[APPLICATION NOTE]

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

AN AUTOMATED DATA ANALYSIS OF THERAPEUTIC INTERFERON PROTEIN USING BIOPHARMALYNX APPLICATION MANAGER

Joomi Ahn, Asish B. Chakraborty, Ying Qing Yu, and John C. Gebler
Waters Corporation, Milford, MA, U.S.

INTRODUCTION

The combination of reversed phase liquid chromatography (RPLC) and electrospray ionization mass spectrometry (ESI-MS) has emerged as one of the most informative analytical platforms for biopharmaceutical characterization. Two complementary workflows have developed based on this technology platform.

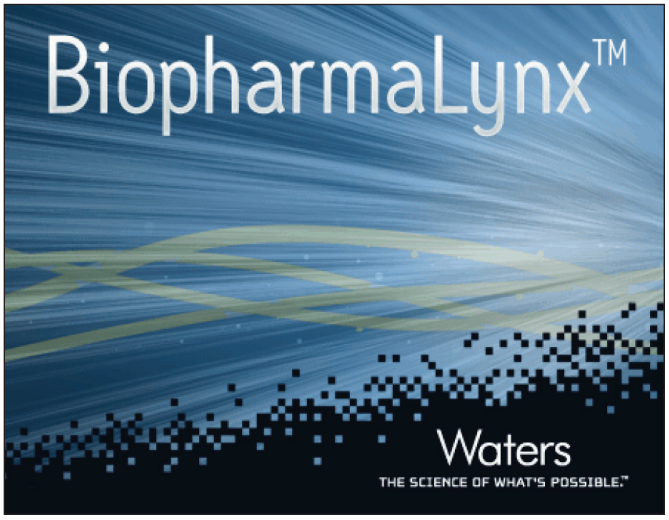
- The so called "top-down" approach applies LC/MS for profiling a biopharmaceutical protein at the intact protein level, obtaining information that confirms molecular weight (MW) and reveals evidence for product heterogeneity or the presence of non-product impurities.
- A second workflow, based on peptide map analysis of an enzymatic digested protein, typifies a "bottom-up" analysis workflow. The analysis complexity introduced by the enzymatic digestion is offset by the large increase in information about the biopharmaceutical sample. This workflow reveals and localizes sites of heterogeneity and rapidly quantifies structural differences between samples.

As HPLC has evolved with UPLC® Technology, and time-of-flight (TOF) mass spectrometry has become more powerful and easier to use, both data quality and quantity have dramatically increased. However, this continued evolution in data acquisition technology results in many laboratories becoming productivity-limited not by LC/MS data acquisition delays, but by the processing of this data.

Intact protein LC/MS analysis has traditionally been a repetitive task of spectral summation, spectral deconvolution, and manual assignments of deconvoluted masses to product variants. As the number of routine intact mass analyses grows, so does the desire for batch processing of intact protein LC/MS data and automated annotation of protein structures to the data.

