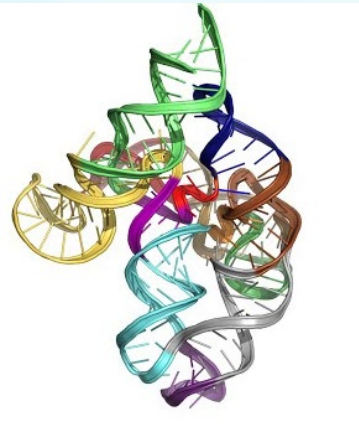
University of Sheffield

Direct Sequence Mapping and Characterization of mRNA Using Mass Spectrometry



- Prof Mark Dickman
- Dept of Chemical and Biological Engineering

>12 HPLC/UHPLC
- IP RP HPLC
- IEX
- SEC

SEC-MALS-DLS, (electrical)
AF4-MALS-DLS and CG-MALS-DLS.

Capillary Electrophoresis
- Fragment Analyzer
- Agilent 7100 CE

UV | CAD | Fluorescence

RNA Analytical Facility

University of Sheffield

Mass Spectrometry
- QExactive HF
- Orbitrap Exploris 240
- Orbitrap Exploris 480

Mass Photometry

Multimodal Spectroscopy (MMS)

Analysis of mRNA therapeutics

mRNA Critical Quality Attributes

- Identity (sequence)
- Content (concentration)
- Integrity (purity/intactness)
- 5’ capping efficiency
- 3’ poly(A) tail length and heterogeneity
- dsRNA impurities
- Residual DNA template ..

Table 2. Characterization and release testing for mRNA Drug Substance

Quality	Attribute	Method
Identity	mRNA sequence identity confirmation	High throughput sequencing (HTS)
		Sanger sequencing
		Reverse Transcriptase – PCR (RT-PCR)
Content	RNA concentration	Quantitative PCR (qPCR)
		Digital PCR (dPCR)
		Ultraviolet Spectroscopy (UV)
Integrity	mRNA intactness	Capillary electrophoresis[D]
		Capillary gel electrophoresis (CGE)[D]
		Agarose gel electrophoresis
Purity	5′ capping efficiency	Reverse-phase liquid chromatography mass spectroscopy (RP-LC-MS/MS)[D]
		Ion pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC)
	3′ poly(A) tail length	Ion pair reversed-phase high-performance liquid chromatography (IP-RP-HPLC)
	Product related impurities - dsRNA	Immunoblot
		Enzyme-linked immunosorbent assay (ELISA)
	Product related impurities - aggregate quantitation	Size exclusion-high-performance liquid chromatography (SEC-HPLC)[D]
	Product related impurities - percentage of fragment mRNA	Reversed-phase HPLC (RP-HPLC)[D]
	Process related impurities-residual DNA template	quantitative PCR (qPCR)
	Process related impurities - quantitation of free/non-incorporated nucleosides	Reverse-phase liquid chromatography mass spectroscopy (RP-LC-MS/MS)[D]
	Process related impurities - residual T7 RNA polymerase content	Enzyme-linked immunosorbent assay (ELISA)
Potency	Expression of target protein	Cell-based assay
Safety	Endotoxin	USP <85>
	Bioburden	USP <61>, <62>, <1115>
Other	Appearance	USP <790>
	Residual solvents	USP <467>
	pH	USP <791>

mRNA identity/sequence

Indirect sequencing methods

Sanger sequencing
Next Generation Sequencing (NGS)

mRNA → (RT-PCR) cDNA

- Loss of information regarding RNA modifications
- Time taken for sample prep/data analysis
- Amplification bias
- Inaccurate poly A tail analysis (poly purine/pyrimidines)

Direct sequencing methods

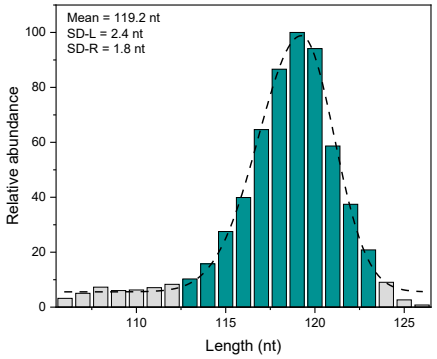
Nanopore sequencing

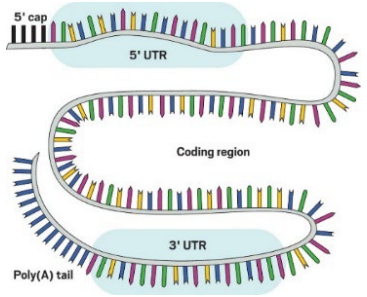
- -Issue of chemical modifications of mRNA
- -Accuracy of base calling
- -Validation of the method?
- Inaccurate poly A tail analysis (poly purine/pyrimidines)

Mass Spectrometry offers a powerful alternative to traditional sequencing methods for direct mRNA sequence mapping

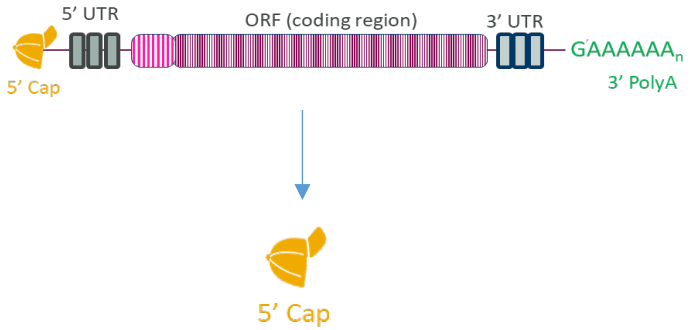
Mass Spectrometry Analysis of mRNA

Analysis of polyA tail
-length and heterogeneity

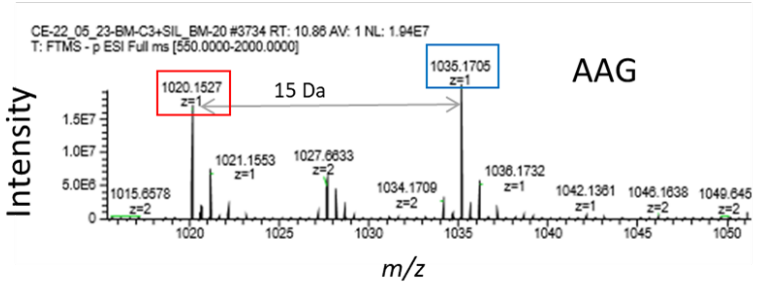


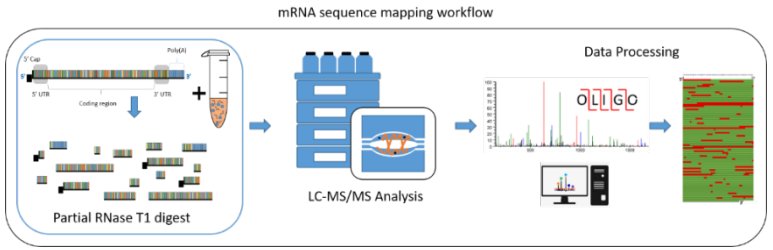


Quantification of 5'-Capping efficiency



Accurate mRNA quantification

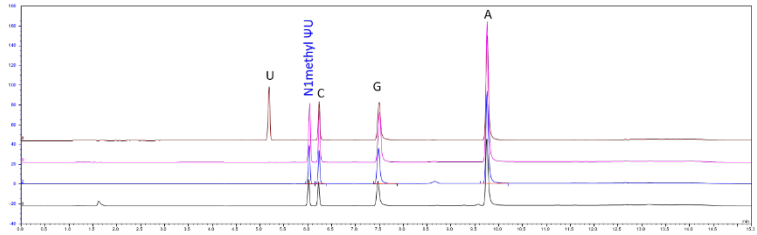




Vanhinsberg C et al., *Anal. Chem.* 2022, 94, 20, 7339–7349

Sequence analysis/identity testing
-validation of RNA chemical modifications

Analysis of nucleosides/NTPs
-base composition analysis of mRNA
-analysis of mRNA quality



