Analytical Challenges, Solutions and Perspective on the Future Analysis and Characterization of mRNA

Khaled Yamout
Analytical Sciences, Quality and Manufacturing Consultant
On behalf of TriLink Biotechnologies

February 2024
Agenda

This presentation will discuss the criticality of establishing appropriate orthogonal analytical methods for a well characterized mRNA.

1. mRNA background and anatomy
2. Analytical consideration for analysis and characterization of mRNA
   - Importance of orthogonal analytical methods
   - Orthogonal ID methods
   - Poly(A) standards
   - Orthogonal Poly(A) analysis
mRNA structure

Nucleoside-modified mRNA

e.g., Ψ, m1Ψ, 5meC

Sequence-engineered mRNA

e.g., GC motif, structural UTR elements

Cap analogs

e.g., CleanCap® capping analogs

5’ cap

5’ untranslated region

Coding region

3’ untranslated region

3’ poly (A) tail

(A)50-120

Initiation codon

Open reading frame

End structures
Analytical consideration for analysis and characterization of mRNA
mRNA critical quality attributes and analytical methods

**Integrity and purity**
- Gel electrophoresis
- Capillary Electrophoresis
- IP-RP-HPLC
- SEC-HPLC
- AX-HPLC

**Identity**
- Sequencing (Sanger, NGS)
- RT-PCR
- "Fingerprinting" by LC-MS/MS
- Base composition Assay

**Content**
- UV-Vis Spectroscopy
- RT-qPCR
- RT-dPCR
- IP-RP-HPLC

**Functionality**
- In-vitro translation/Western blot (Cell Free assay)
- Cell-based assays (contract out)

**Impurities**
- DNA: qPCR
- Protein: Nano Orange; BCA
- NTPs: AX-HPLC
- Solvents: GC
- dsRNA: Immunoblot, ELISA

**Others**
- Appearance: USP <1>, <790>
- pH: USP <791>

**Safety**
- Endotoxin: USP <85>
- Bioburden: USP <61>, <62>, <1115>
Case study | Orthogonal mRNA ID methods

1. Development of three (3) orthogonal ID methods for mRNA:
   
   a. Sequencing (Sanger)
   
   b. Finger Printing LC-MS/MS
   
   c. Base composition by LC

2. Case study using Fluc mRNA with a mixture of Uridine and N1-Methylpseudouridine as compared to a FLuc WT mRNA
<table>
<thead>
<tr>
<th>Sample</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Fluc WT</td>
<td>The sequence was a match</td>
</tr>
<tr>
<td>Uridine and N-Me-PseudoU mixture</td>
<td>The sequence was a match</td>
</tr>
</tbody>
</table>
mRNA fingerprinting workflow by LC-MS/MS

**Partial mRNA digestion using magnetic beads with immobilized RNase T1**

1. Controlled enzymatic Digestion with RNase T1 mag beads.

2. Incubate Reaction for 2-30 min depending on construct.

3. Separate digested mRNA fragments from mag beads.

4. Analyze by LC-MS/MS.

5. Analyze data using Biopharma Finder Software.

---

**Experimental**

- RNA sample
- Enzymatic digestion
- LC-MS/MS analysis
- MS/MS data

**In-silico**

- RNA sequences
- In silico digestion
- Theoretical precursors and product ions mass
- Precursor ions and product ions matching
- Scoring and ranking
- Map modification onto RNA sequence

Yu, et al 2017
mRNA fingerprinting by LC-MS/MS

Target sequence coverage achieved for different constructs with and without modified NTPs

- **eGFP (996nt) WT:** target sequence coverage 87.9%
- **FLuc (1929nt) WT:** target sequence coverage 93.7%
- **Cas9 (4521nt) WT:** target sequence coverage 97.3%

<table>
<thead>
<tr>
<th></th>
<th>Sequence Coverage (%)</th>
<th>RSD (n = 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFP</td>
<td>87.9</td>
<td>0.0</td>
</tr>
<tr>
<td>FLuc</td>
<td>93.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Cas9</td>
<td>93.3</td>
<td>1.1</td>
</tr>
</tbody>
</table>

- TIC chromatogram highlighting the peaks used to calculate sequence coverage.
- Unhighlighted peaks were not used for identification.

