

## ProMass for MassLynx Software for Oligonucleotide Analysis

Vera B. Ivleva, Sean M. McCarthy, Ying Qing Yu, and Martin Gilar  
Waters Corporation, Milford, MA, USA

### APPLICATION BENEFITS

ProMass for MassLynx® Software enables batch mass deconvolution and data analysis for UPLC-MS data generated on both single quadrupole and time-of-flight Waters® MS instruments. ProMass for MassLynx is the only software available on the market that performs automatic assignment of the high-complexity chromatograms of an oligonucleotide. The software allows significant time savings in labs that routinely perform quality control and study *in-vitro* metabolism of native and modified samples.

### WATERS SOLUTIONS

[ACQUITY UPLC®](#)

[SYNAPT® HDMS®](#)

[Xevo® QToF MS](#)

[UPLC® Oligonucleotide Separation  
Technology Columns](#)

[MassLynx 4.1 Software](#)

### KEYWORDS

ProMass Software, batch mass deconvolution, peak assignments, oligonucleotide, RNA, siRNA

### INTRODUCTION

Oligonucleotides are an emerging class of biopharmaceutical compounds that includes antisense oligonucleotides (AO) and short interfering RNA oligonucleotides (siRNA), among others. Oligonucleotides are produced via automated stepwise synthetic methods. While synthetic procedures offer high yields, by-products of synthesis need to be removed and the desired product must be characterized prior to use in therapeutic applications.

UPLC-MS methods have been successfully applied to the analysis of the oligonucleotides. Due to the presence of failed products and a variety of modifications, the chromatograms can be complex and difficult to interpret even when LC-MS data are available. Manual processing of LC-MS data is laborious and time-consuming.

ProMass Software, currently used for LC-MS data processing by many laboratories analyzing oligonucleotides, has been recently adopted for MassLynx Software. ProMass for MassLynx enables automated data analysis in a high-throughput manner. This application note describes the utility of the software for characterization of truncated oligonucleotides, modifications, and other impurities.

## EXPERIMENTAL

**Oligonucleotide samples**

MassPREP OST standard (Waters) consisting of 15, 20, 25, 30, and 35 nucleotides (nt) long of oligodeoxythymidines (poly-T), ~ 1 nmole of each oligo per vial.

21 nt: 5'-UUC UGU AAU CUC UUG UCU ATT-3'

20 nt: 5'-UC UGU AAU CUC UUG UCU ATT-3'

(Integrated DNA Technologies, Coralville, Iowa)

All samples were reconstituted in 0.1-M triethylamine acetate (TEAA) to 40 pmole/ $\mu$ L.

**LC conditions**

LC system:	Waters ACQUITY® SQD
LC column:	Waters ACQUITY UPLC Oligonucleotide Separation Technology (OST) C <sub>18</sub> 1.7 $\mu$ m, 135 Å, 2.1 x 50 mm
Column temp.:	60 °C
Sample injected:	2.5 $\mu$ L
Flow rate:	0.2 mL/min
Weak wash:	Water
Strong wash:	50% methanol in water
Mobile phase A:	aqueous solution of 15 mM triethylamine (TEA, Fluka 23,962-3) and 400 mM hexafluoroisopropanol (HFIP, Fluka 32,522-8), pH 7.9; solutions were prepared by weight
Mobile phase B:	50% solvent A, 50% methanol, v/v

**MS conditions**

MS system:	ACQUITY SQD
Capillary:	4 kV
Sampling cone:	40 V
Extraction cone:	3 V
Source temp.:	150 °C
Desolvation temp.:	350 °C
Desolv. gas flow:	700 L/hr

**SYNAPT HDMS**

Capillary:	3 kV
Sampling cone:	35 V
Extraction cone:	3 V
Source temp.:	150 °C
Desolvation temp.:	300 °C
Desolv. gas flow:	500 L/hr
Lockmass calibration:	CsI 10 mg/mL (water/isopropanol, 1:1), 5 $\mu$ L/min flow rate, set mass = 1685.765

## RESULTS AND DISCUSSION

UPLC separation of the components was accomplished in less than 10 minutes using the MS-compatible ion-pairing mobile phase TEA/HFIP system (Figure 1). Manual assignment of the chromatographic peaks can be laborious and time-consuming, since each peak needs to be deconvoluted.

