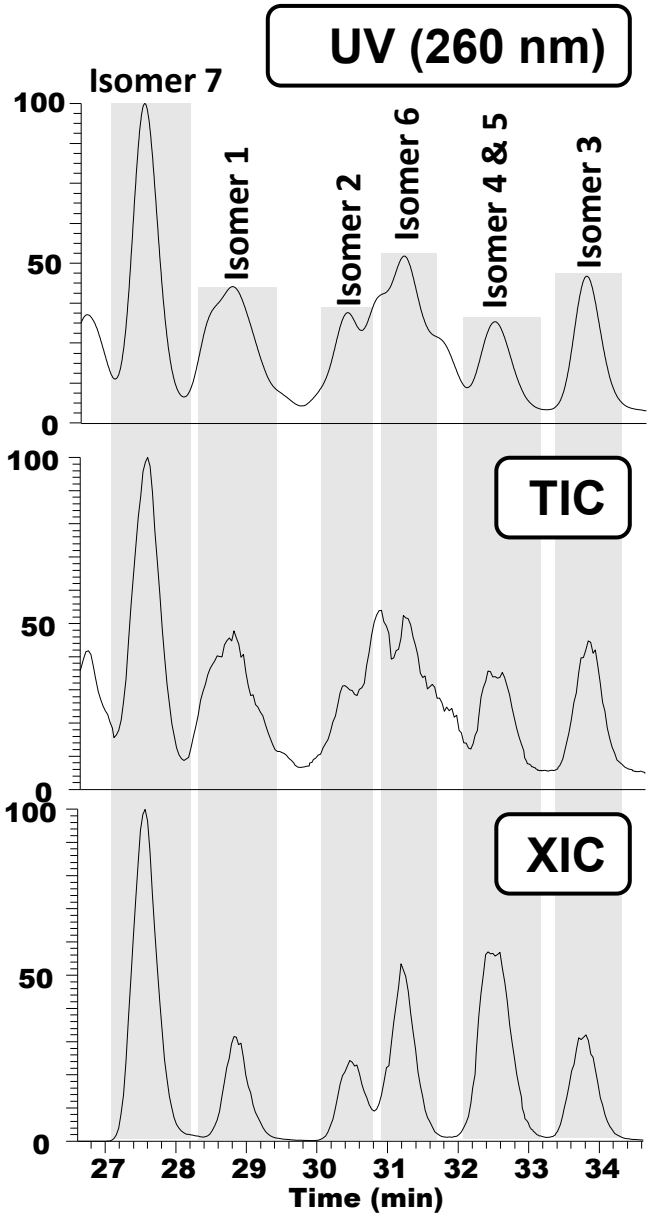
Poly(A) Tail Distribution Characterized

Moderate Shifts in Poly(A) Length Distribution Are Detectable



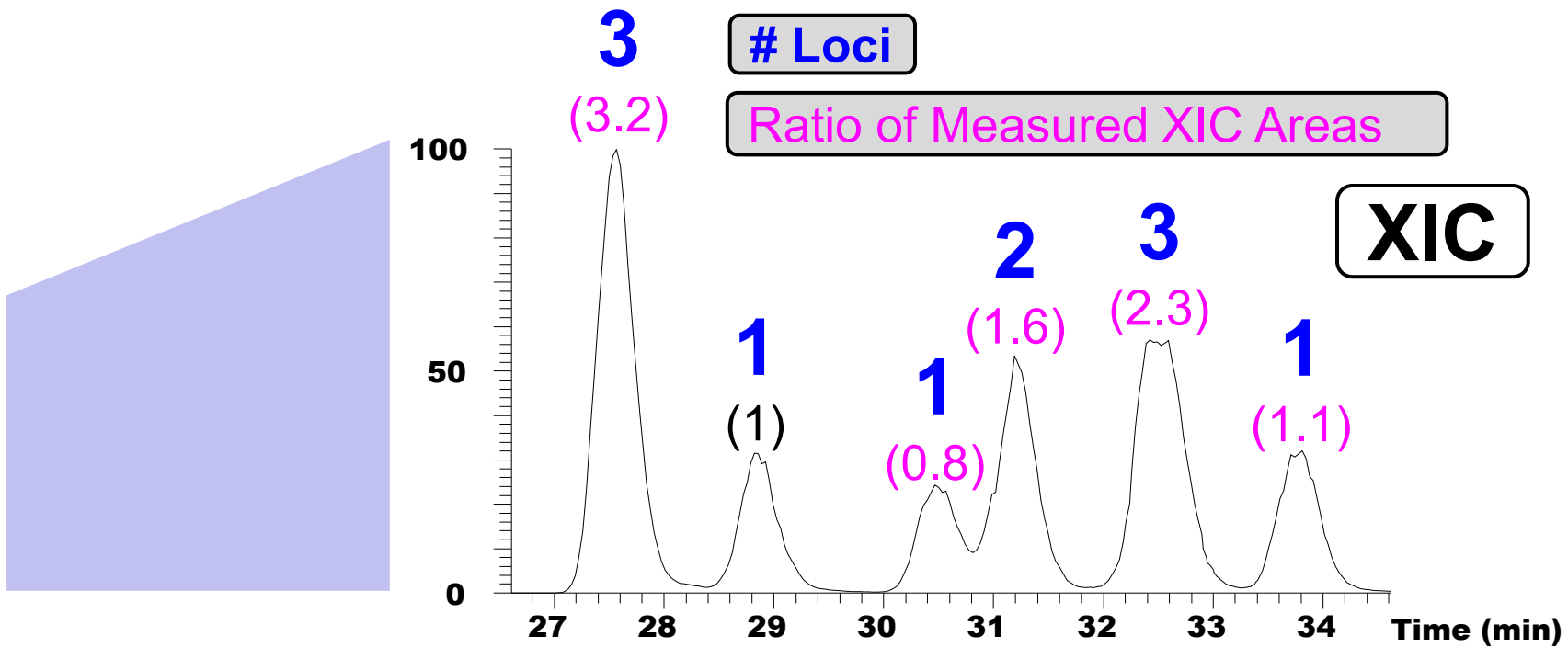
Detection of Multiple Loci

Measured XIC Areas of Non-Unique Sequence Isomers Correlate with their Number of Loci in the Full Length mRNA Sequence



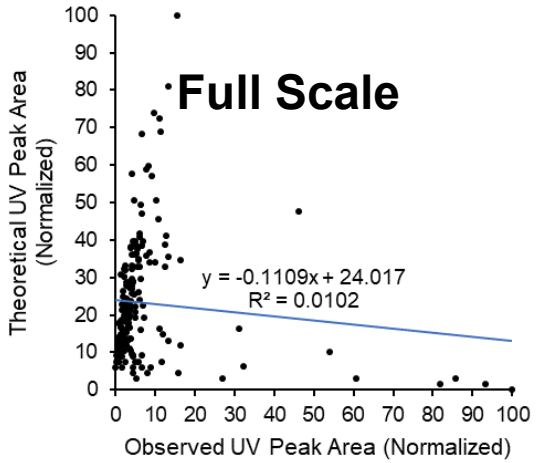
Theoretical				Observed	
Oligonucleotide	Sequence	# of Loci	Monoisotopic Mass (Da)	Retention Time (min)	Ratio of XIC Areas Normalized to Isomer 1
Isomer 1	CVAAG	1	1646.2453	28.9	1
Isomer 2	VCAAG	1		30.5	0.8
Isomer 3	AVCAG	1		33.8	1.1
Isomer 4	AVACG	1		32.5	2.3
Isomer 5	AACVG	2			
Isomer 6	VACAG	2		31.2	1.6
Isomer 7	CAAVG	3		27.6	3.2