~ 100% coverage of ORF plus UTRs
Accuracy and reproducibility of the method

- Cas9_1
- eGFP_1
- Fluc_1
- Average Mass Identified: Cas9_2
- R² = 1
- R² = 0.903
- R² = 0.931

- Cas9_2
- eGFP_2
- Fluc_2
- Average Mass Identified: eGFP_2
- R² = 1
- R² = 0.903
- R² = 0.921

- Cas9_3
- eGFP_3
- Fluc_3
- Average Mass Identified: Fluc_2
- R² = 1
- R² = 0.931
- R² = 0.921
Fingerprinting analysis of uridine and modified uridine sample

**Uridine and N-Me-PseudoU mixture**

**Control**
Base composition assay

Digest mRNA sample using snake venom phosphodiesterase

Converting to nucleosides using Calf Intestinal Phosphatase (CIP)

Analysis and quantitation of nucleosides sample by RP-HPLC against a nucleosides bases standard curve

Representative Calibration Curves
Linear range 0.4 – 200 µM
Base composition assay

[Cytidine, Uridine, N1-Methylψ, Guanosine, Adenosine graphs]

©2023 Maravai LifeSciences. This document and its contents are confidential and proprietary and are not to be shared or redistributed without express consent.
Base composition assay

Fluc WT mRNA control (blue) overlayed with N1081 modified + UPT spiked sample (black)
## Base composition assay

### µM Present in Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th># C</th>
<th># U</th>
<th># G</th>
<th># A</th>
<th># N1-MethylΨ</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluc WT theoretical</td>
<td>594</td>
<td>273</td>
<td>555</td>
<td>500</td>
<td>0</td>
<td>1922</td>
</tr>
<tr>
<td>Fluc with N1-MethylΨ spiked with UTP</td>
<td>567</td>
<td>26</td>
<td>585</td>
<td>502</td>
<td>242</td>
<td>1922</td>
</tr>
<tr>
<td>Fluc WT control</td>
<td>569</td>
<td>262</td>
<td>587</td>
<td>504</td>
<td>0</td>
<td>1922</td>
</tr>
</tbody>
</table>
Summary | Orthogonal mRNA ID methods

1. Each ID methods provide unique information regarding the identity of mRNA
   a. Sequencing (Sanger) provides a sequence coverage of the ORF but would not directly differentiate between WT and modified nucleotides.
   b. Finger Printing LC-MS/MS provides sequence coverage of the ORF and UTR’s and indication of any possible nucleotides mismatching.
   c. Base composition by LC provides specific details regarding the base composition of the sequences and how does it match to the theoretical sequence base composition.

2. When combined, the three orthogonal ID methods can provide a tool for investigational and some detailed levels of ID testing.
Methodologies for characterization of Poly(A) tail

Due to the size of the mRNA, direct analysis of mRNA by Liquid Chromatography - Mass Spectrometry (LC-MS) can be challenging at time. As such, this requires to digest the mRNA prior to analysis. Main methods of digesting mRNA:

• Annealing at the 3’ UTR using DNAzyme. This yield the poly (A) + x Bases and untailed mRNA- x Bases
• Enzymatic digestion such as RNase which yield a complete digestion except for the Poly (A) portion.
• Base composition

Once the sample is prep, then it can be analyzed by a wide variety of analytical methods

Recently, various publication on this subject and quite a few approaches are still being vetted and evaluated for fit for use in a QC environment.
Poly(A) tail enzymatic digestion method by LC-MS

Enzymatic digestion of mRNA using RNase A and T1

AAAAA
AAAAAAA
AAAA

Analysis by LCMS

Poly(A) tail remains intact
Fluc mRNA poly(A) tail by LC-MS

Poly(A) tail distribution

- Number of As: 122, 123, 124, 125, 126
- Intensity: 2.50E+05, 2.00E+05, 1.50E+05, 1.00E+05, 5.00E+04
Determination of poly(A) tail length by base composition method

1. Use of developed ID composition assay to determine poly (A) length

   a. Tested 6 lots of Fluc mRNA made by the same IVT process and backbone were analyzed against an mRNA without a Poly (A) tail.

   b. Samples were prepared and analyzed for the amount of A present.

   c. Poly (A) length was determine based on the concentration of A and that was compared to values of Poly (A) by LC-MS method.