The manual peak assignment is based on two pieces of information: the highly accurate mass and the retention order of oligonucleotides, which depend on their relative length (shorter oligonucleotides elute before the longer ones). Shorter, 5' truncated failure sequences, typical by-products of synthesis, can be rather confidently assigned. Nevertheless, the manual data analysis is a very time-consuming process.

In order to facilitate faster and more convenient data analysis, ProMass for MassLynx Software was adopted for UPLC-MS chromatogram data processing.

Deconvolution of the components/peaks in the mass chromatograms is performed in the background using ZNova deconvolution algorithm. The software then compares deconvoluted masses of analytes to user-provided expected mass/sequence and the corresponding peak(s) are putatively assigned. Oligonucleotide-failed sequences can also be assigned in a similar fashion.

Figure 2 illustrates the software output, after data analysis of chromatogram shown in Figure 1. Two oligonucleotides, 35- and 30-oligodeoxythymidine sequences, were provided by the user. Software calculated their masses and identified two peaks eluting at 10.28 min and 9.64 min, corresponding to the expected masses. Information about mass error and relative abundance is also provided.

When specified in the processing method, 5' or 3' truncated failure sequences (or both) will also be interrogated by the software. Figure 2 illustrates the output of 5' truncated sequences above the chosen threshold. The dominant peaks of 15, 20, 25, 30, and 35 nt were assigned in an LC-MS chromatogram (see table in Figure 2). The retention time, mass error, and sequence assignment in relation to the target sequence are listed in the table. The target sequence in this case was 35 nt oligodeoxythymidine, the last dominant peak in the chromatogram.

ProMass for MassLynx Software typically processes the UPLC-MS chromatograms immediately after the data acquisition while the next sample in the queue is being analyzed by UPLC-MS. Alternatively, the previously acquired data batch can be assigned for off-line post-processing using the existing MassLynx sample list, which contains appropriate information required by ProMass (methods, deconvolution ZNova parameters, and sequence).