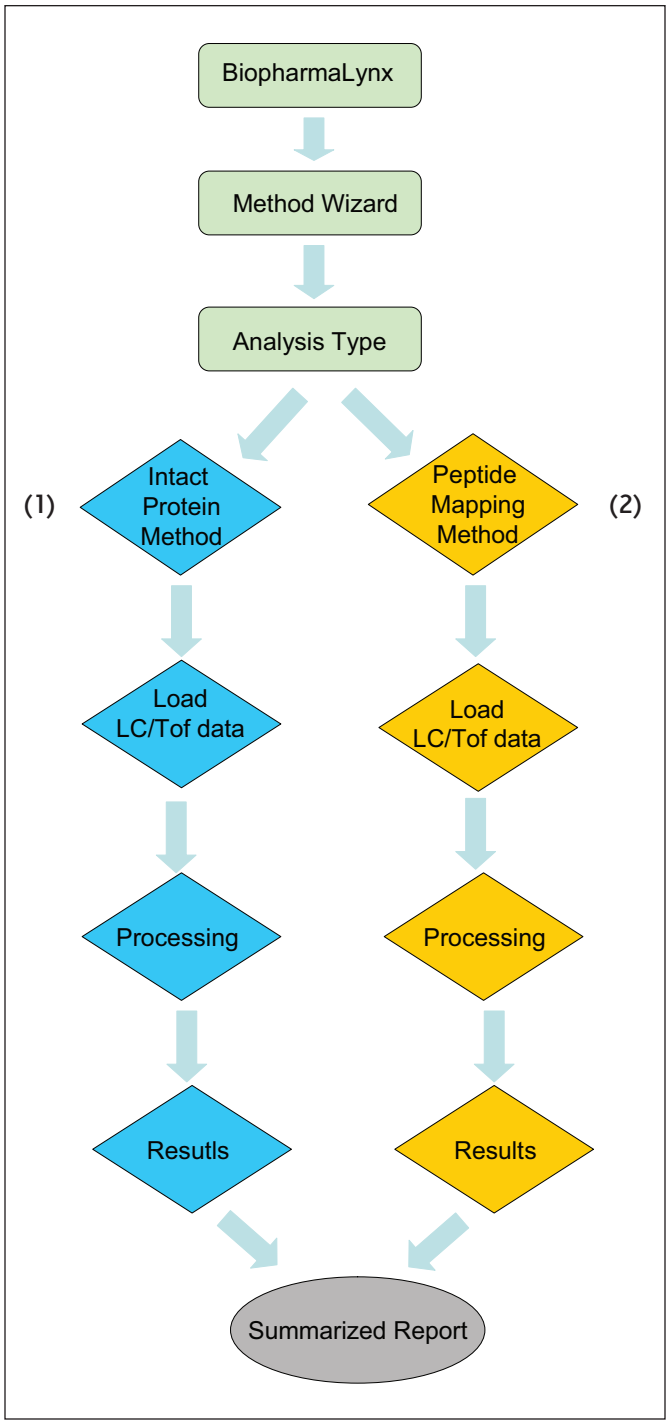
LC/MS peptide mapping analysis involves much greater levels of repetitive data processing for each run of acquired data, and presents significant challenges when inter-run comparisons are required. Many labs have turned to proteomic software tools to automate data processing, only to find that these tools are rather blunt for biopharmaceutical analysis. The proteomic focus on protein-level identification rather than peptide-level characterization results in overall poor characterization of a protein, and a significant lack of information about the sample. Proteomic software also lacks appropriate functionality for structural investigations, such as disulfide mapping, glycopeptide characterization, and quantitative investigations into sample-sample differences.

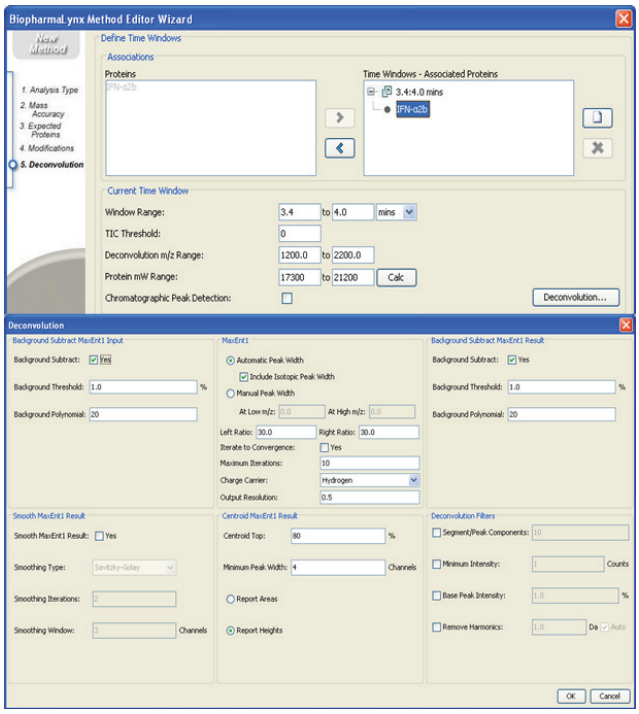
In this application note, we focus on the application of BiopharmaLynx™ Software, an informatics tool that automates processing and analysis of both intact mass and peptide mapping biopharmaceutical workflows. A case example is presented for recombinant Interferon produced by two cell culture conditions, and the ability of the software to process, annotate, and compare the results from the LC/TOF-MS analysis of these two samples. The benefits of this combined intact mass-peptide mapping approach for readily detecting and highlighting the structural differences between these samples is illustrated.