Sequence mapping of mRNA therapeutics using LC MS

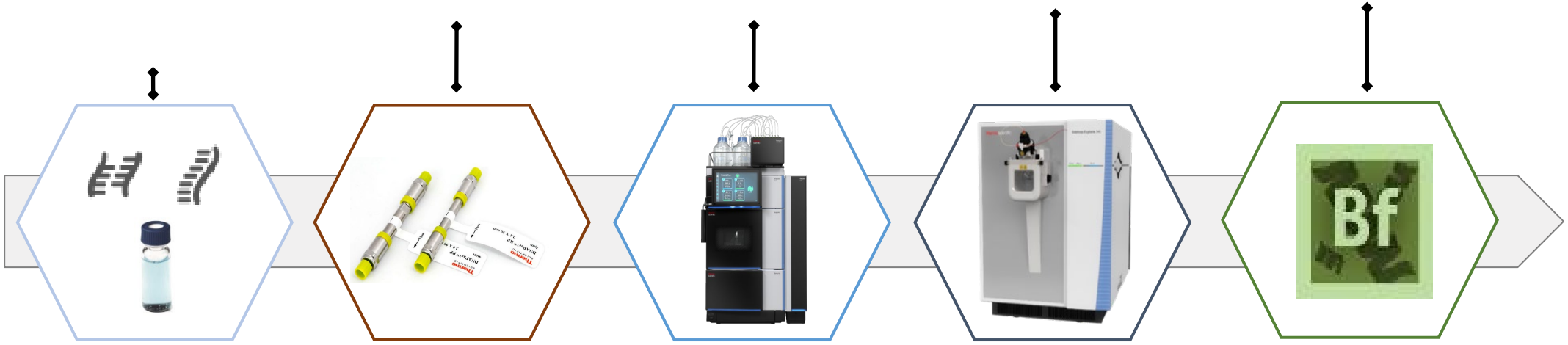
Digested RNA with Nuclease T1 immobilized on magnetic beads

Thermo Scientific™ DNAPac™ RP UHPLC columns

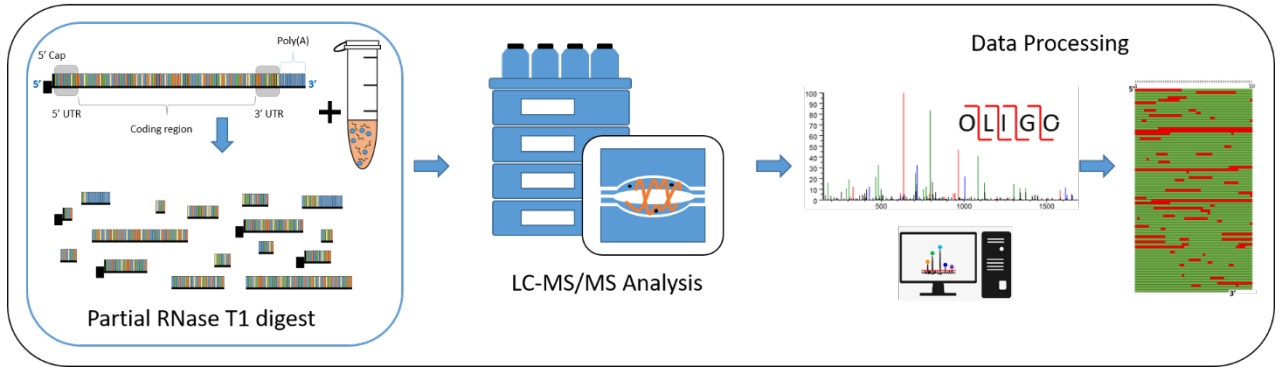
Thermo Scientific™ Vanquish™ UHPLC system