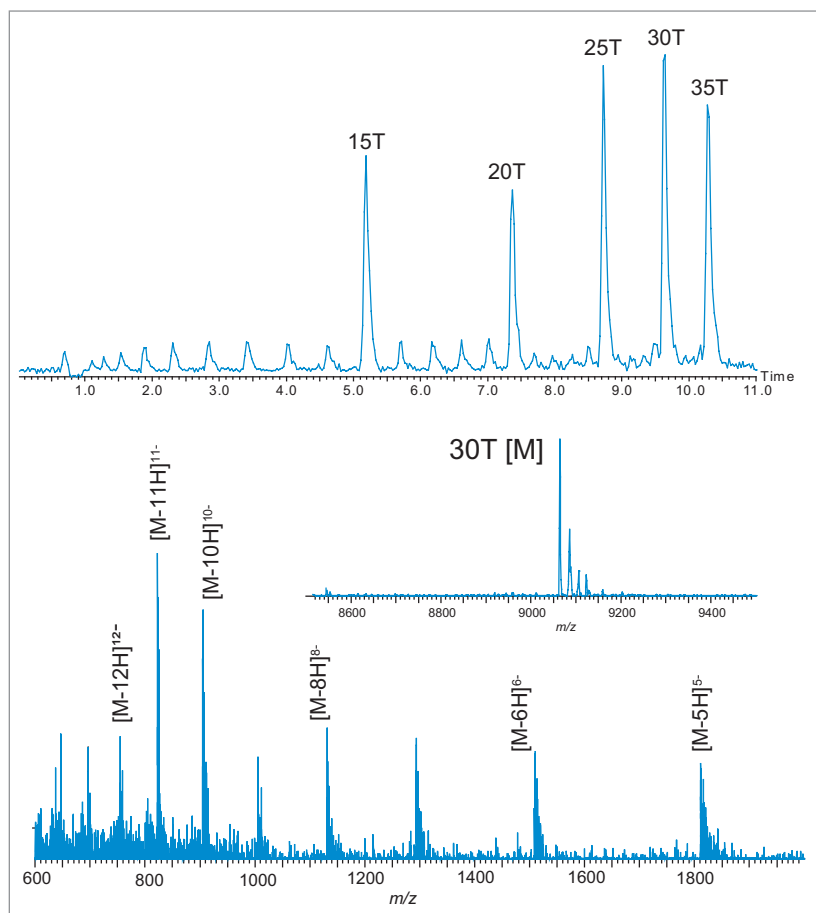


Figure 1. MassLynx chromatogram of OST standards; mass spectrum of 30 T (MaxEnt™ 1 deconvoluted spectrum on inset).

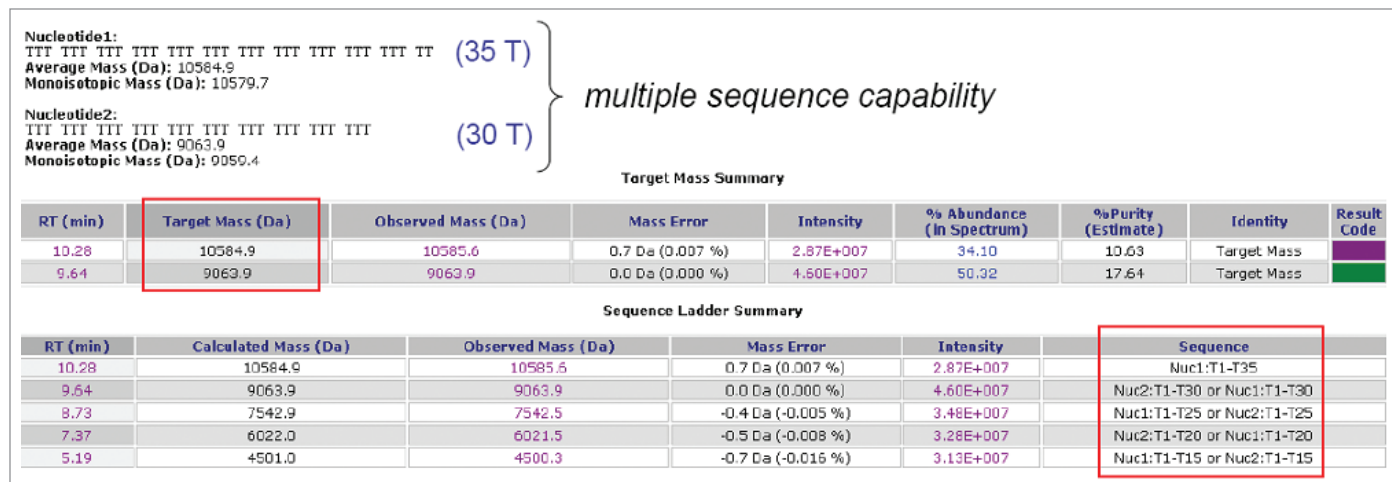


Figure 2. ACQUITY SQD MS data for DNA oligonucleotides.

The workflow of ProMass for MassLynx Software is presented in Figure 3. The software package provides a MassLynx sample list format with explanation of the function of each column field. Correct configuration of the sample list is essential for successful data processing. The user provides the raw data, defines the parameter files, and the optional details about the target product search.

## TWO PARAMETER FILES ARE NECESSARY FOR DATA PROCESSING

The parameter file of the application is called ProMassBridge, which provides a connection between MassLynx and ProMass. ProMassBridge takes care of the peak detection and reports the data to ProMass. ProMass software deconvolutes the spectra and reports the assignment of the peaks as an HTML file either in the form of tables, an Excel spreadsheet, or visualized as a 96-well plate.

In the ProMassBridge parameter file, the user defines the desired retention time range, smoothing and subtracting settings, and LockMass correction. The other parameter file is the one of ProMass called ZNova, which defines the input and output masses, ion polarity, deconvolution settings, target mass, impurities tolerance thresholds, and options for the data report.