<table>
<thead>
<tr>
<th>FLuc mRNA</th>
<th>Lot 1</th>
<th>Lot 2</th>
<th>Lot 3</th>
<th>Lot 4</th>
<th>Lot 5</th>
<th>Lot 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed # A</td>
<td>498</td>
<td>496</td>
<td>498</td>
<td>500</td>
<td>502</td>
<td>504</td>
</tr>
<tr>
<td>theoretical # A without tail</td>
<td>380</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Tail #</td>
<td>118</td>
<td>116</td>
<td>118</td>
<td>120</td>
<td>122</td>
<td>124</td>
</tr>
<tr>
<td>Tail Length by LCMS</td>
<td>124</td>
<td>123</td>
<td>123</td>
<td>123</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>% difference</td>
<td>95%</td>
<td>94%</td>
<td>96%</td>
<td>98%</td>
<td>98%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Direct analysis of poly(A) tail DNAzyme cleavage method

Reaction steps
- DNAzyme reaction
- Annealing
- Cleavage

HPLC-UV
- Cleaved eGFP mRNA
  - 127-mer polyA tail cleaved from eGFP mRNA

HPLC-MS-TIC
- Cleaved eGFP mRNA
- Synthetic 120-mer polyA
Another approach for poly(A) tail analysis

Deconvoluted spectra of synthetic 120-mer poly(A) tail

Deconvoluted spectra of poly(A) tail cleaved by DNAzyme from eGFP mRNA
Synthetic poly(A) standards

- This has been a cross functional collaborative effort to provide better solution to enhance customer’s journey
  - AS/QC team
  - Oligo team
  - Core mRNA

- Five synthetic Poly (A)s 39, 59, 79, 99, 120
- Analysis of synthetic Poly (A)s by various methods
  - Size exclusion
  - Modified IPRP-HPLC
  - IPRP-HPLC
  - Octylamine HPLC
  - AX-HPLC
  - Fragment analyzer
MS deconvoluted data for size determination

<table>
<thead>
<tr>
<th>RT (min)</th>
<th>Base Peak Mass (Da)</th>
<th>Intensity</th>
<th>Spectral Quality</th>
<th>LC/MS Peak Area</th>
<th>LC/MS Area Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.362</td>
<td>12775.4</td>
<td>1.01E+007</td>
<td>ok</td>
<td>6.33E-008</td>
<td>95.91</td>
</tr>
<tr>
<td>12.979</td>
<td>19360.1</td>
<td>7.92E+007</td>
<td>ok</td>
<td>2.40E-009</td>
<td>99.84</td>
</tr>
<tr>
<td>13.309</td>
<td>25944.1</td>
<td>9.19E+007</td>
<td>ok</td>
<td>4.93E-009</td>
<td>99.95</td>
</tr>
<tr>
<td>13.463</td>
<td>32528.2</td>
<td>7.27E+007</td>
<td>ok</td>
<td>7.05E-009</td>
<td>99.98</td>
</tr>
<tr>
<td>13.528</td>
<td>39441.1</td>
<td>1.42E+008</td>
<td>ok</td>
<td>9.68E-009</td>
<td>99.98</td>
</tr>
</tbody>
</table>
Poly(A) tail mixture

Length vs relative retention time standard curves

\[ y = 65.72x - 779.89 \]
\[ R^2 = 0.8762 \]

\[ y = 0.4145x - 3.5514 \]
\[ R^2 = 0.9617 \]

\[ y = 77.125x^2 - 1923x + 12025 \]
\[ R^2 = 0.9931 \]
## Poly(A) standards results

<table>
<thead>
<tr>
<th>Length</th>
<th>Target Mass (Da)</th>
<th>Observed (Da)</th>
<th>LCMS Area Percent PolyA method</th>
</tr>
</thead>
<tbody>
<tr>
<td>120 A’s</td>
<td>39443.2</td>
<td>39442.7</td>
<td>99.99</td>
</tr>
<tr>
<td>99 A’s</td>
<td>32529.8</td>
<td>32529.1</td>
<td>99.97</td>
</tr>
<tr>
<td>79 A’s</td>
<td>25945.6</td>
<td>25945.1</td>
<td>99.94</td>
</tr>
<tr>
<td>59 A’s</td>
<td>19361.4</td>
<td>19360.4</td>
<td>99.96</td>
</tr>
<tr>
<td>39 A’s</td>
<td>12777.2</td>
<td>12776.2</td>
<td>99.88</td>
</tr>
</tbody>
</table>
Length versus RT graphs constructed with synthetic poly(A) tail standards

Waters Octylamine Method
\[ y = 2.874959x + 17.149191 \]
\[ R^2 = 0.984670 \]

AX Method
\[ y = 13.466801x - 6.192984 \]
\[ R^2 = 0.999476 \]

SEC Method
\[ y = -0.337145x + 3.708222 \]
\[ R^2 = 0.998078 \]

IPRP Method
\[ y = 24.104082x - 40.023611 \]
\[ R^2 = 0.988167 \]