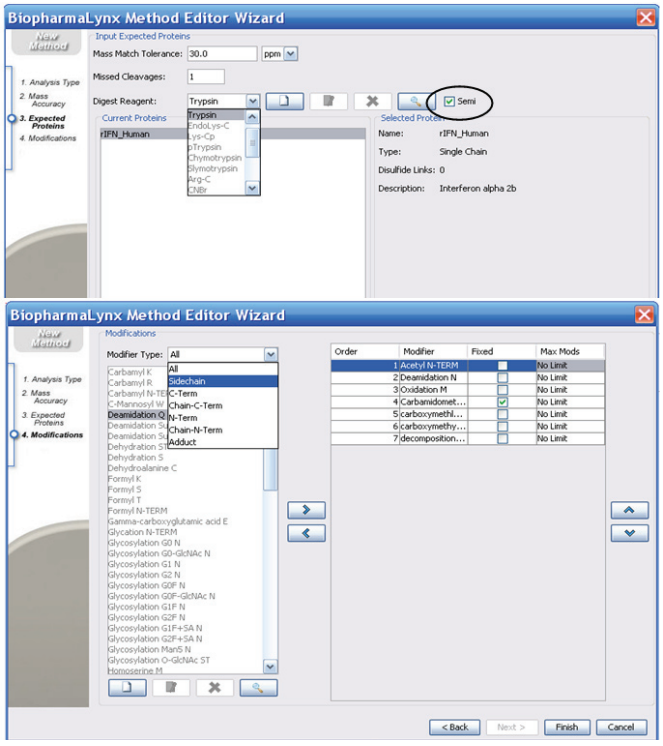
METHODOLOGY

The biopharmaceutical workflow using BiopharmaLynx 1.1





(1) Method wizard: Intact protein.



(2) Method wizard: Peptide mapping.

RESULTS

Intact Interferon protein analysis

Our experimental goals are to obtain intact protein MW, compare the MW of Interferon expressed from two cell culture conditions (control vs. analyte), provide purity assessment, and search for post-translational modifications (PTMs). Details of the LC/MS experiments are listed in the Experimental Details section. An example of the BiopharmaLynx Intact Protein Analysis Method Editor is shown in the BiopharmaLynx 1.1 workflow.

There are multiple ways to display results in the BiopharmaLynx browser. Each display reveals a different aspect of the data. For example, the display of *Total Ion Chromatogram* (TIC) gives the chromatographic retention time for each component, the *Raw Spectrum View* displays the charge envelope of the intact protein, the *Processed Spectra View* shows the deconvoluted mass spectrum and relative intensities of components, and the peak match data table summarizes all the above information in a spreadsheet format (Figures 1 and 2).

Figure 1. Results display for control and analyte Interferon protein.

A) Total Ion chromatogram (TIC) where the main Interferon peak is highlighted. The later eluting peaks represent product-related impurities..

B) Raw Spectrum view: m/z of multiply charged ions of Interferon (+9 to +15 ion series were selected and summed for mass deconvolution).