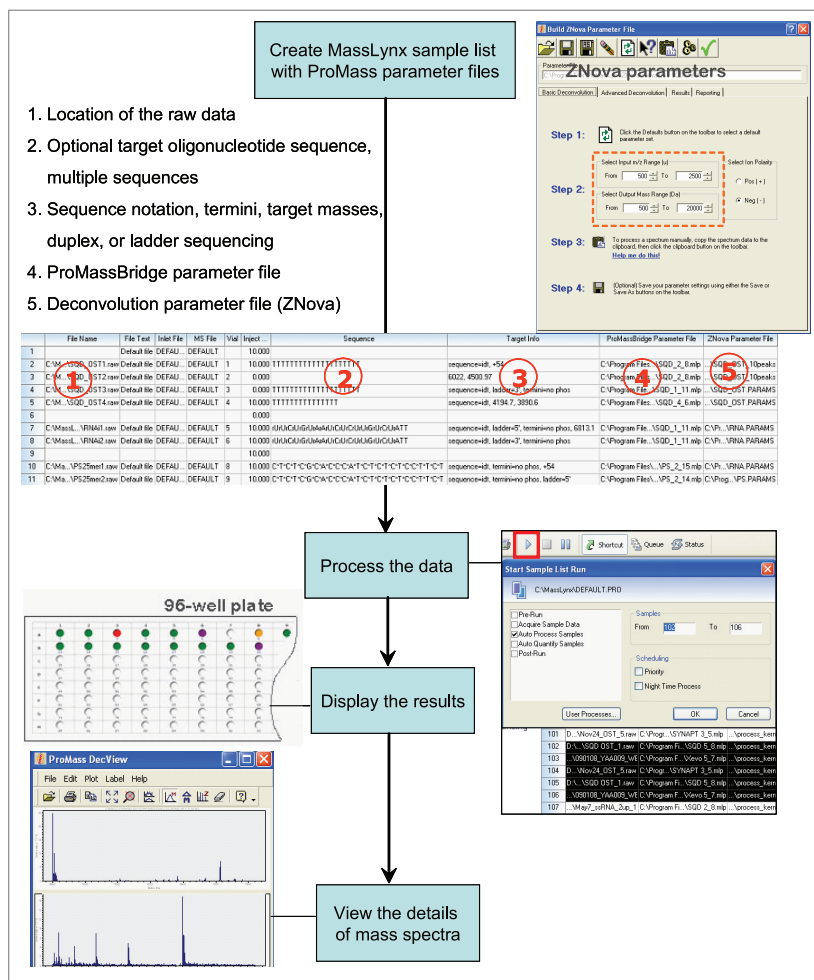


Figure 3. Overview of LC-MS data processing with ProMass for MassLynx.

In the sample set, the user has the option to define the analyte(s) of interest either as target mass, sequence, or combination of both. A variety of nucleotides modifications is supported, such as DNA, RNA, LNA, 2'-O-methylation, 2'-fluorination, or phosphorothioate modification. The user can specify any modification according to the Integrated DNA Technology (IDT) notations, or define custom modifications. The target information may include the mass of the targeted product, the adducts, termini modifications, and a request for ladder sequencing or processing of a duplex.

In high-throughput analysis mode, the user can display data summary as a 96-well plate with color-coded entries (Figure 3). The colors represent the result of the targeted mass(s) search. It is offered in interactive format to allow navigation through the chromatograms and the details of mass spectra. There is a general chromatographic summary listing the peaks assigned within the retention time range defined in ProMassBridge parameter file. The mass spectra corresponding to each chromatographic peak can be viewed in detail in the ProMass Viewer window.

In order to test the correctness of ProMass for MassLynx, the data of OST standards acquired on an ACQUITY SQD single quadrupole mass spectrometer were processed (Figure 2). Multisequence capability was investigated using 30 T and 35 T OST homologs as two targeted mass searches. Generally, more than two sequences can be entered. The sequences of both homologs were supplied, and the ProMass for MassLynx calculated the theoretical masses and searched for the corresponding values within the specified time range.