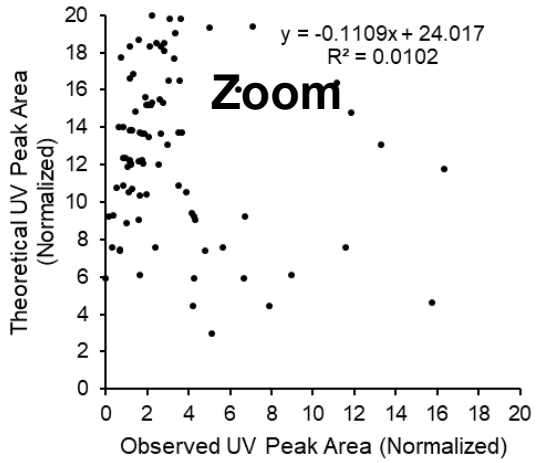
V = N1-methyl pseudouridine



Measured UV Areas Across Oligonucleotide Map Correlate with Theoretical UV Areas but only when Accounting for All Loci of Non-Unique Oligonucleotides

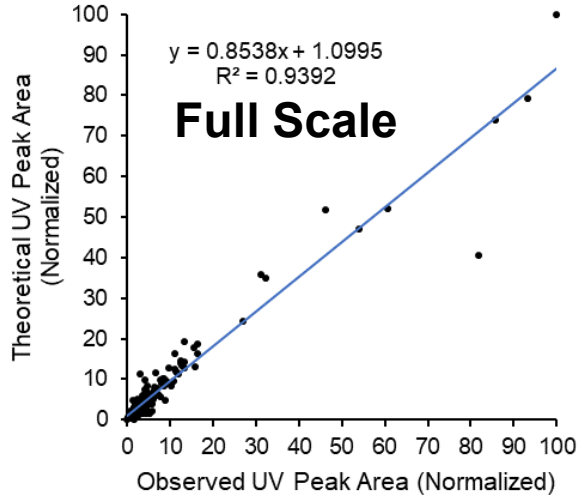
Not Accounting for Multiple Loci of Non-Unique Oligonucleotides

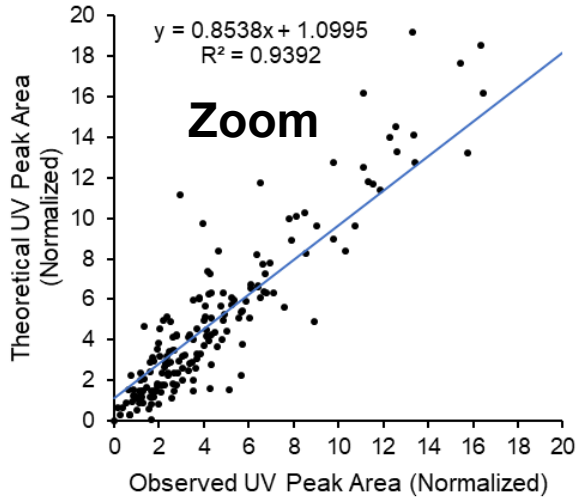




Poor Correlation

Accounting for Multiple Loci of Non-Unique Oligonucleotides



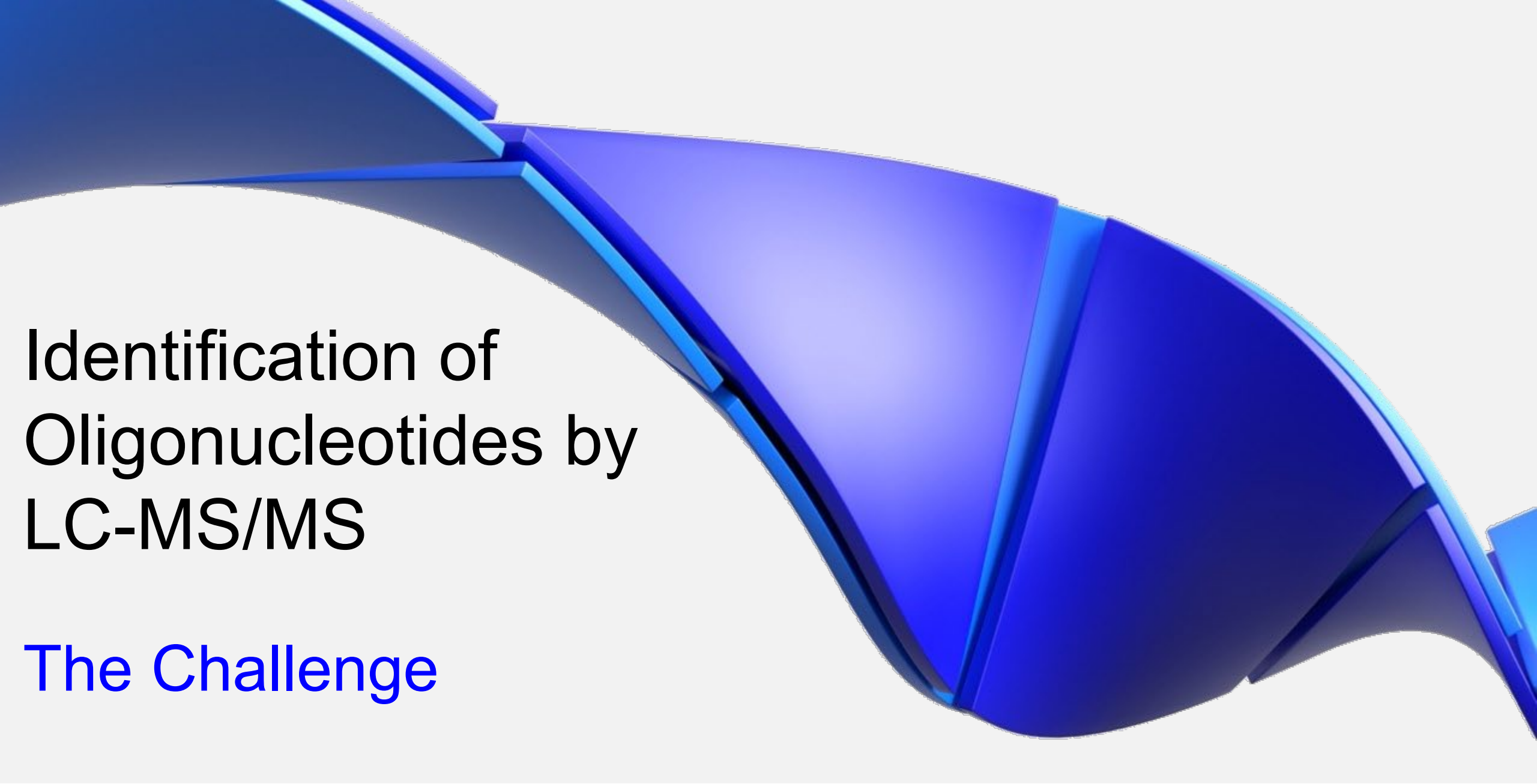


Good Correlation

Theoretical peak areas were calculated using oligonucleotide compositions and NMR-derived extinction coefficients for pdG*, pdA*, pdC*, and N1-methylpseudouridine monophosphate¥

*Cavaluzzi MJ, Borer PN. Nucleic Acids Res. 2004 Jan 13;32(1):e13.