Thermo Scientific™ Orbitrap Exploris™ 240 mass spectrometer

Thermo Scientific™ BioPharma Finder™ 5.0 software

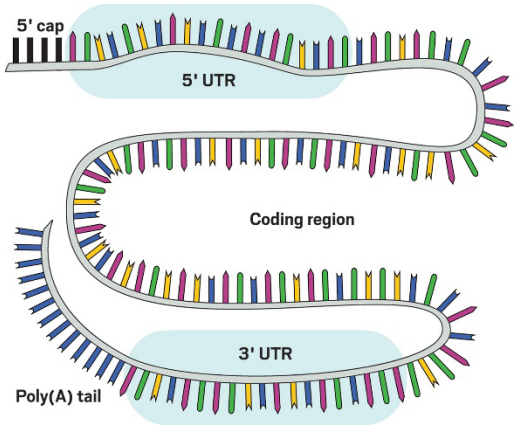


mRNA sequence mapping workflow

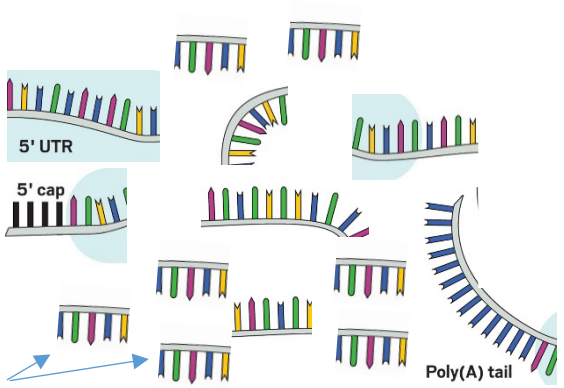


Vanhinsberg C et al., *Anal Chem* May 2022

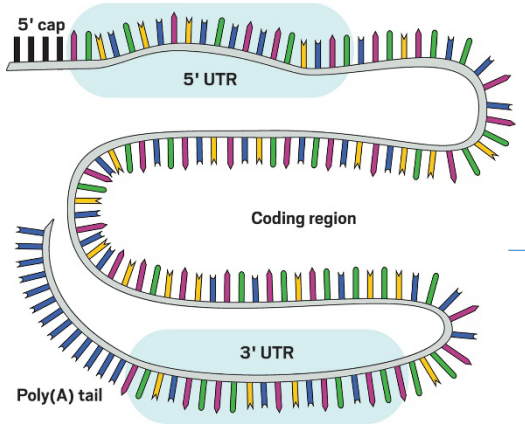
RNase digestions



Complete RNase T1 digestion

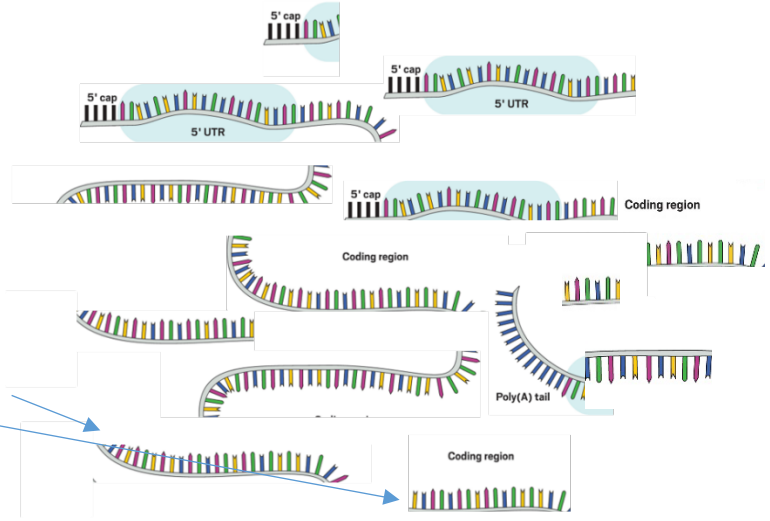


Non-unique oligos



Partial RNase T1 digestion

RNase T1 immobilized on magnetic beads



unique oligos

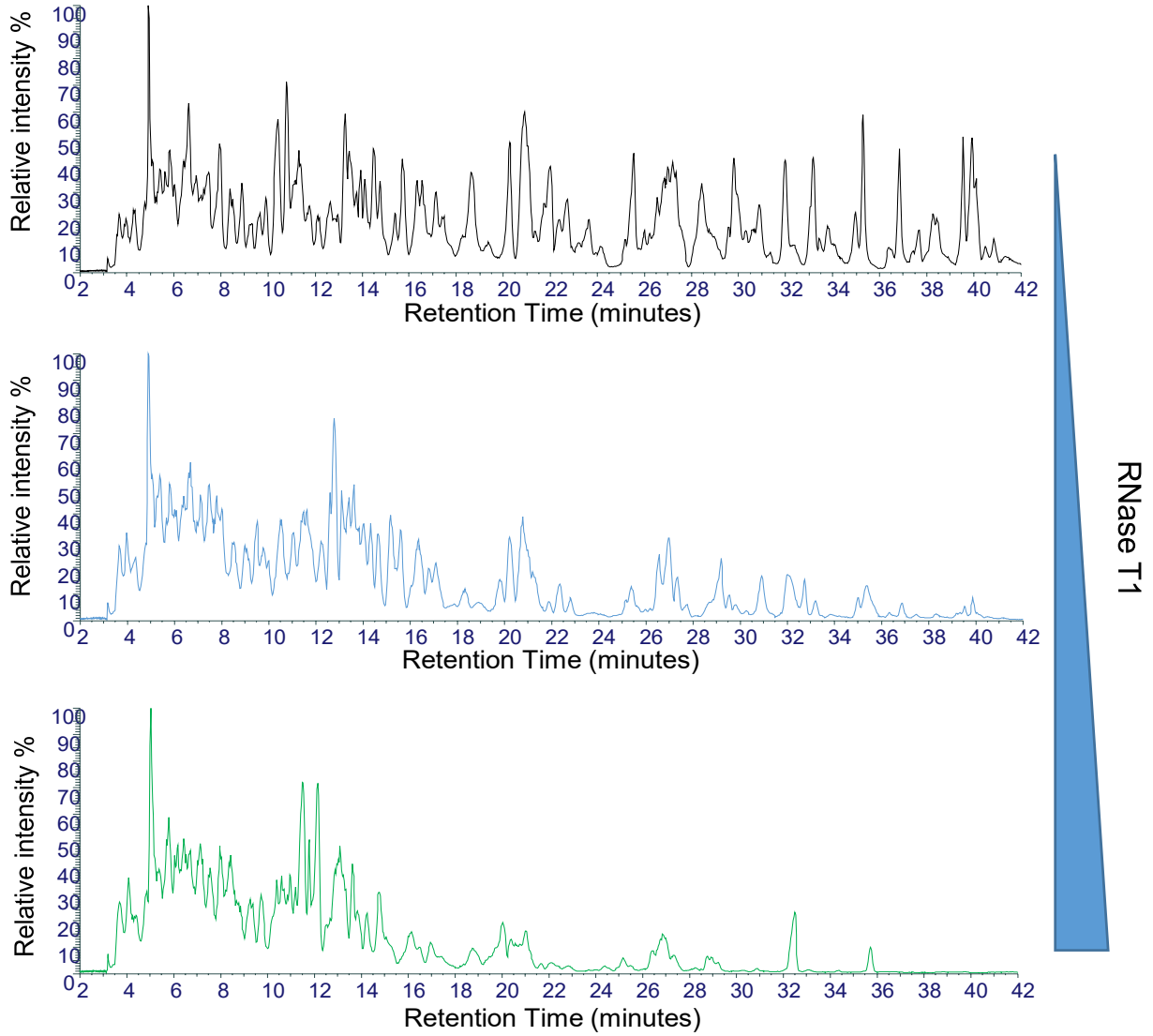
RNase digestions

Thermo Scientific™ SMART Digest™ Kits

RNase T1/A immobilized on magnetic beads

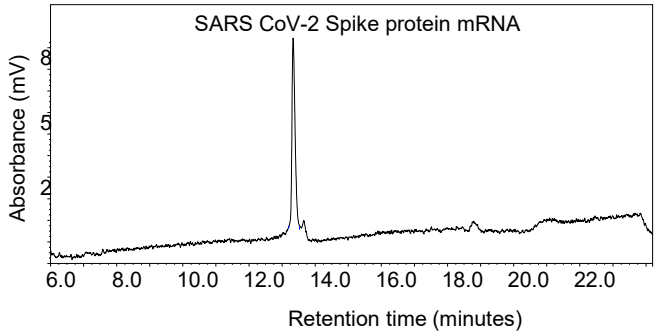
Simple control of partial RNase digests

Thermo Scientific™ KingFisher™ Duo Prime purification system for automated RNase T1 digestions