The deconvoluted mass of 30 T species resulted in 0.0 Da mass error, and was marked in green, meaning that the target mass found in chromatogram is the most abundant component within the suggested 0.02% mass error tolerance. The 35 T homolog was marked in purple, which happens when other components present in the spectrum have greater than 30% abundance. The deconvoluted masses of the failed sequences of the targeted 35 T oligonucleotide (30 T, 25 T, 20 T, 15 T) were reported in the sequence ladder summary. ProMass for MassLynx successfully identified the targeted masses and picked the correct picks as truncated sequences. The results correlated with manually deconvoluted masses of OST homologs (Figure 1).

To process the OST chromatogram, the reporting option in the ZNova parameter file was chosen to define the threshold of the chromatographic peaks selected for deconvolution so only the major homologs of the truncated sequences were taken into consideration (Figure 4). Assignment of the entire chromatogram is possible, as was demonstrated in Figure 4 in an example of 21-mer RNAi data obtained on an SYNAPT HDMS instrument. The molecular weight of the intact molecule was calculated based on the provided sequence and deconvoluted by ProMass for MassLynx with 0.011% mass error. In the Target Info column the request for the ladder sequencing was denoted as "ladder=5".

Generally, the search for the truncated sequences from either 5' or 3' ends is available as are internal cleavages. As a result, 17 truncated sequences were reported in the sequence ladder summary. The structures of these sequences are reported in the summary table rather than assigned directly on the chromatogram.

Using the automatic peak deconvolution, most of the 21-mer sequence of RNAi was confirmed in a matter of a few minutes. The speed of UPLC separation coupled with automated ProMass for MassLynx processing makes this method fast and efficient. The interactive interface covers the vast range of the user's needs: from a quick summary report to the details of the mass spectra.