PolyA generic LCMS
\[ y = 0.303091x^2 - 7.511900x + 48.135621 \]
\[ R^2 = 0.985648 \]
MS results for Cas9 80A mRNA– generic Poly(A) tail LC-MS

Theoretical

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cas9 digested</td>
<td>80</td>
</tr>
<tr>
<td>Cas9 cleaved</td>
<td>85</td>
</tr>
</tbody>
</table>

8 nt difference between the cleavage and RNase treatment

Result does not match
5-mer of 3’UTR region with tail
MS results for eGFP 120A mRNA – generic Poly(A) tail LC-MS

Theoretical

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFP digested</td>
<td>120</td>
</tr>
<tr>
<td>eGFP cleaved</td>
<td>127</td>
</tr>
</tbody>
</table>

7 nt difference between the cleavage and RNase treatment
Result matches expectations
7-mer of 3’UTR region with tail
Calculated Poly(A) tail length summary of orthogonal methods

<table>
<thead>
<tr>
<th>Calculated Length</th>
<th>Octylamine</th>
<th>AX</th>
<th>SEC</th>
<th>mRNA IPRP</th>
<th>Generic PolyA LCMS</th>
<th>Theoretical Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFP digested</td>
<td>117</td>
<td>122</td>
<td>120</td>
<td>130</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td>Cas9 digested</td>
<td>89</td>
<td>84</td>
<td>86</td>
<td>79</td>
<td>82</td>
<td>80</td>
</tr>
<tr>
<td>eGFP cleaved</td>
<td>126</td>
<td>125</td>
<td>141</td>
<td>130</td>
<td>116</td>
<td>127</td>
</tr>
<tr>
<td>Cas9 cleaved</td>
<td>86</td>
<td>79</td>
<td>74</td>
<td></td>
<td></td>
<td>85</td>
</tr>
</tbody>
</table>

• Poly(A) tail lengths were calculated with length vs RT graphs shown on the previous slide.
Summary

1. Robust analytical methods for mRNA characterization
2. Orthogonal ID Methods
3. Multiple approaches for poly(A) analysis
4. Poly(A) standards

<table>
<thead>
<tr>
<th>Identity, purity and characterization testing</th>
<th>Safety and general testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Identification</td>
<td>• Safety (microbial and endotoxin)</td>
</tr>
<tr>
<td>• Quantitation</td>
<td>• Stability</td>
</tr>
<tr>
<td>• Product characterization</td>
<td>• Qualitative assessment</td>
</tr>
<tr>
<td>• Functionality</td>
<td></td>
</tr>
<tr>
<td>• Impurity profiling</td>
<td></td>
</tr>
</tbody>
</table>

**Custom services**
Construct-specific method development, qualification/validation, analytical testing and stability studies

Achieve your analytical objectives from process development to scale-up and cGMP manufacturing
## Analytical methods lifecycle by phase

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Discovery</th>
<th>Preclinical/Toxicology</th>
<th>Phase I</th>
<th>Late Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of material</td>
<td>RUO material</td>
<td>Engineering material</td>
<td>Phase I GMP material</td>
<td>Late Phase GMP material</td>
</tr>
<tr>
<td>Level of analytical method evaluation</td>
<td>Feasibility recommended</td>
<td>Feasibility</td>
<td>Method* qualification/ verification</td>
<td>Method* validation</td>
</tr>
<tr>
<td></td>
<td>Generic method based on structure</td>
<td>Generic method based on structure</td>
<td>Specific method</td>
<td>Specific method</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Specificity**</td>
<td>• Specificity**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Linearity</td>
<td>• Linearity</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Dynamic range</td>
<td>• Dynamic range</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Accuracy</td>
<td>• Accuracy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Precision**/IP</td>
<td>• Precision**/IP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• LOD**/LOQ</td>
<td>• LOD**/LOQ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Robustness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Sample stability</td>
</tr>
</tbody>
</table>

*For intended use e.g. to support safety, integrity, strength, purity, and quality

** Qualitative assessment
Analytical Services and Quality Control laboratory expansion
TriLink’s Analytical Sciences Center of Excellence (ASCE)

Centralized hub to drive innovation in nucleic acid analytical methodologies

- Develops additional methodologies for characterization of nucleic acids
- Provides method development, qualification, validation, stability testing, and product characterization
- Offers standalone support for RUO and non-clinical mRNA
ACKNOWLEDGEMENTS

TriLink BioTechnologies®
Analytical Services team

Contact
trilinkbiotech.com/contact-us
Thank you