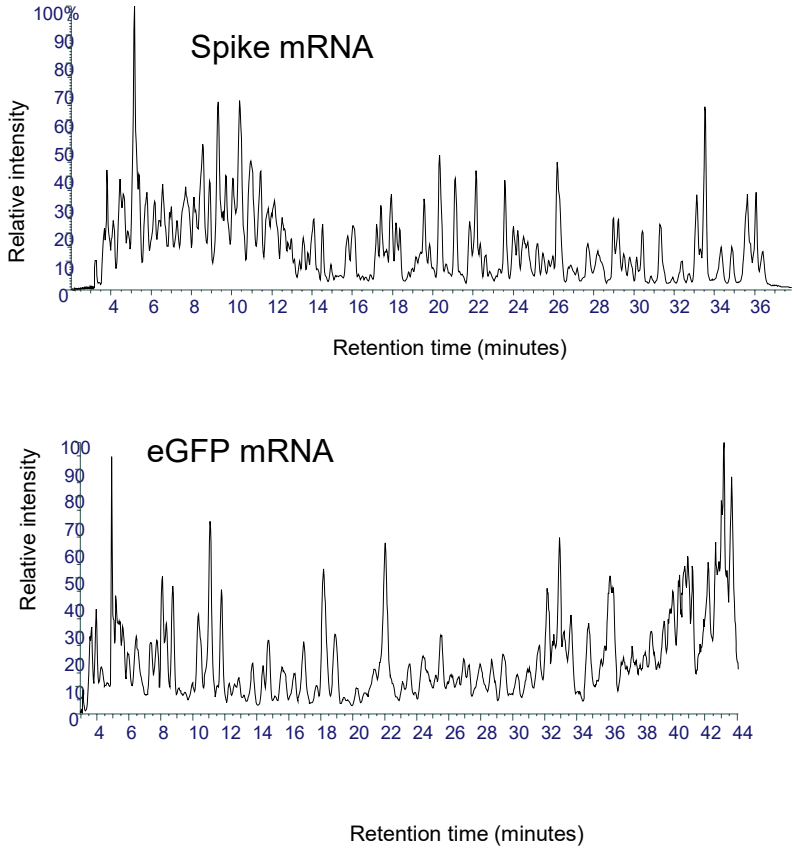
Sequence mapping of mRNA therapeutics

Production and Purification of mRNA



Partial T1 digest

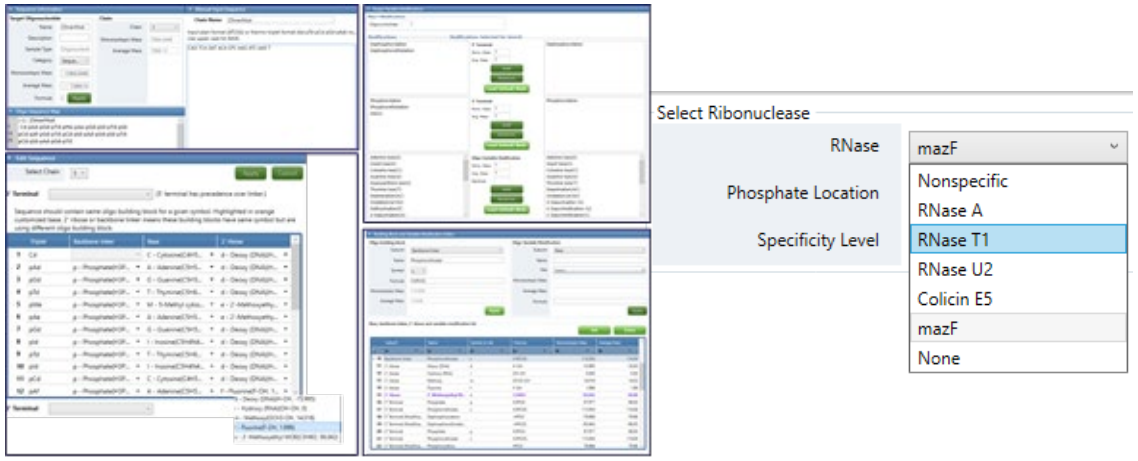
IP RP HPLC MS/MS Analysis



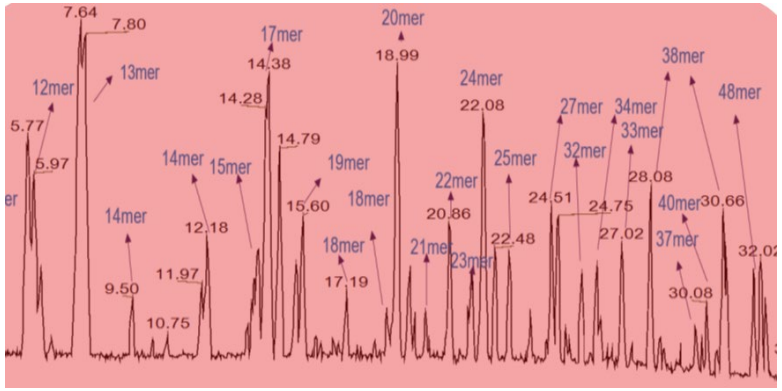
Data analysis – sequence mapping

mRNA analysis - Software

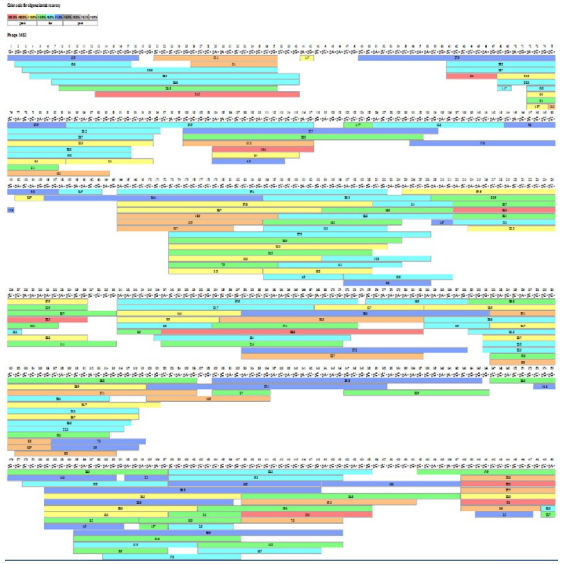
Sequence input → Ribonuclease selection → mRNA component detection → Identification confirmation



Ribonuclease selection includes common RNases



Chromatogram shading showing the pink identified digestion fragments from the mRNA sample.