¥Emperically determined at Pfizer

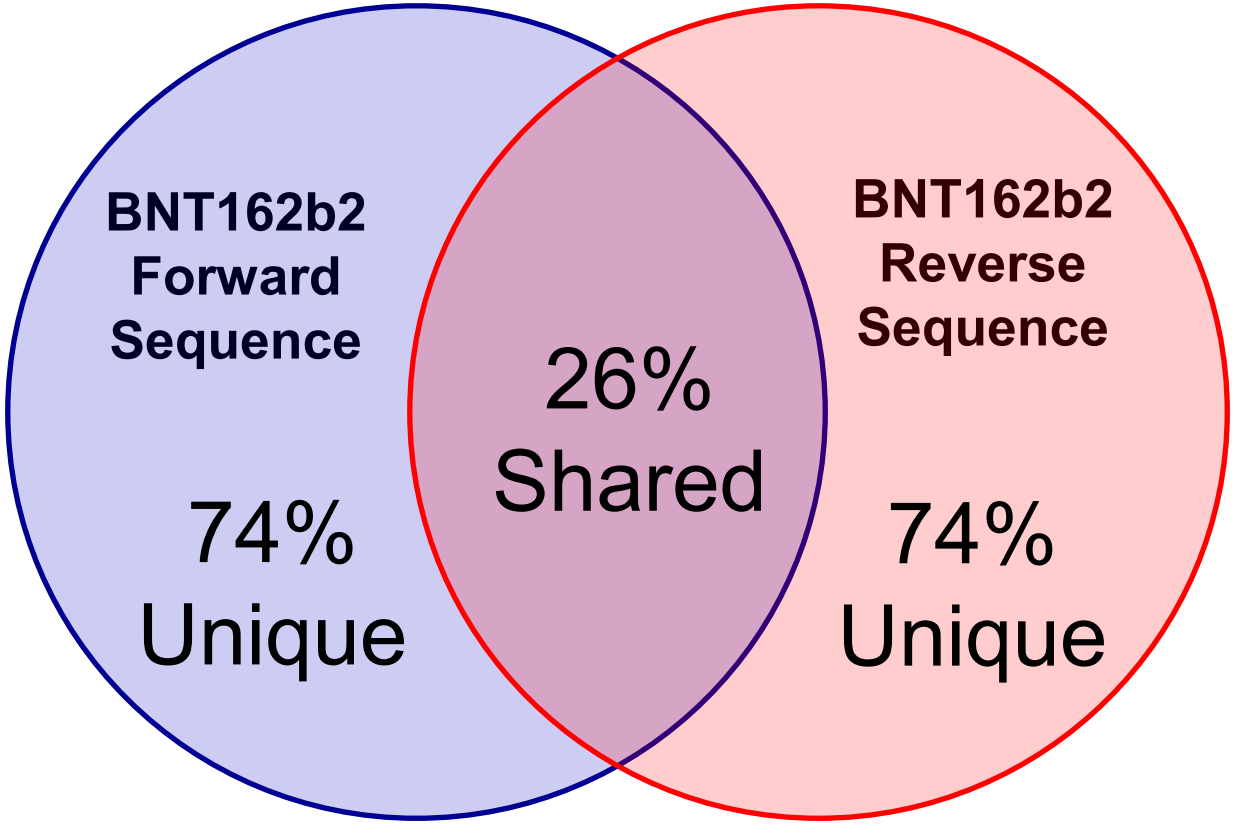


Identification of Oligonucleotides by LC-MS/MS

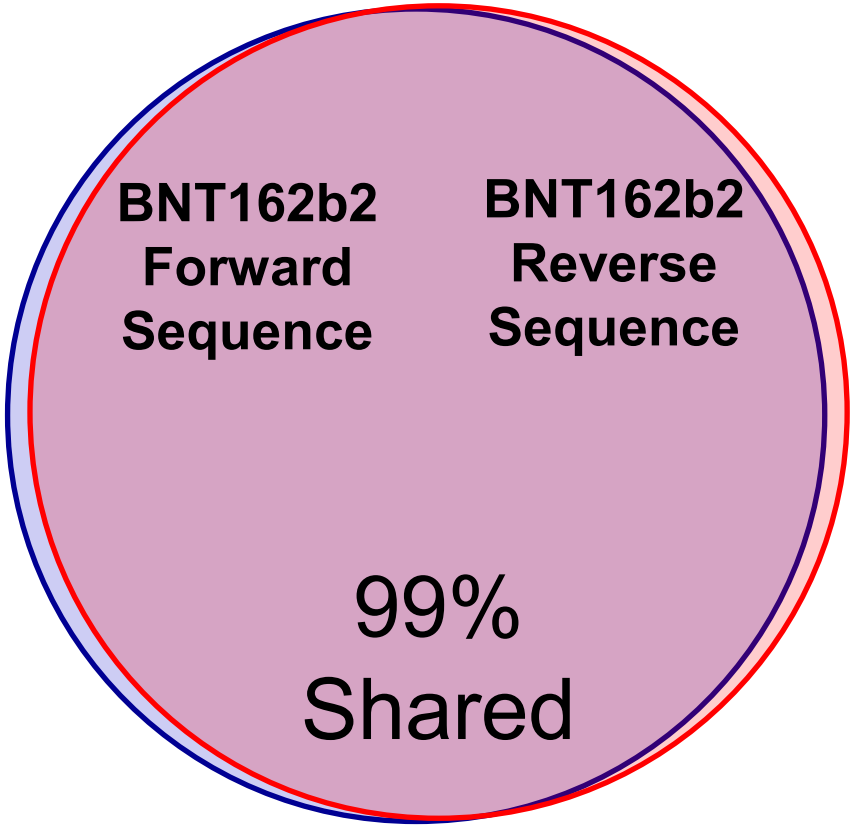
The Challenge

MS/MS Fragmentation is Critical For Differentiating Oligonucleotide Sequence Isomers Generated Through Enzymatic Digestion by RNaseT1

Most Oligonucleotide Sequences are Unique



Most Oligonucleotide Masses are Not Unique



Identification of Oligonucleotides by LC-MS/MS

HCD Fragmentation Study

Higher Energy Collision Dissociation (HCD) Parameters Optimized to Generate Fragmentation Appropriate for Oligonucleotide Mapping