C) Deconvoluted intact Interferon mass spectra in mirror view. The Δmass of + 42 Da observed between control and analyte results from mono-acetylation of the analyte sample. Also, minor components (~6%) representing oxidation (+16 Da) of the major form were observed for both samples.

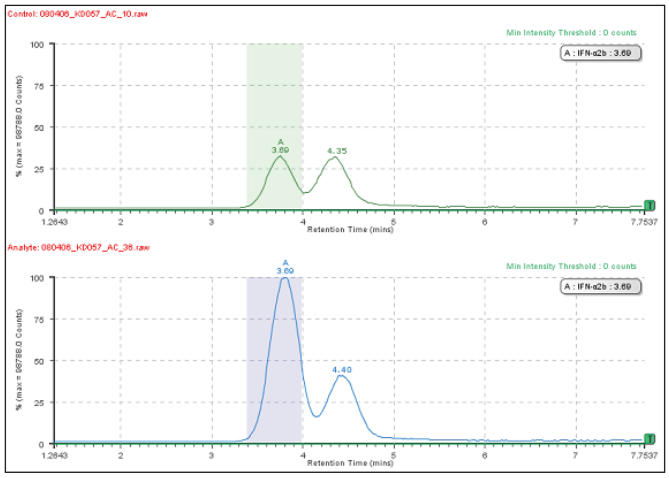


Figure 1A. Total ion chromatogram (TIC) view.

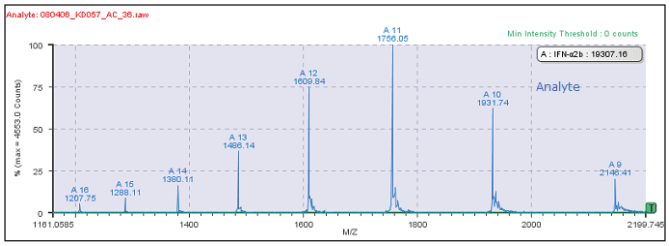


Figure 1B. Raw Spectrum view.

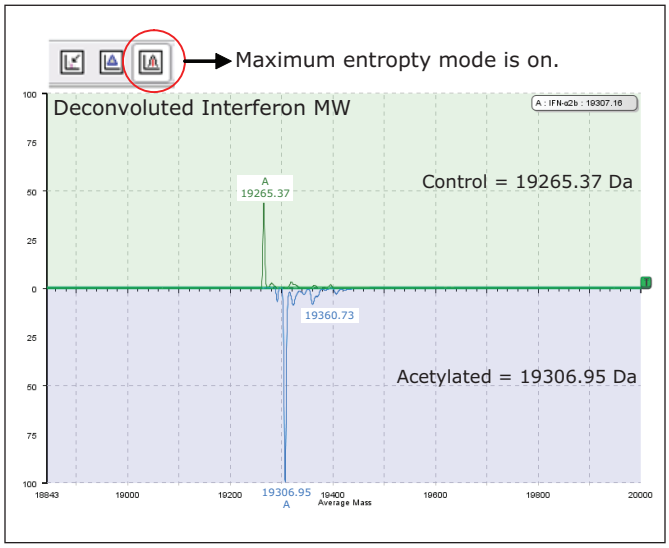


Figure 1C. Processed spectra view.

Chromatogram | Spectrum | Coverage Map | Protein Digests | Peak Match Data

Protein	Modifiers	Calculated Protein Mass (Da)	Control Intensity (Counts)	Control Mass Error (ppm)	Analyte Intensity (Counts)
IFN-α2b		19265.1543	24497.1	11.3	
IFN-α2b	Oxidation M(1)	19281.1484	1409.8	3.9	
IFN-α2b	Acetyl N-TERM(1)	19307.1641			55893.1
IFN-α2b	Acetyl N-TERM(1),Oxidation M(1)	19323.1602			5362.1
IFN-α2b	Acetyl N-TERM(1),Oxidation M(4)	19371.1445			2841.1

Figure 2. Peak match data table displays the mass spectral results annotated with the targeted protein search results. Information such as modifications, intact protein mass, and component are listed. These search results can be customized, filtered, and sorted by any category. Intensity thresholds can be adjusted to prevent false assignment to lower level spectral noise peaks.

