Roundtable Session 1 – Table 7 – MS of RNA and Oligonucleotides: Characterization and Quantitation of PQAs

Facilitator: Steven Broome, *ProteinMetrics*

Scribe: Brian Gau, *Pfizer*

Abstract:

This roundtable will explore and discuss commonly characterized RNA and oligonucleotides Product Quality Attributes (PQAs), as well as the challenges faced by analysts. The goal is for everyone to gain a better understanding of the critical questions being addressed using mass spectrometry today and to come away with practical tips and suggestions. The discussion will focus on challenges related to mass spectrometry analysis of oligonucleotide PQAs. The challenges may be related to any aspect of the oligonucleotide PQA analysis workflow, including sample preparation, data acquisition, and data analysis. We will share existing solutions to challenges and discuss possible solutions to outstanding challenges. This discussion will be valuable for those with significant experience and for those who are just starting to learn about oligonucleotide characterization using mass spectrometry. Even if you are unable to answer some of the discussion topic questions, listening and asking questions will help everyone in the group learn and exchange ideas.

Notes:

The attendees were expressly interested in RNA characterization, spanning small siRNA, RNA conjugates, and large mRNA. Antisence DNA and other small synthetic oligonucleotide characterization was discussed in the context of comparison to RNA quality attributes and analytical methods. In general, PQAs encompass the profiling of the 5’- and 3’-end heterogeneity, sequence confirmation, and intact mass. With the exception of intact mass, these can be accomplished by “bottom-up” endonuclease digestion and LC-MS analysis of the product oligonucleotides.

Unlike synthetic DNA/RNA, synthetic impurity profiling by MS is not done for RNA that is produced by in vivo from cells or by in vitro transcription from a DNA plasmid. The heterogeneity of the 5’-end is affected by several factors. Extra sequence “variants” (i.e. +1 or +2 nucleotides) may be present in addition to the uncapped species, which is usually a 5’-triphosphate and 5’-diphosphate. Impurities in the capping raw material can carry through if they’re capable of being incorporated. Too much RNase H can give rise to non-specific cuts if the 5’-end is probed using an annealing oligo approach. The 3’-end heterogeneity is usually transcriptional-slippage-derived incorporation of more than the intended # of adenosine nucleotides, if the RNA is transcribed from DNA containing the poly(A) message (which is not the case endogenously).

Though one company has recently shown data on the intact mass of a 0.7 MDa mRNA by MS, intact mass analysis of mRNA is best done by Mass Photometry or possibly Charge Detection Mass Spectrometry (CDMS). The highly charged phosphodiester backbone of RNA at physiological pH fosters significant ion pairing with salt and metal adducts. These can make CDMS less robust the mass photometry. Like MALDI-MS, mass photometry must be properly calibrated with biomolecules or standards which are comparable in size or larger than the target RNA.

There are several endonuclease digestion enzymes that may be used for the bottom-up analysis of RNA. RNase T1, RNase 4, and MazF are commercially available. Waters will soon be announcing other RNase enzymes with different specificity. RNase T1 cuts on the 3' of every G and is very efficient. "Dose make the poison": too much enzyme will lead to nonspecific cuts at other sites. RNase T1 catalyzes the hydrolysis of the phosphodiester bond and the hydrolysis of the cyclic phosphate product of the 1st reaction to give a phosphate at the 3' end, in separate reactions. A hallmark of RNase T1 non-specific cuts is cyclic phosphate rather than phosphate at the 3' end for a non-G 3'-end.

An advantage to using RNase T1 over other enzymes is that because of its cut-site frequency and exquisite sensitivity, the expected oligonucleotides of all lengths from 1-70 (sequence depending) are produced reproducibly. With appropriate ion pair reversed phase (IP-RP) chromatography, these may be spread out and separated such that the IP-RPLC UV chromatogram, though complex, is highly reproducible and may be completely characterized.

The advantage to using other enzymes like MazF and RNase 4 is that they cut with less frequency owing either to their cut-site motif and/or lower efficiency. This results in many more product oligonucleotides that map to exclusively one locus in the target mRNA sequence. This allows for an LC-MS based unique-sequence coverage analysis, with the possibility of complete sequence elucidation by overlapping contiguous oligonucleotides. The chromatogram tends to be more crowded in the later-eluting unique-sequence region for IP-RPLC.

It is important to note that, unlike proteins, there are already very good contiguous sequencing technologies (i.e. NGS) that are more fit-for-purpose w.r.t. sequence confirmation for mRNA. LC-MS-based oligonucleotide mapping is seen as an orthogonal approach to NGS, much as LC-MS peptide mapping HCP analysis is orthogonal to HCP ELISA.

The pioneering work in synthetic DNA characterization and oligonucleotide chromatography has provided for mature chromatography and mass spectrometry solutions for the analysis of oligonucleotide digests. Ion pair reversed phase UHPLC can use several different ion pair reagents in the mobile phase. For mRNA oligonucleotides, the triethylamine (TEA) / hexafluoroisopropanol (HFIP) combination works very well, with methanol as the strong solvent. Typically a tertiary amine such as TEA is used as the IP reagent: the pH is adjusted with a volatile buffer such as HFIP to give rise to triethylammonium, which couples electrostatically to the negatively charged phosphodiester backbone. Using 0.1% TEA/ 1% HFIP in MPA and MPB with MPB = 50% methanol, and using the Waters Premier Oligonucleotide 2.1 x 150 mM column, the full spread of RNase T1 digest oligos from a 4000+ nt mRNA can be separated such that oligos of different sizes group to distinct chromatogram regions. No other kind of chromatography, not HILIC, not RPLC, can give this kind of exquisite separation based on # of nucleotides. This separation is b/c of the sticky alkyl-ammonium ion-pair with phosphate.

A dedicated HPLC system for ion-paired analysis is recommended owing the stickiness of IP reagents. Manufacturers do have very good flushing protocols (offline from the MS), but these may require ~2 Day of downtime. A dedicated MS is not necessary; however, be mindful of the HPLC system and look out for low MW ions attributed to the IP reagents.

The mass spectrometry simply requires a commitment to working in negative mode. The source should be hot, and tuned with the appropriate vendor tuning mixture in negative mode. The gas rates will depend on flow rate.

Fragmentation by CID gives a characteristic series of 5'-end containing and 3'-end containing ions with the right energy usually only sampling one or a few ion types. HCD fragmentation will give more ion types which can make for a more crowded MS/MS, which can be advantageous in some cases and disadvantageous in others. Each requires tuning for energy…it may be best to leave some ion signal as unfragmented to limit the amount of "internal" fragment ions born from more than one fragmentation event—these are less useful for sequencing. Other fragmentation types can be useful, esp. UVPD.

The commercial software available (instrument vendor and 3rd party) is good for identifying almost all products in a digest; where the software may fail is when two sequence isomers co-elute and thus co-fragment.

Characterizing the poly(A) tail requires better MS resolution because of the convolution owing to the transcriptional-slippage-derived incorporation of more than the intended # of adenosine nucleotides. This may not be a problem if the mRNA is derived by means not using poly(A) encoded in the template DNA.

MS is but one part of the analytical package for RNA characterization. Translation competency of the RNA DS (if this is the MOA) cannot be evaluated by MS, but MS may be used as a tool to confirm or monitor the presence of the expressed target in an in vitro or in vivo expression study. This should be coordinated with flow cytometry or other analytical methodologies.