Fragmentation of RNA Oligonucleotides is Complex

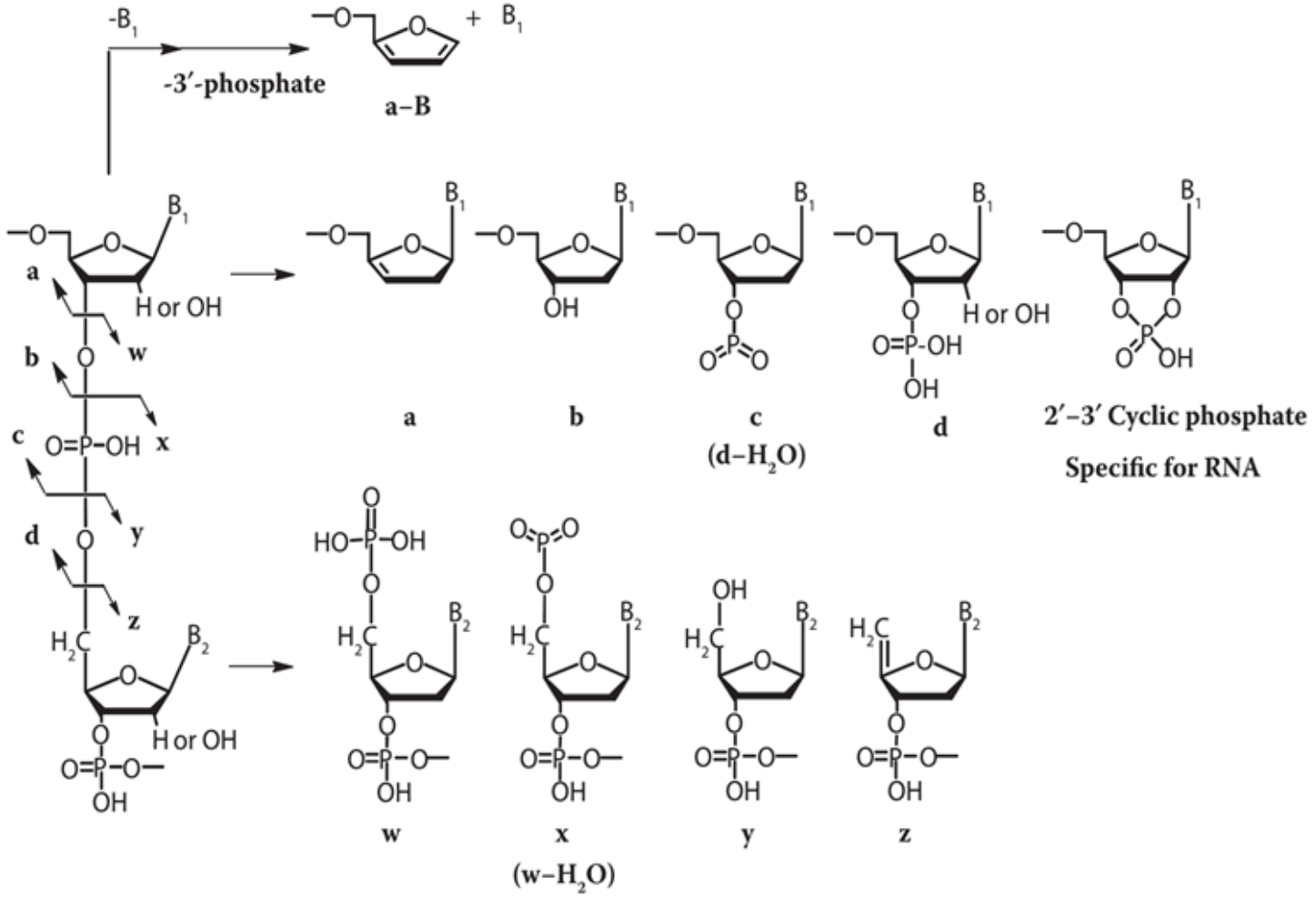
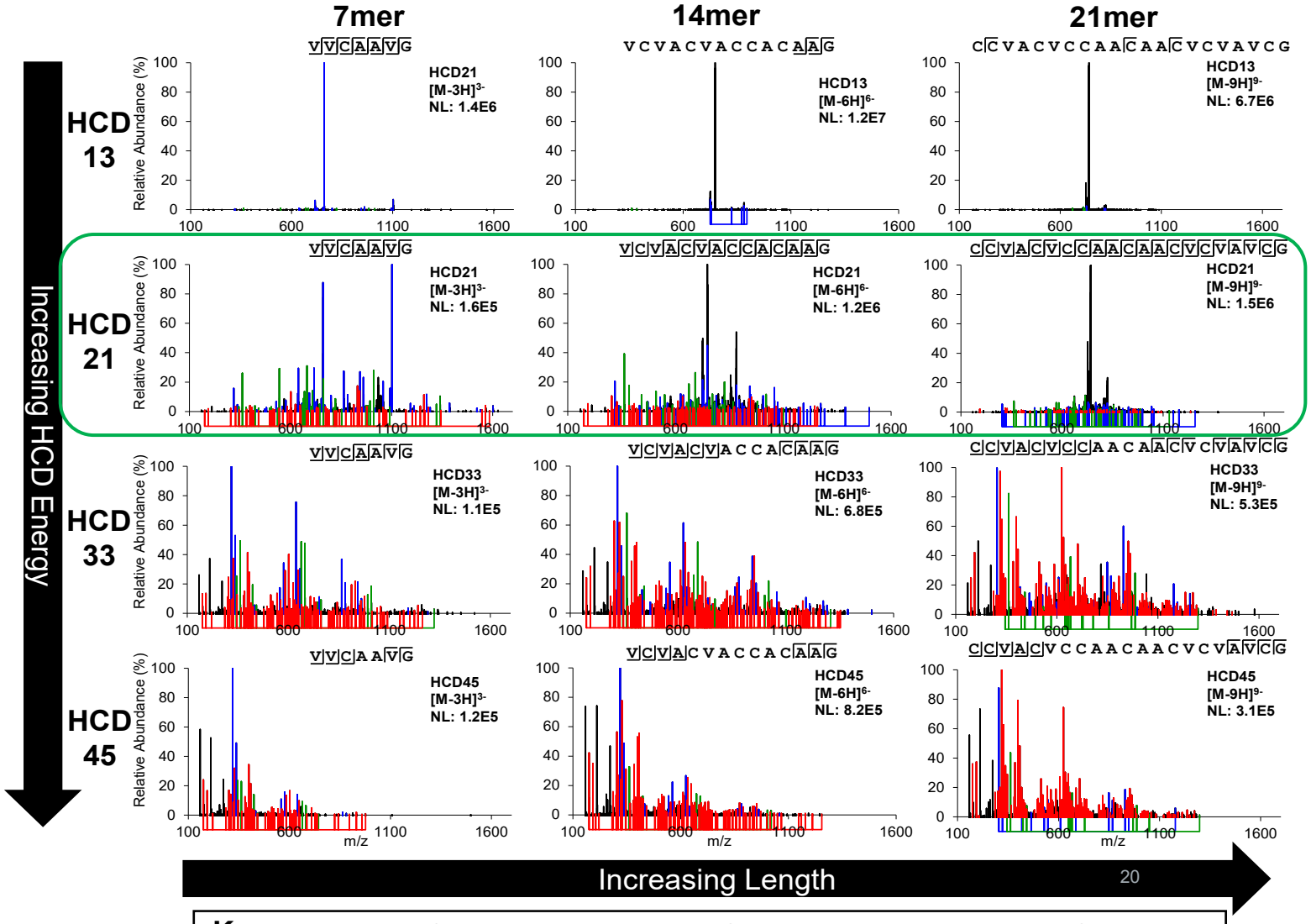


Figure adapted from McLuckey SA, Van Berkel GJ, Glish GL. Tandem mass spectrometry of small, multiply charged oligonucleotides. J Am Soc Mass Spectrom. 1992 Jan;3(1):60-70.