Peptide mapping analysis

With the knowledge gained from the intact protein analysis, peptide mapping analysis in BiopharmaLynx can focus on more specific questions. A common goal for peptide mapping is to confirm protein identity and characterize the primary protein structure with a high coverage LC/MS peptide map. Secondary goals usually involve annotating modifications (e.g., finding the acetylation site suggested by analyte intact protein data), and producing comprehensive sample-to-sample comparisons. Again, these results can be viewed and compared in different formats (sequence coverage map, TIC, spectrum or a data table).

The high mass accuracy from TOF MS data (< 10 ppm routinely with lock mass correction) enables confidence in producing assignments of ions to peptides generated by proteolytic digests. To more comprehensively assign all the ions detected during the LC/MS analysis, the analysis of semi-digested can be enabled. Even under optimal conditions, the electrospray process can generate in-source fragments of labile peptides, and other ions representing neutral loss of water and ammonia. BiopharmaLynx 1.1 now recognizes and annotates these ions, and allows users to expand the range of modifications and MS adducts beyond those found in the default library.

Figure 3 shows results in a protein sequence coverage map display. High protein coverage (97.0% and 96.4% in control and analyte respectively) was achieved. BiopharmaLynx also identified the acetylation site as the N-terminus cysteine in the analyte. This finding correlates the intact protein results (Figure 1C). The color coding scheme depicted in the left panel of Figure 3 allows users to readily visualize control vs. analyte sample differences. As an example, the acetylation unique to the analyte sample is highlighted with an orange hue.