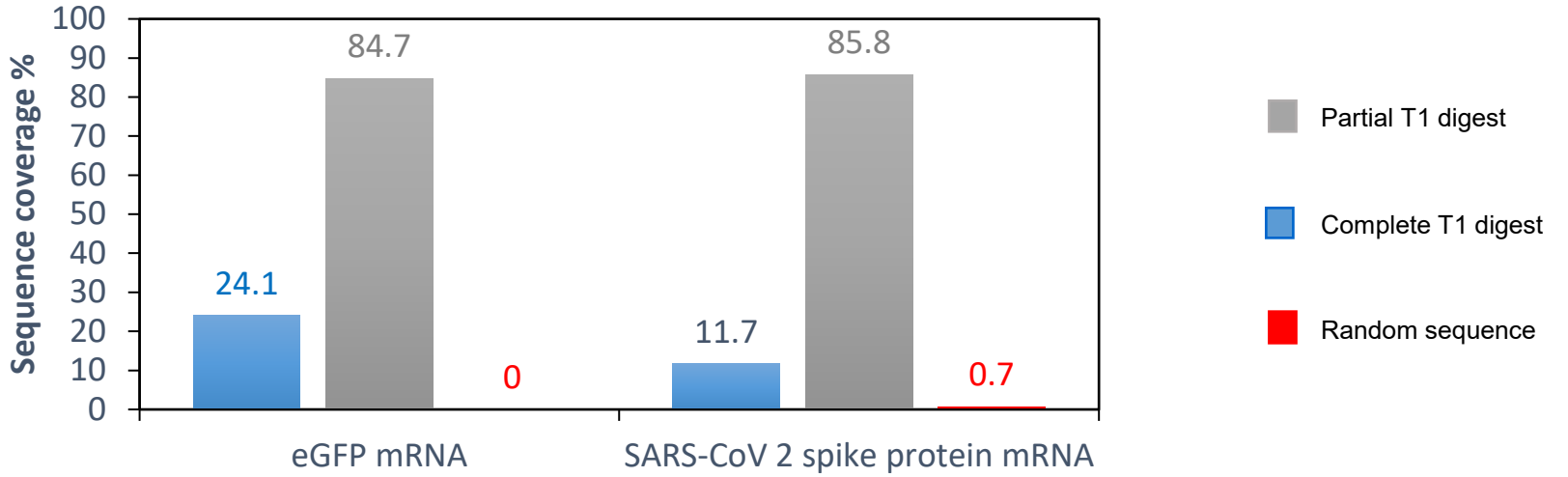


Sequence coverage map, automatic annotation and % coverage calculation

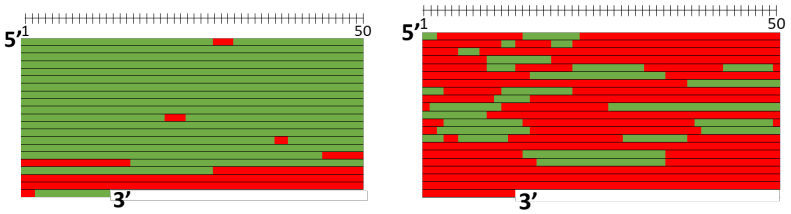
Sequence mapping of mRNA therapeutics

Single Analysis

Only using unique oligo identifications

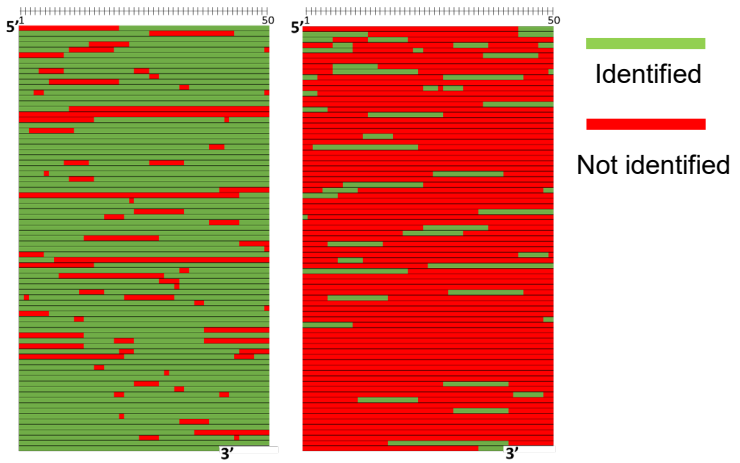


eGFP mRNA



Partial T1 digest Complete T1 digest

SARS-CoV 2 spike protein mRNA



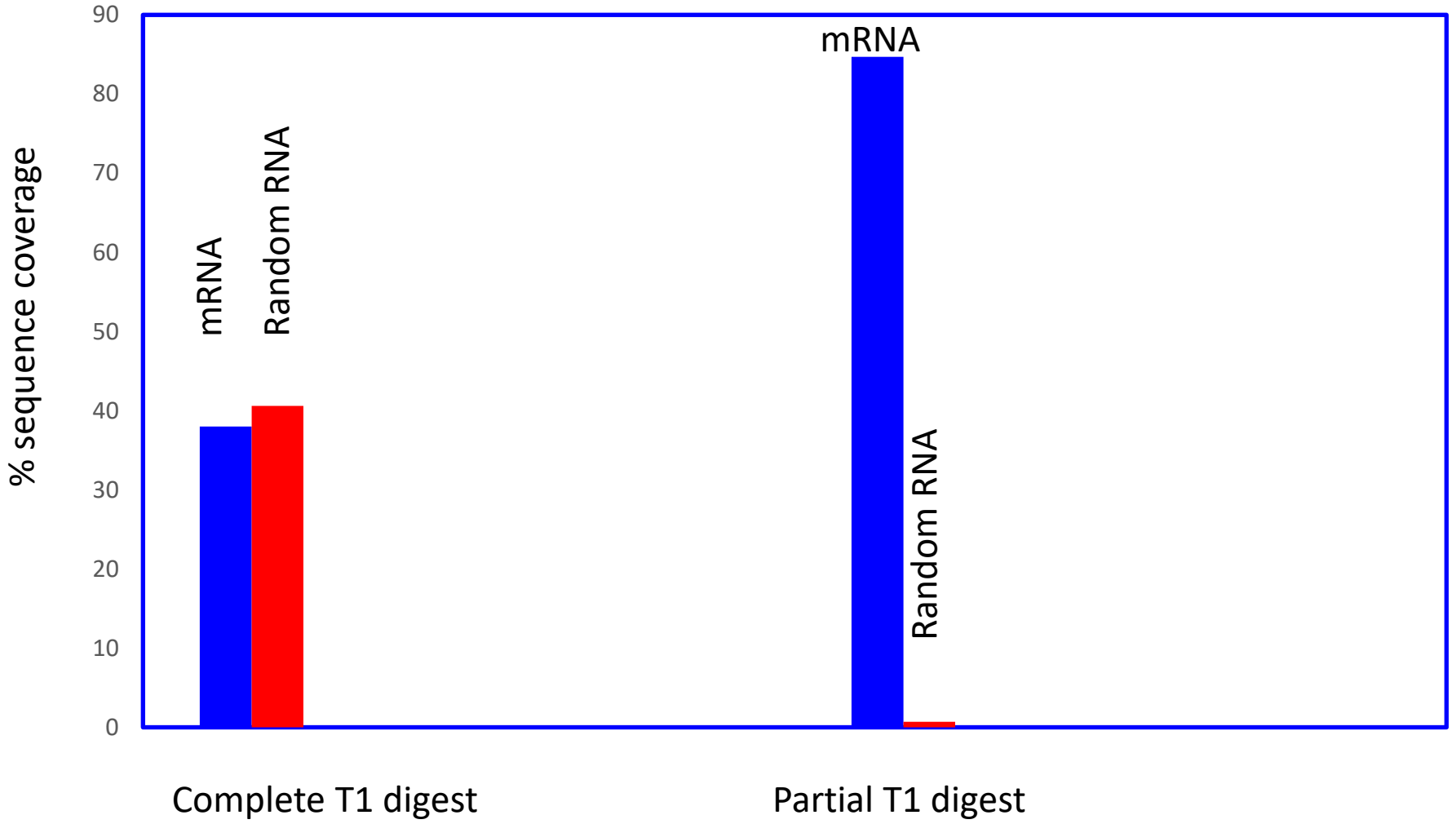
Partial T1 digest Complete T1 digest

- >90% sequence coverage <90 mins
- Identification and validation of mRNA chemical modifications

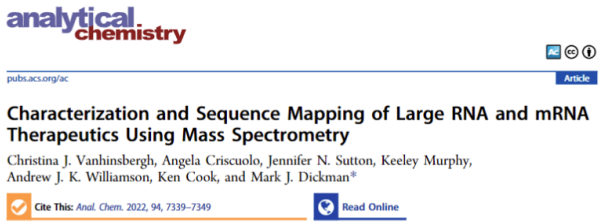
Sequence mapping of mRNA therapeutics

Sequence coverage **from non unique** oligos (map multiple times)