HCD Collision Energy Optimized at Stepped CE 17, 21, 25

Oligonucleotide Fragment Ion Coverage as a Function of HCD Energy & Length



Increasing Length

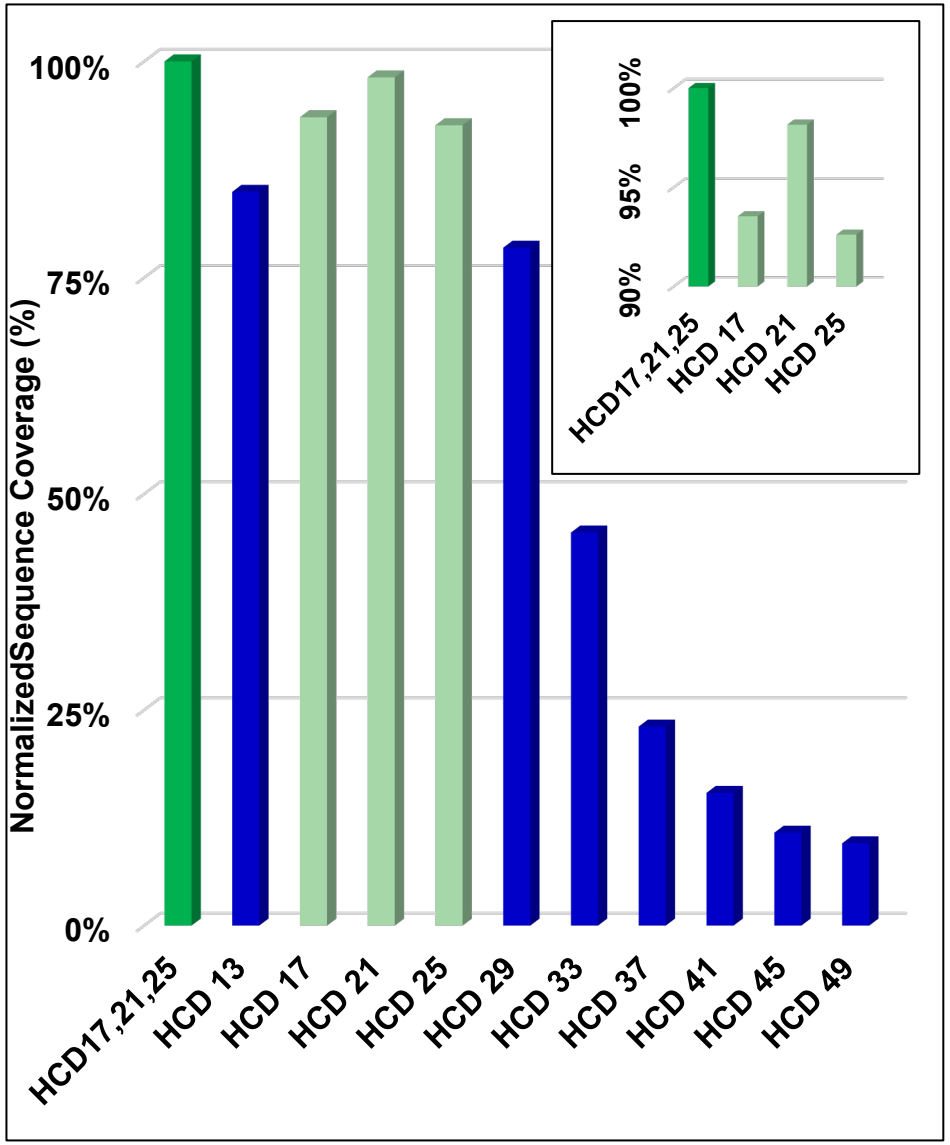
20

Key — 5′ fragment ion — 3′ fragment ion — Internal fragment ion

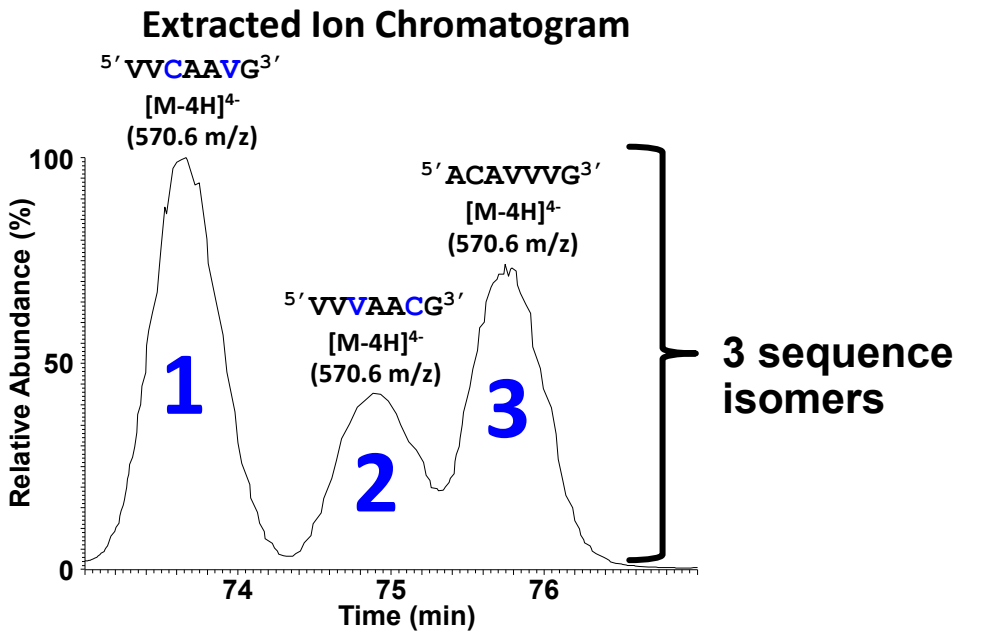
Charge densities are fixed at 0.4 charge / base

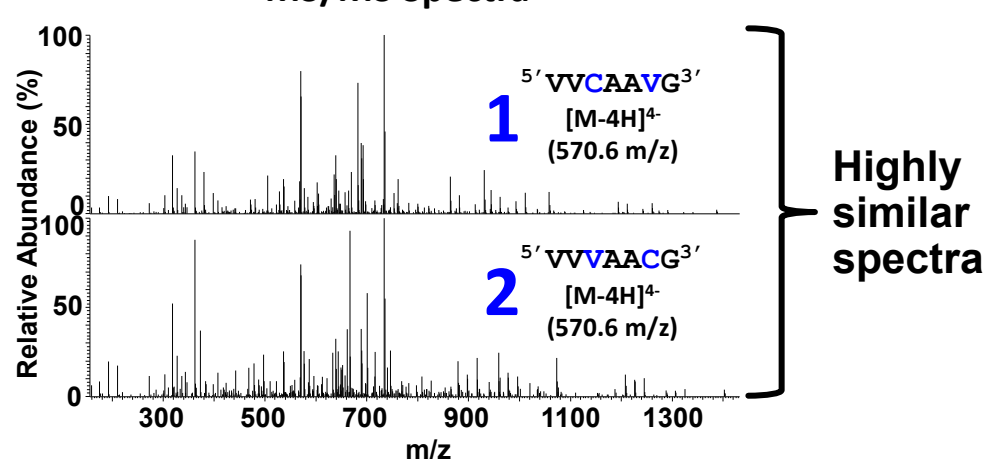
V = N1-methyl pseudouridine

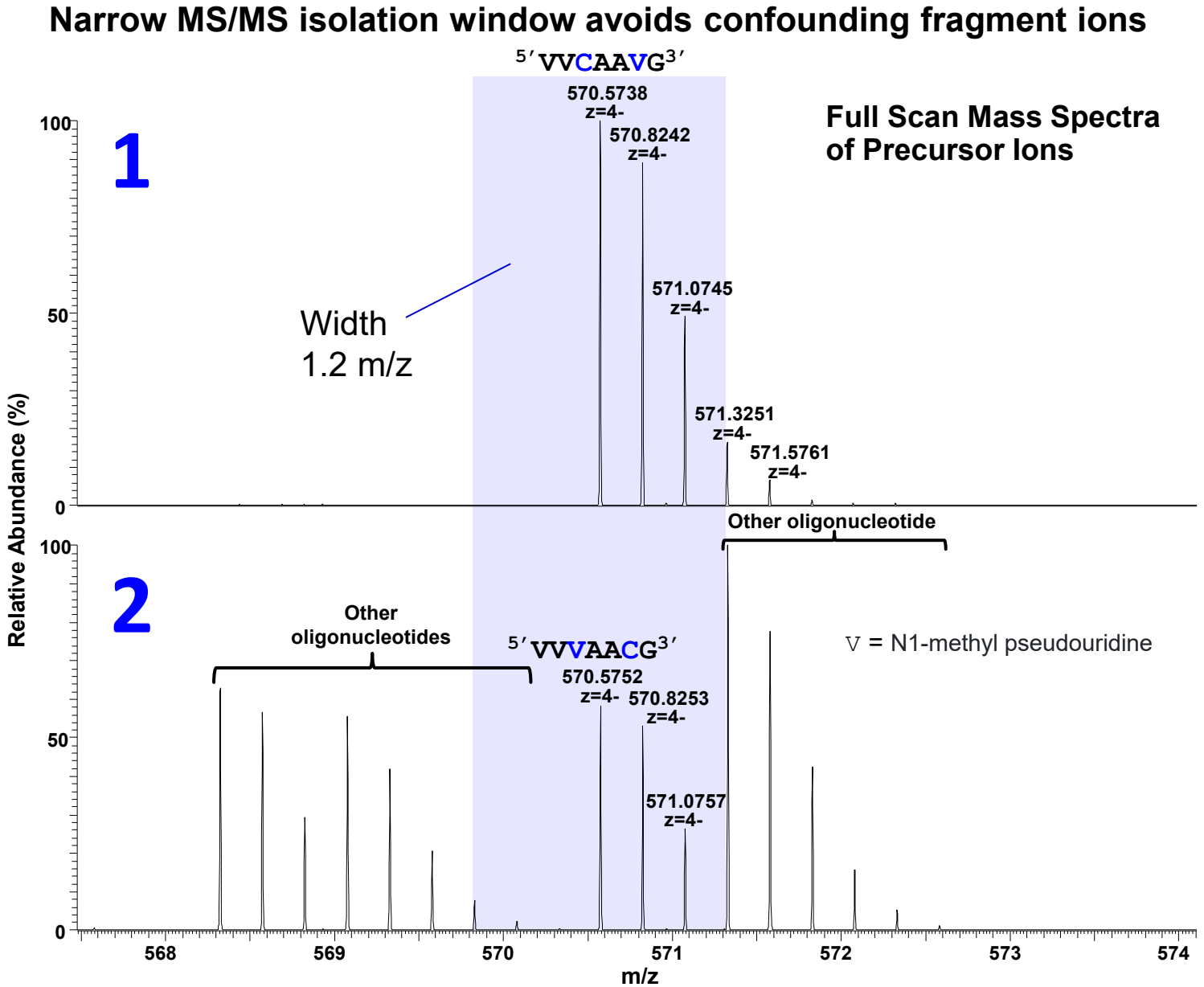
Normalized BNT162b2 Sequence Coverage Across All Oligonucleotides as a Function of HCD Energy