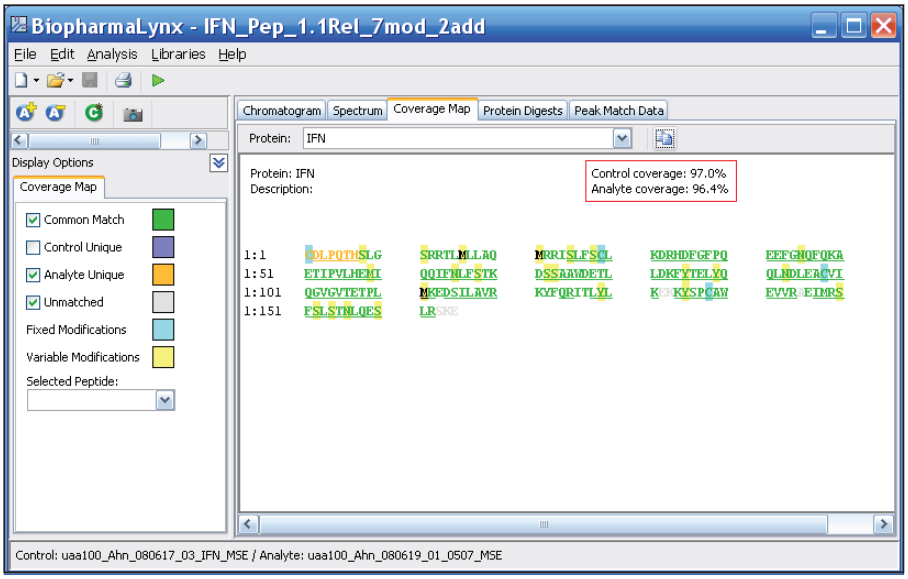
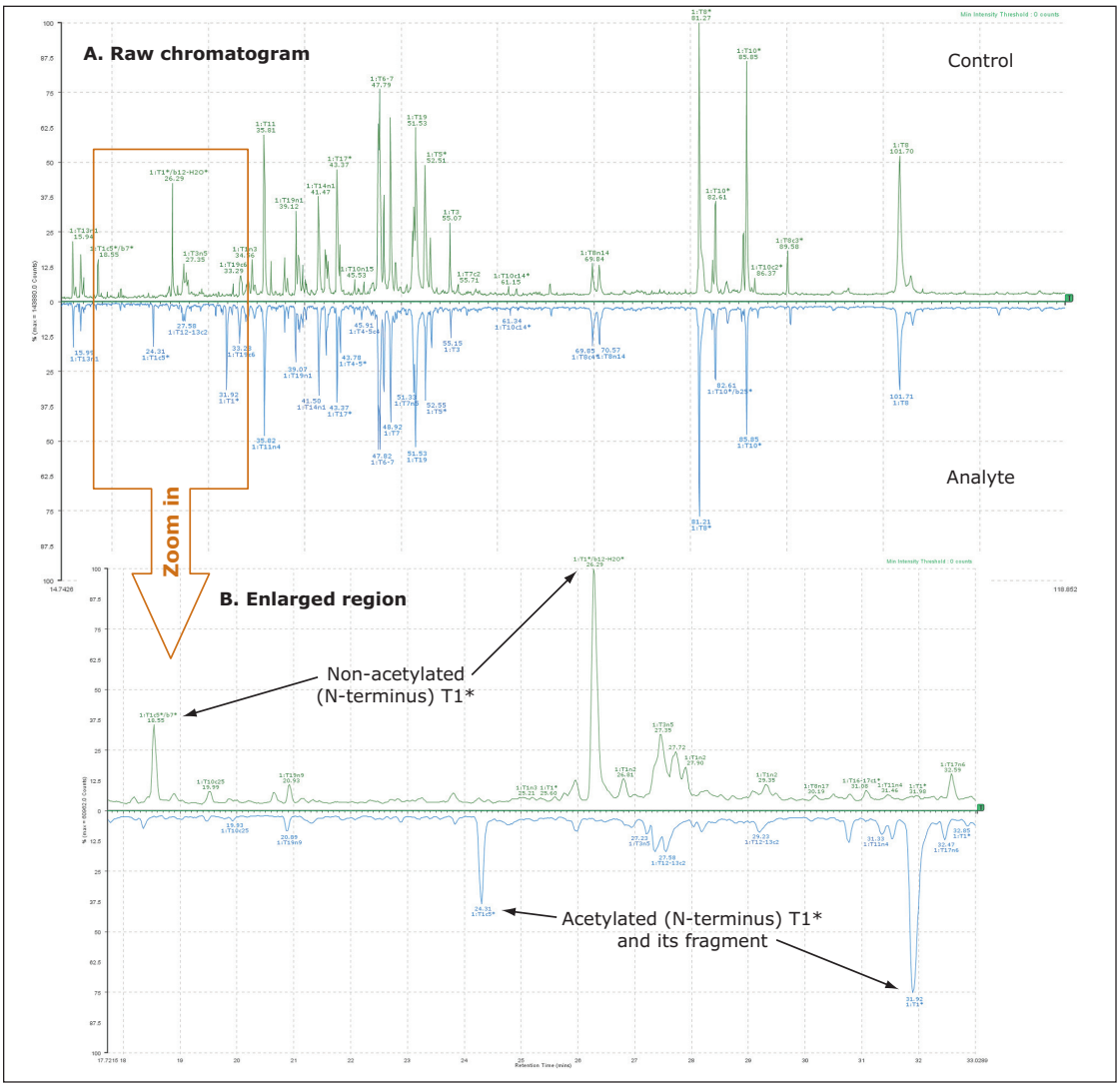


Figure 3. Protein sequence coverage for control and analyte Interferon samples.

The raw TIC display depicts the overall elution pattern for all peptides detected during the LC/MS map analysis. Visualizing significant differences between samples can often be accomplished using the mirror plot view of the data (Figure 4A). Tryptic peptide peaks are automatically annotated using Tryptic peptide fragment numbers (T_n) and retention time in chromatograms.