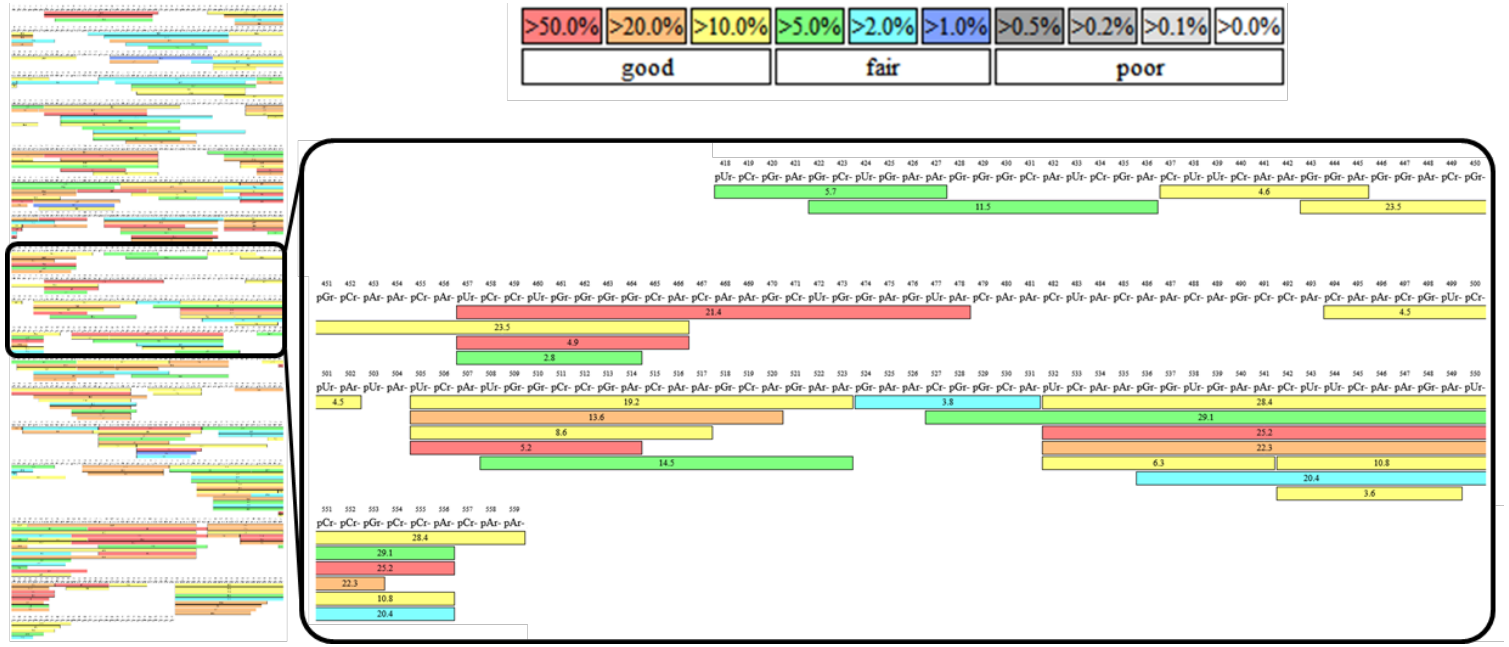
Sequence coverage from **unique** oligos (map multiple times)



Vanhinsberg C et al., *Anal Chem* May 2022



Generation of linear mRNA sequence maps using novel visualisation software



Partial RNase T1 digest generates large number of overlapping oligoribonucleotides

Linear mRNA sequence maps

Intensity of the colour represents a combined confidence score

Number of overlapping fragments and their individual confidences from the original BPF oligo identification

Sequence mapping data output and visualisation

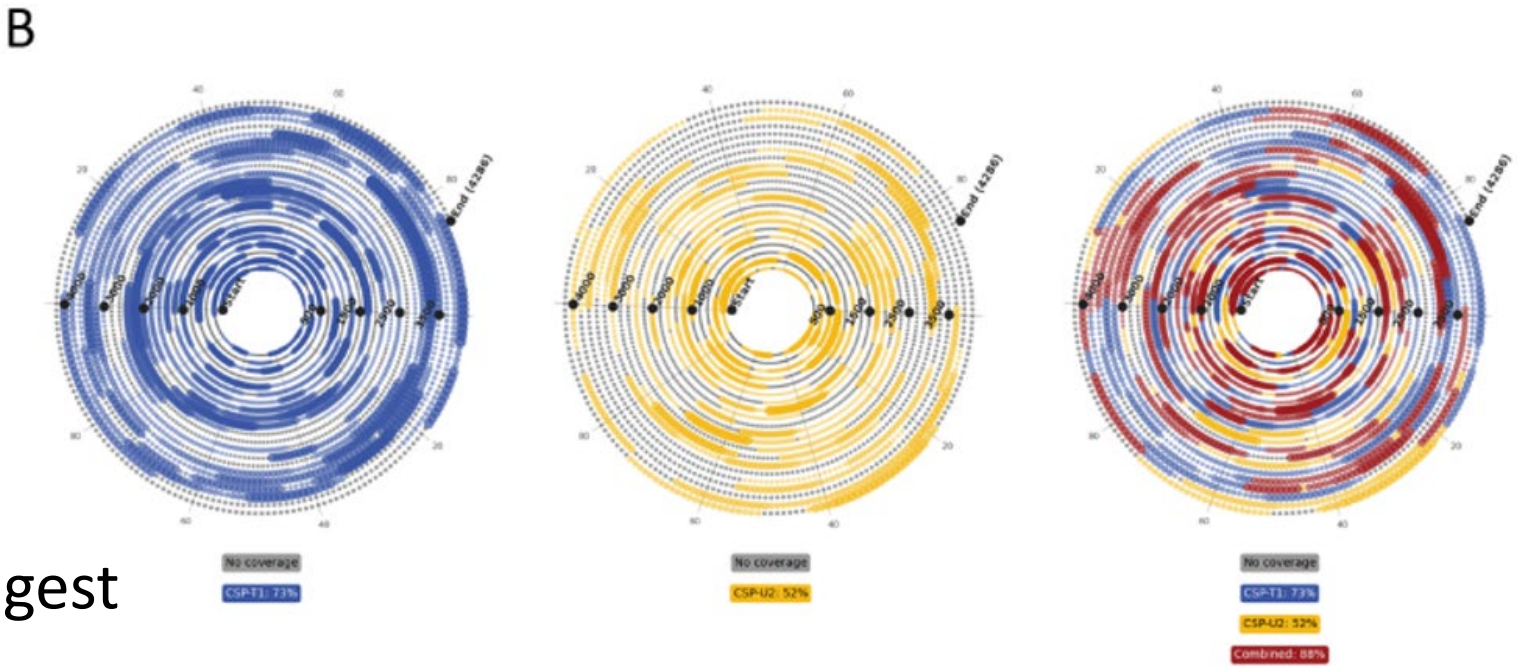
COVID Spike mRNA (>4000 nt)