Applying Optimized HCD to Differentiate 2 Sequence Isomers Differing by a Single Exchange in Base Positions

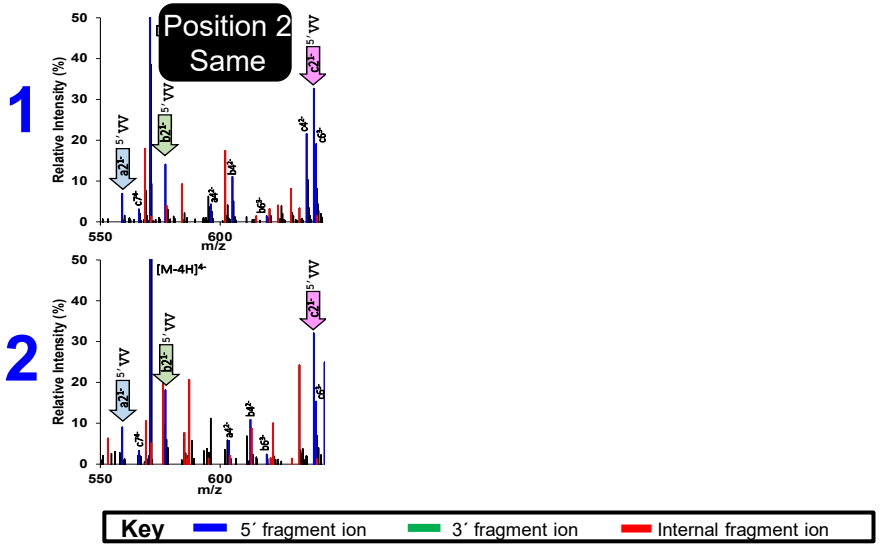






Optimal Fragmentation Enables Differentiation of Highly Similar Sequence Isomers

Observed 5′ MS/MS fragments



1 5′VVCAAVG3′ [M-4H]$^{4-}$

V = N1-methyl pseudouridine

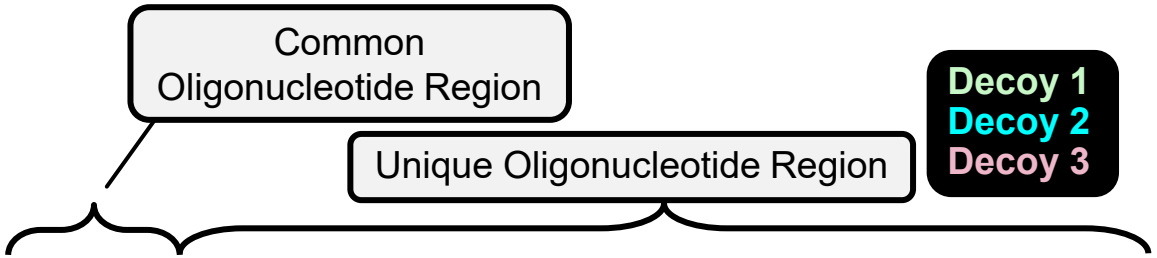
Observed 5′ fragments					
a	b	c	d	#	
		(319.0)$^{1-}$	(337.0)$^{1-}$	1	V
(559.1)$^{1-}$	(577.1)$^{1-}$	(639.1)$^{1-}$	(657.1)$^{1-}$	2	V

2 5′VVVAACG3′ [M-4H]$^{4-}$

Observed 5′ fragments					
a	b	c	d	#	
		(319.0)$^{1-}$	(337.0)$^{1-}$	1	V
(559.1)$^{1-}$	(577.1)$^{1-}$	(639.1)$^{1-}$		2	V

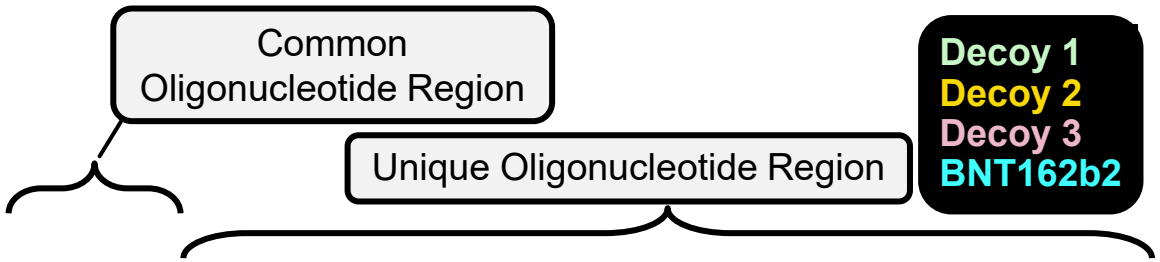
Fidelity of Automated Oligonucleotide Identifications Comprehensively Verified by Decoy Searching