A segment of the chromatogram in Figure 4A was enlarged to show peaks that do not align between the control and the analyte samples (Figure 4B). These peaks reflect T1 (CDLPQTHSLGSR) modification differences: the control T1 peptide was *carbamidomethylated* (during protein reduction/alkylation step) while the analyte T1 peptide was *carbamidomethylated* and acetylated.



(eg. T --- Tryptic peptides (T1, first peptide from N-terminus)
** --- post translational modifications*
n/c --- partial enzymatic cleavage from C- terminus (or N-terminus), e.g., T1/c5 or T1/n2
y/b --- y (or b) series of in-source fragment ions of T1, eg. T1/y11 or T1/b3

Figure 4. (A) TIC (mirror view) of the two Interferon peptide maps. A region of the chromatogram that shows unique peptides for each sample is highlighted. (B) The enlarged chromatogram region indicates that these unique peptides were all generated from Interferon T1 tryptic peptide and a semi-tryptic fragment of this peptide that differ between samples by acetylation of the N-terminus.

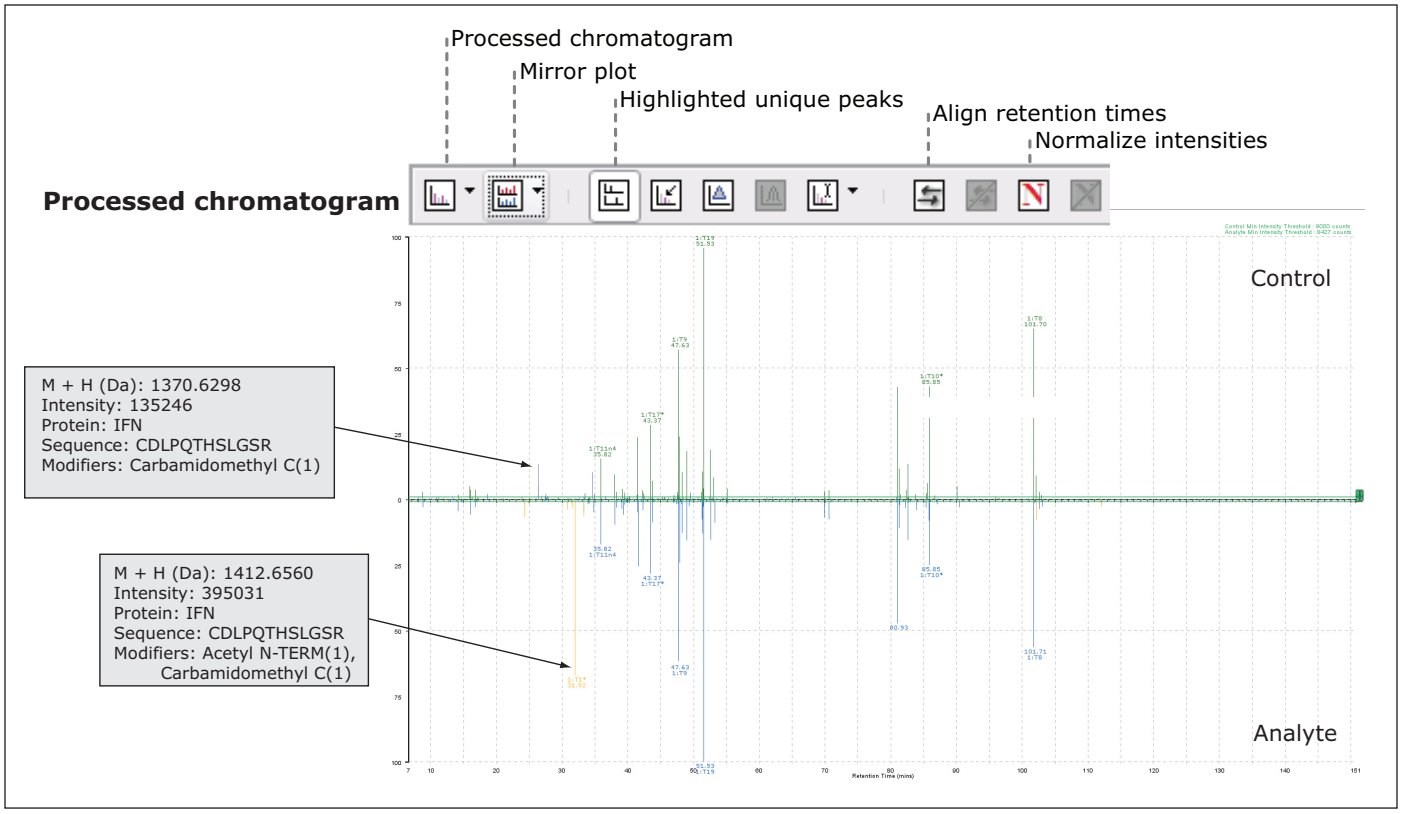


Figure 5. The processed chromatograms displayed in mirror plot with highlighted unique peaks in purple and orange colors for control and analyte interferon peptides.

More sample information is highlighted in Figure 5, which displays the mapping results as a component-centric processed chromatogram view. Each centroided "stick" in the processed chromatogram represents the summed intensity (ion counts) of all isotopes from all detected charge states of a peptide over its full chromatographic elution profile. This simplified view enables several useful tools potentially needed for efficient sample-sample comparisons, including alignment of LC retention times and normalization of ion intensities for better relative quantitation results. Additional information about peptide assignment is accessible by moving the mouse over any given peak.

CONCLUSION