Partial RNase T1

Partial RNase U2

Combined

Spiral mRNA sequence maps

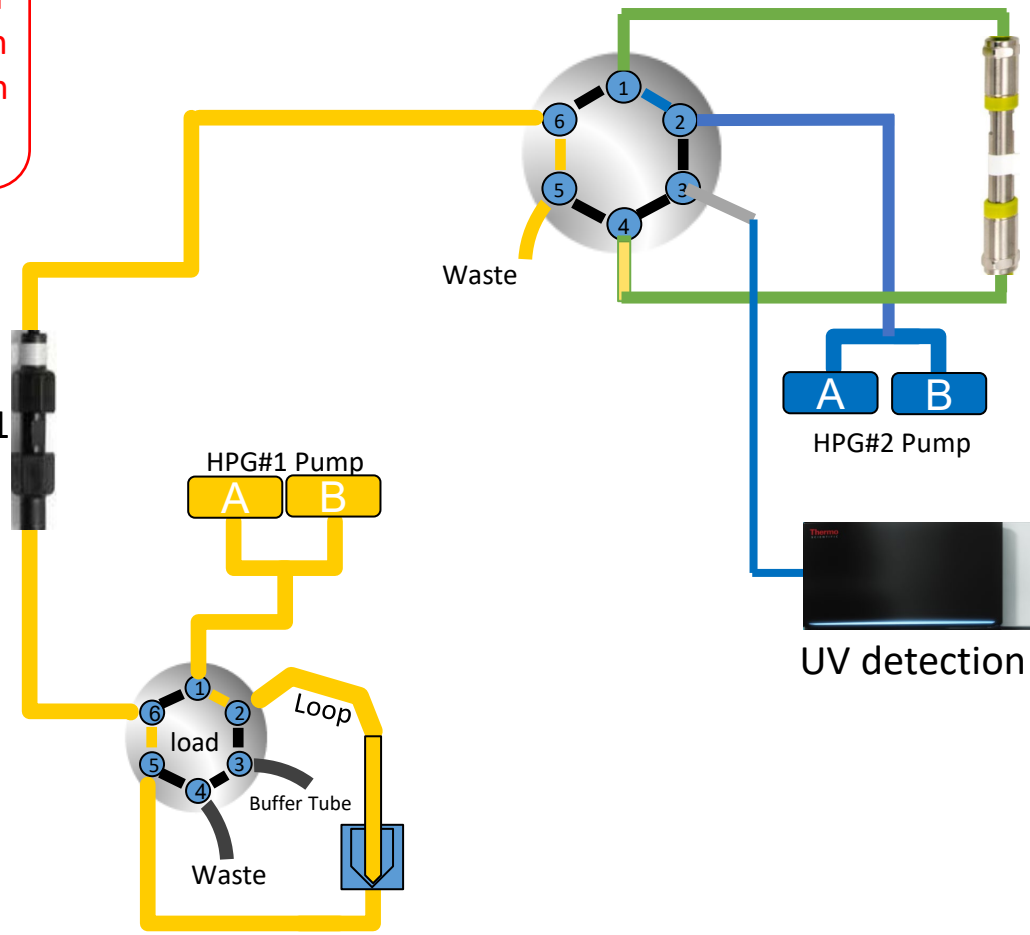


- Combining multiple RNase digest
- Integrating multiple overlapping fragments
- Simple visualisation of mRNA sequence maps

Welbourne E *et al.,* Mass spectrometry-based mRNA sequence mapping via complementary partial RNase digests and bespoke visualisation tools. *Analyst* 2025.

Online 2D LC MS analysis of partial RNase T1 digests

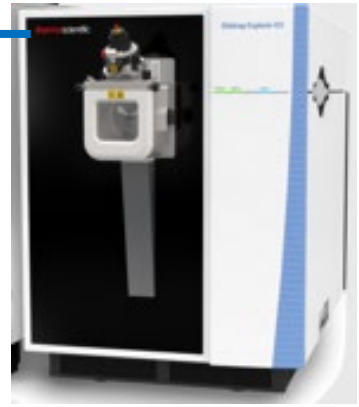
Flow rate control residence time in RNase T1 column (10-100 µl/min)



DNAPac RP 2.1 x 100 mm
IP RP HPLC

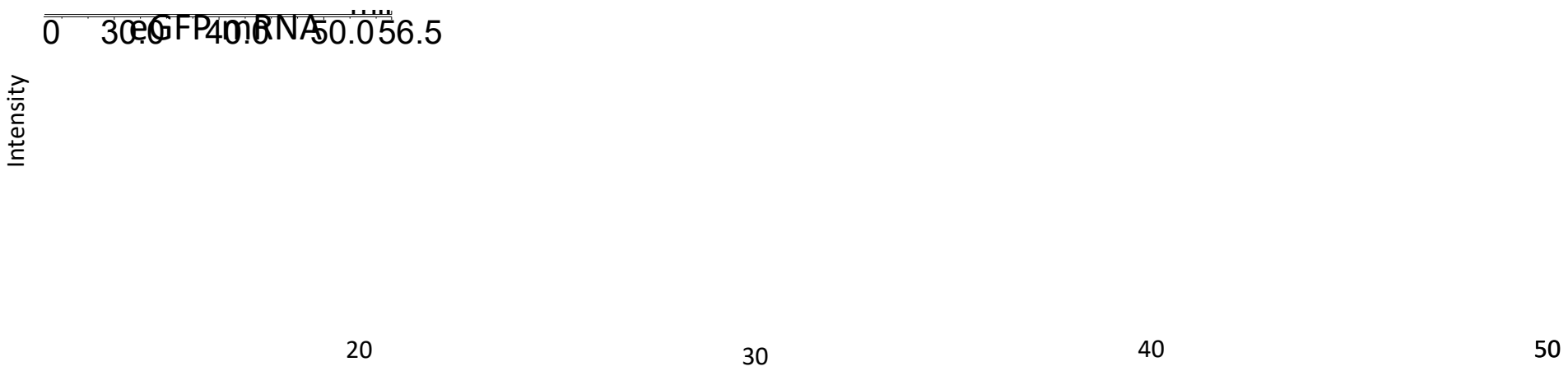
2D IP RP HPLC
Traps oligo fragments and separations using analytical gradient

Orbitrap Exploris 240



Online 2D LC MS analysis of partial RNase T1 digests

1D – TEAA, 50 µl/min, 25 °C
Valve switch 10 min
2D – TEA/HFIP, 200 µl/min, 25 °C

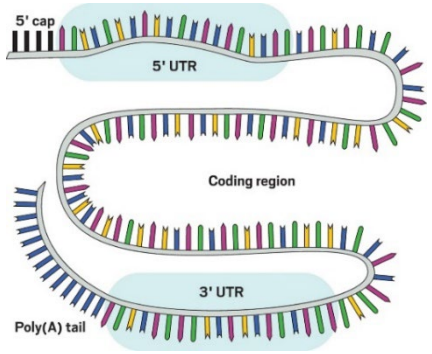


10 µg eGFP mRNA injected > 90% sequence coverage (based on unique oligonucleotides)

- Fully automated 2D LC MS analysis
- Sequence mapping performed in <60 min
- >90 sequence coverage 10 µg mRNA

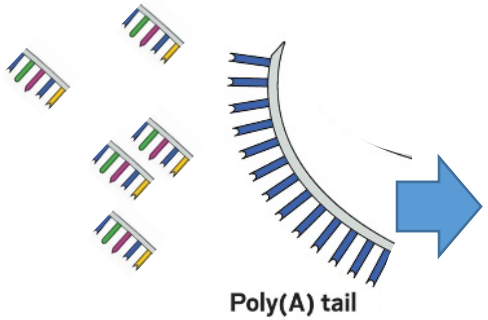
Analysis of polyA tail length and heterogeneity

mRNA



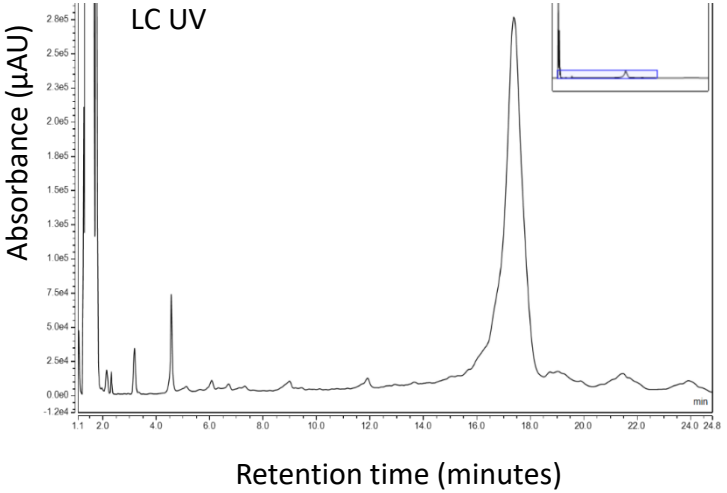
Complete RNase T1 digest

Beverly, M. et al. *Anal Bioanal Chem* **2018**, *410* (6), 1667–1677.