Decoy search excluding BNT162b2 mRNA construct



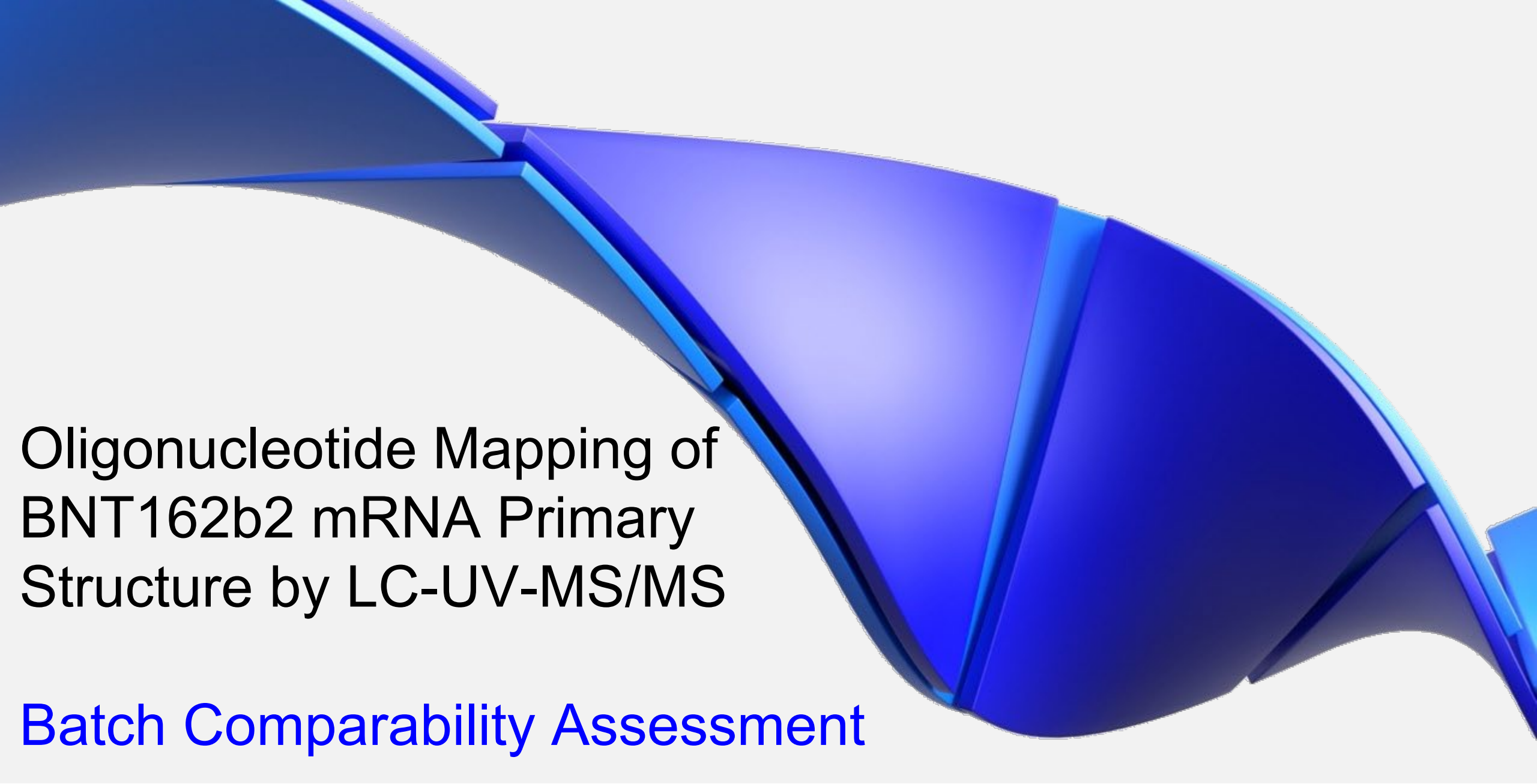
No Preferential Match

Decoy search including BNT162b2 mRNA construct



Preferential Match to BNT162b2

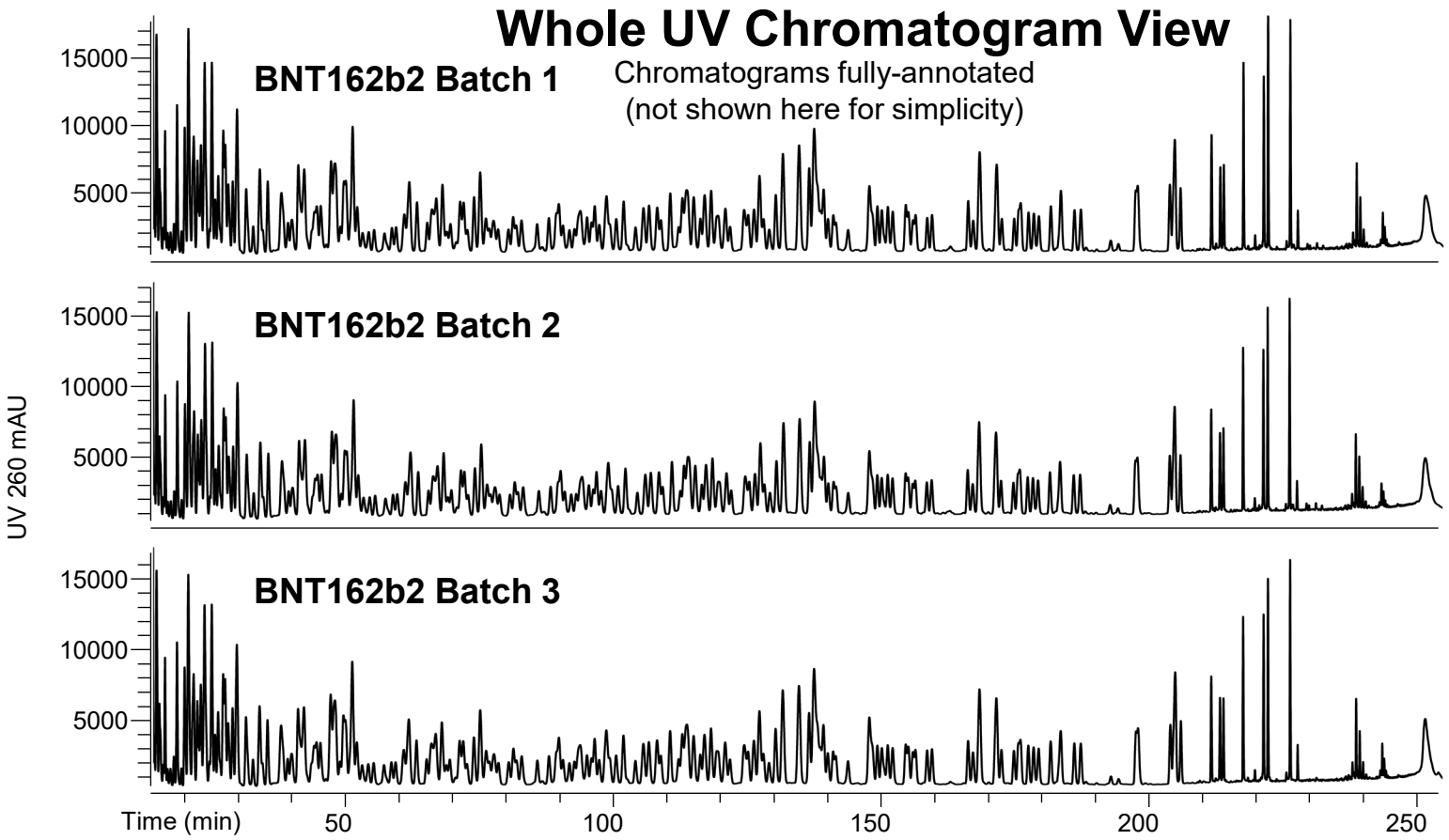
Pfizer

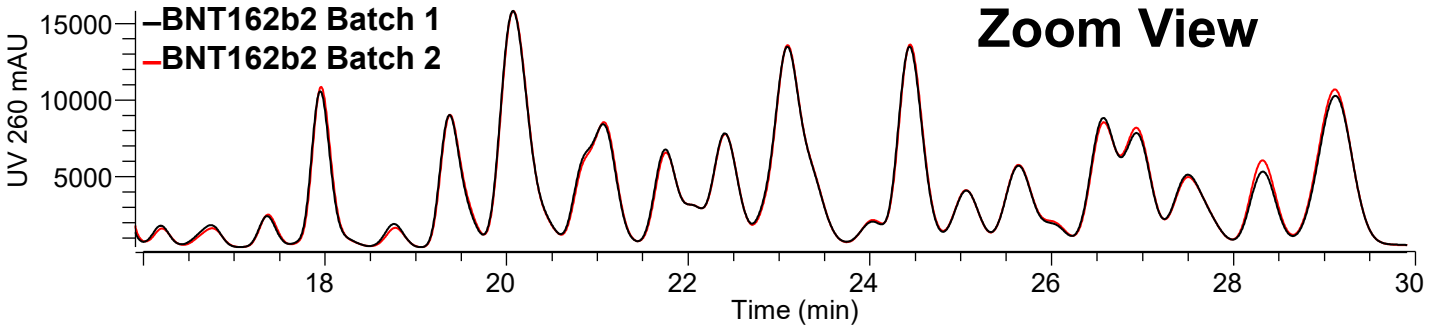


Oligonucleotide Mapping of BNT162b2 mRNA Primary Structure by LC-UV-MS/MS

Batch Comparability Assessment

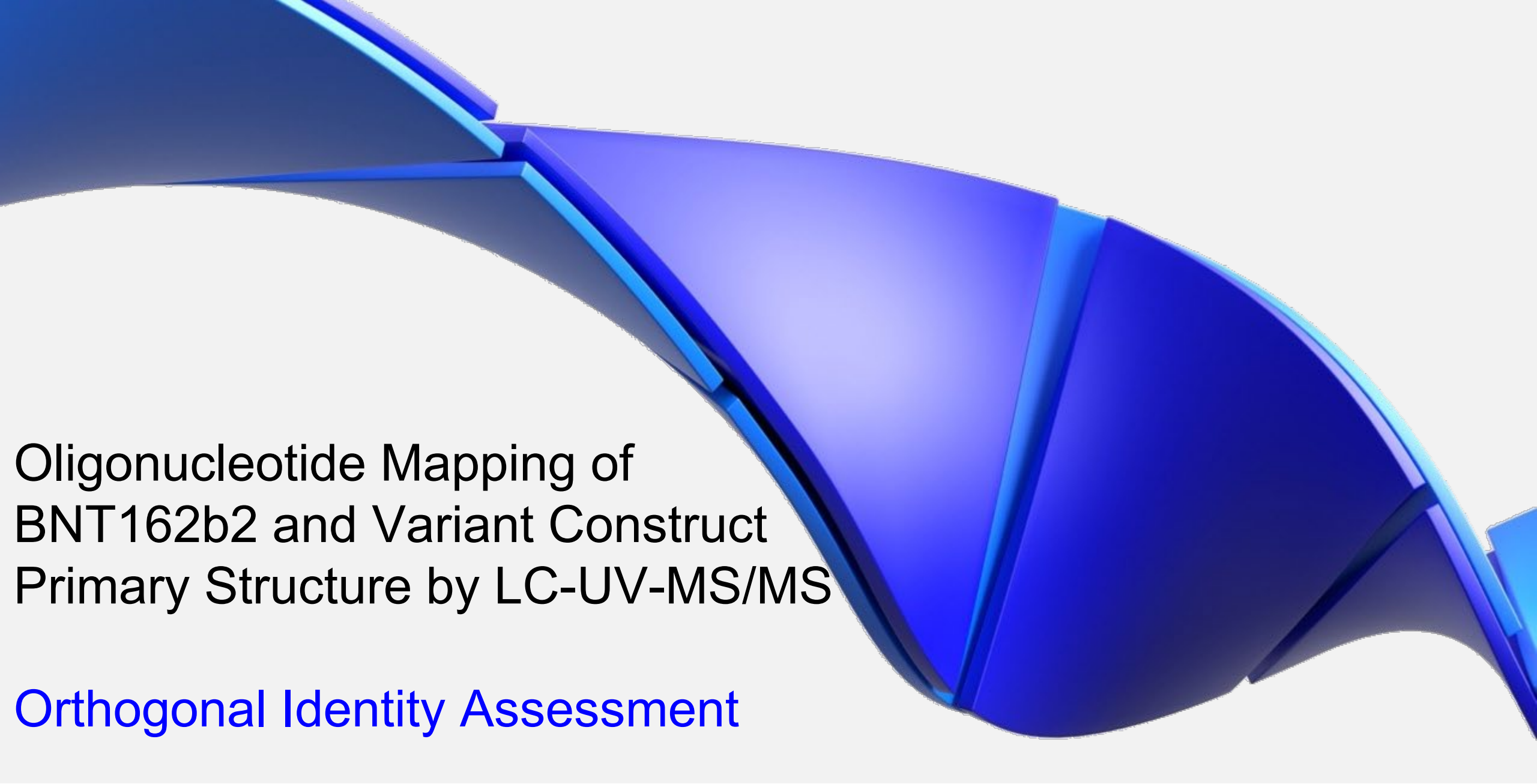
Oligonucleotide Mapping Enables Assessment of mRNA Batch Comparability





Oligonucleotide Maps Demonstrate Comparability of Multiple BNT162b2 mRNA Drug Substance Batches

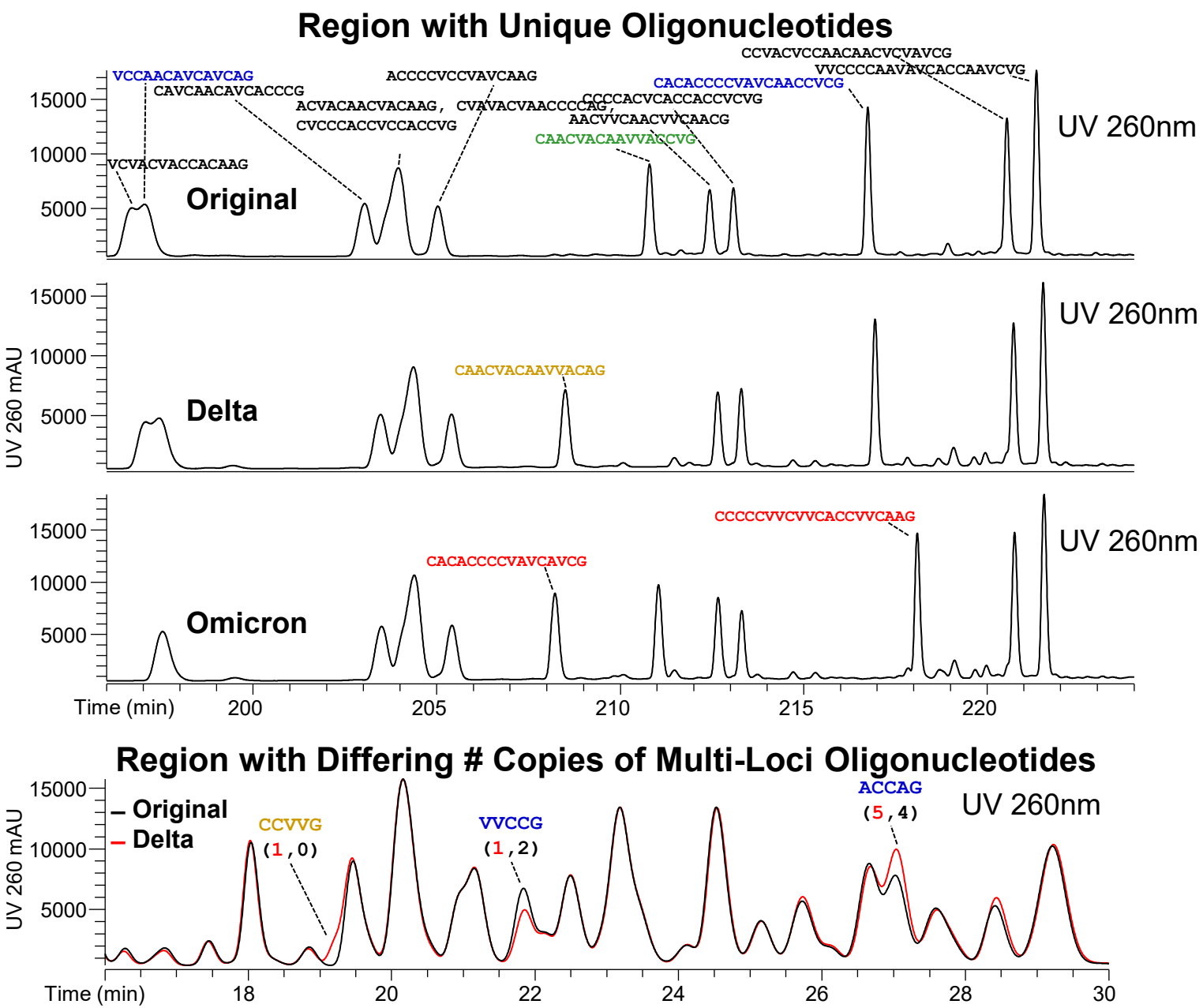
- Side-by-side analyses are highly robust
- Chromatographic peaks overlay well



Oligonucleotide Mapping of BNT162b2 and Variant Construct Primary Structure by LC-UV-MS/MS

Orthogonal Identity Assessment

Oligonucleotide Mapping Enables Comparison of mRNA for Variant Constructs



Key

Present in Original, Delta, & Omicron
Present in Original & Delta
Present in Original & Omicron
Present in Delta
Present in Omicron

Oligonucleotide mapping detects:

- Oligonucleotides unique to one or more constructs
- Differences in number of copies of multi-loci (non-unique) oligonucleotides