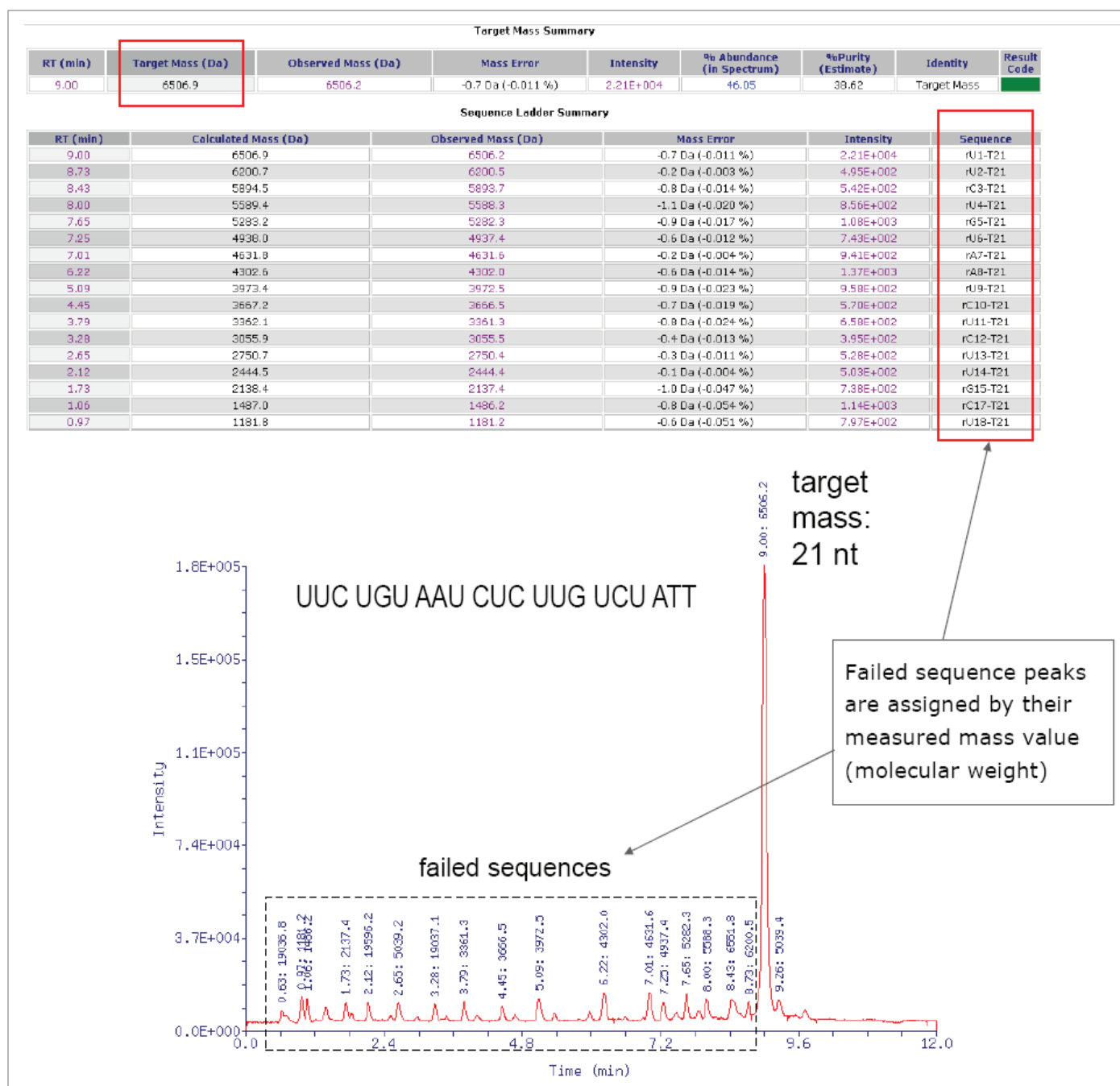


Figure 4. ProMass (for MassLynx) results visualized in browser as tables, chromatograms, and MS spectra. SYNAPT HDMS results for 21 nt siRNA oligonucleotide.

Analysis of the RNAi duplex is feasible if MS conditions allow for detecting the intact duplex peak. Two complementary RNAi strands were injected simultaneously to anneal into a duplex on a column. Consequently, ProMass for MassLynx identified both single RNAi strands and the peak corresponding to the sum of their masses composing the duplex mass (Figure 5). In order to perform this task, the target information should specify "sequence=duplex". In most instances, the duplex partially falls apart onto single RNAi strands upon electrospray ionization, so the same strand can appear twice on the MS chromatogram – once as a part of the decomposed duplex and again as the excess of upper or lower strand during the injection. If there is more than one chromatographic peak with the same mass, ProMass for MassLynx will match the one having highest abundance and report that one in the target mass summary report table. The quantifying aspect of ProMass for MassLynx was verified using 20 T and 25 T OST standards with injected concentration range 1 through 16 pmole.

As an alternative to the automatic high-throughput processing, MassLynx spectra can be deconvoluted manually one at a time simply by copying the spectrum list from MassLynx. It can be useful during the tuning of deconvolution parameters or for quickly processing the spectrum of interest.