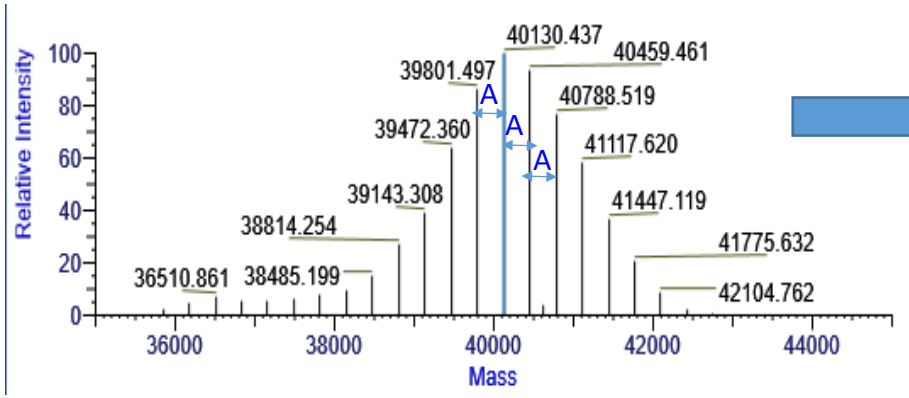


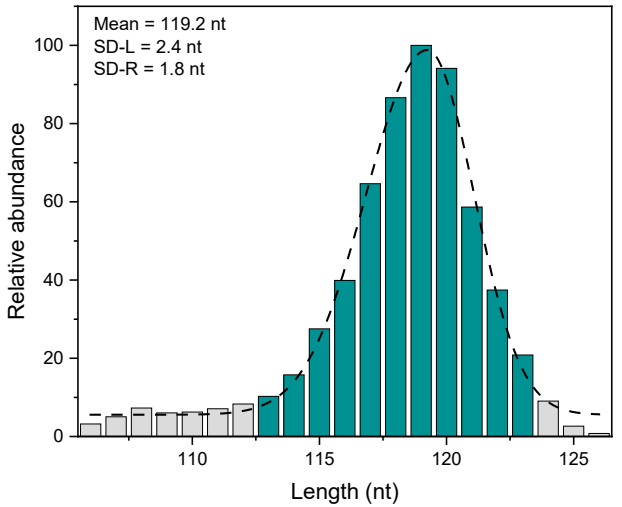
Low pressure mode

PolyA tail fragment

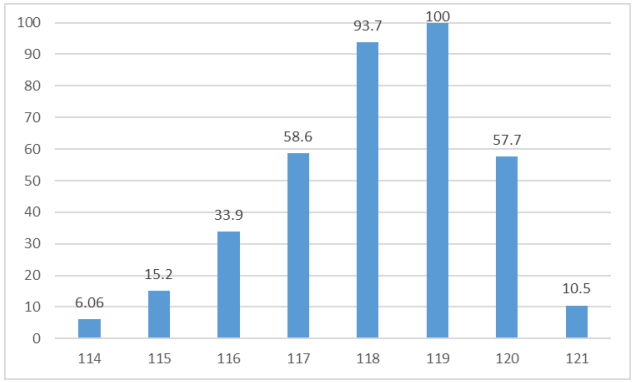


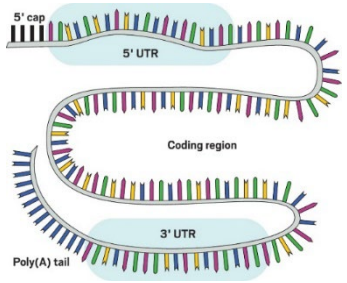
Mass Deconvolution





LC MS characterisation of mRNA therapeutics





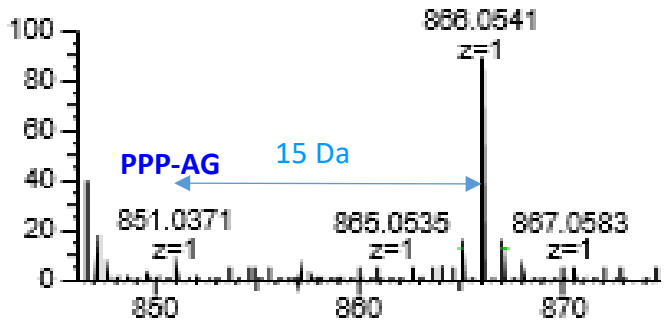
RNase T1

Partial RNase T1

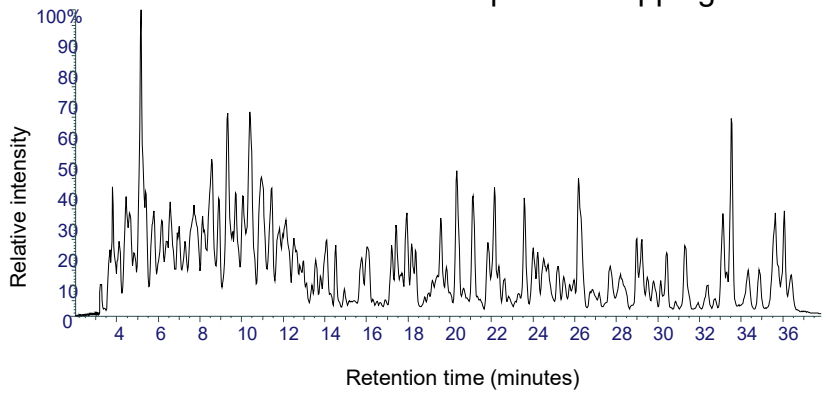
RNase T1

15N spike reference

PPP-AG



Direct mRNA sequence mapping



- Combine direct MS sequencing analysis (MS/MS sequencing)
  - + polyA tail MS characterisation
  - + identification of 5’ Cap (Capping efficiency)

- 100% sequence coverage –based on unique oligonucleotides

Conclusions

- Established a dedicated RNA analytical facility-methods for the analysis of CQAs
- Direct sequence mapping using LC MS enabled >90% sequence coverage of mRNA in <90 mins
- Developed powerful software and visualisation tools for analysis of combined partial RNase digests for mRNA sequence mapping
- Novel online 2D LC methods for mRNA sequence mapping -partial RNase T1
- LC MS offers a powerful rapid approach for the analysis of polyA tail length and heterogeneity

Acknowledgments

mRNA LC MS:
Dr Christina Vanhinsbergh
Dr Emma Welbourne
Jessica Dale
Dr Caroline Evans

Current group:
Paul Coombes
Gareth Owen
Zoe Ashcroft
Alex Webb
Mollie Glenister
Jessica Dale
Dan Buckledee
George Muir
Adam Miah
Lauren Wright

mRNA manufacturing:
Dr Adithya Nair
Dr Zoltán Kis
Kate Loveday

Software:
Royce Copley
Kesler Isoko
Peyman Z. Moghadam (UCL)

ThermoFisher SCIENTIFIC

Jim Stephenson
Josh Hinkle
Scott Kronewitter
John Syka
Ken Cook
Andrew Williamson
Angela Criscuolo
Jon Bardsley
Jennifer Sutton
Keeley Murphy

EPSRC Engineering and Physical Sciences Research Council

BBSRC bioscience for the future

R3 RNA READINESS + RESPONSE leap wellcomeleap.org

CEPI