Conclusion

- Oligonucleotide mapping via LC-UV-MS/MS directly interrogates the primary structure of RNA, enabling enhanced structural understanding for mRNA vaccines, genetic therapies, and other RNA molecules

- Semi-automated workflow produces a reproducible and fully-annotated oligonucleotide map
  - Fully-annotated LC/UV chromatogram
  - Sequence coverage map (up to 100% sequence coverage - e.g. BNT162b2)
  - Microheterogeneity assessment of 5′ terminus capping and 3′ terminus poly(A) tail length

- MS/MS fragmentation was optimized and fidelity of identifications verified by decoy sequence searching

- Oligonucleotide mapping assisted the development and commercialization of the Comirnaty® vaccine against SARS-CoV-2
  - Elucidation of Structure (3.2.S.3.1)
  - Comparability (3.2.S.2.6)
  - Data supported regulatory filings to health authorities in 180+ markets

- Gau, B.C., Dawdy, A.W., *et al. Sci Rep* **13**, 9038 (2023)
  - Step-by-step protocol and VBA-enabled data analysis tools are publicly available

Special Thanks

Brian Gau (Pfizer)

Lead oligonucleotide map co-developer & mass spectrometry lead on Comirnaty mRNA vaccine against SARS-CoV-2

BioNTech

ThermoFisher Scientific
Protein Metrics
Agilent
Waters