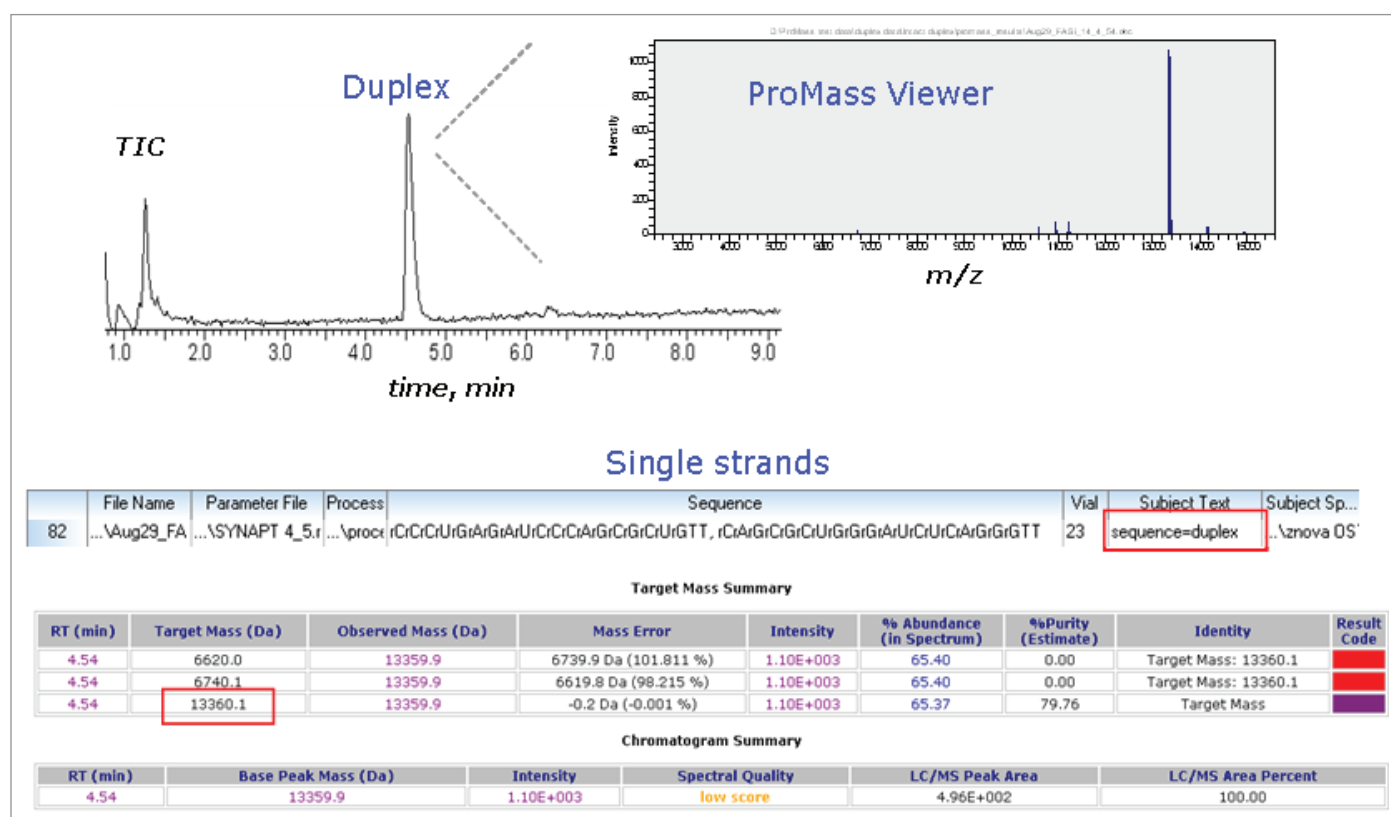


Figure 5. Analysis of intact RNA duplex.



## CONCLUSIONS

ProMass for MassLynx Software enables batch mass deconvolution and data analysis for UPLC-MS data generated on Waters MS instruments: single quadrupole and QToF. The software was successfully adopted for the data acquired in the MassLynx project. Confirmation of the analyte mass is consistent with the molecular weight of the entered target sequence.

ProMass for MassLynx assigns additional peaks on a chromatogram such as failed sequences/fragments and oligonucleotide modifications.

Multiple sequences and/or duplex oligonucleotides can be analyzed, which is a common task for the therapeutics applications.

The software significantly reduces the analysis times and enables automated UPLC-MS data processing in high-throughput mode. The interactively viewed color-coded results in web-based format are user-friendly and allow the researcher to navigate through the chromatographic data and the details of the mass spectra.

ProMass for MassLynx is the only software available on the market that performs automatic assignment of the high-complexity chromatograms of an oligonucleotide. The software allows significant time savings in labs that routinely perform quality control and study *in-vitro* metabolism of native and modified samples.

# Waters

THE SCIENCE OF WHAT'S POSSIBLE.®

Waters, The Science of What's Possible, ACQUITY, ACQUITY UPLC, UPLC, Xevo, MassLynx, and SYNAPT are registered trademarks of Waters Corporation. MaxEnt is a trademark of Waters Corporation. All other trademarks are the property of their respective owners.

©2010 Waters Corporation. Produced in the U.S.A. October 2016 720003577EN AG-PDF

Waters Corporation  
34 Maple Street  
Milford, MA 01757 U.S.A.  
T: 1 508 478 2000  
F: 1 508 872 1990  
[www.waters.com](http://www.waters.com)