In this study, we have demonstrated how automated BiopharmaLynx data processing promotes rapid comparisons of two batches of Interferon produced by different expression systems. Both the top-down intact protein analysis and the bottom-up peptide mapping analysis supported the finding that the two batches differed primarily by the presence of N-terminal acetylation in one sample.

BiopharmaLynx is a key addition to the Waters UPLC/TOF-MS system solution for biotherapeutic protein characterization. BiopharmaLynx automates most aspects of data processing, protein/peptide annotation, sample comparison, and report generation, which are currently viewed as the major productivity bottlenecks for most biopharmaceutical characterization laboratories. Routine characterization studies that take well-trained scientists several days to weeks to manually process can be batch processed by BiopharmaLynx in minutes. The laboratory productivity gained by allowing scientists to focus on higher-value activities rather than routine data analysis should permit development organizations to expand the breadth of their activities and streamline timelines existing investigations.

Waters

THE SCIENCE OF WHAT'S POSSIBLE.™

January 2009 720002950en 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

SUPPLIMENTAL EXPERIMENTAL

Intact Protein LC/MS

Sample preparation

Two recombinant Interferon expressed under different cell conditions were used (control and analyte). Interferon samples were reconstituted in 50 mM ammonium bicarbonate buffer, micro-centrifuged, and injected onto the column for LC/MS analysis. Additional details on the method can be found in Waters Application note 720002107en.

LC conditions

LC system:	Waters® ACQUITY UPLC® System
Column:	MassPREP™ Micro Desalting Column (2.1 x 5 mm)
Mobile phase A:	0.1% Formic acid in water
Mobile phase B:	0.1% Formic acid in acetonitrile
Gradient:	Total run time: 10 min 5-10 %B in 0.5 min, 0.2 mL/min (waste) Then 10-20 %B in 0.01 min, 0.2 mL/min (MS) Then 20-45 %B in 7.1 min (MS), 0.2 mL/min Then 3-sawtooth 90-5 %B in 0.5 min, 0.5 mL/min
Column temp.:	80 °C

MS conditions

MS system:	Waters LCT Premier™ ESI-TOF MS
Ionization mode:	ESI Positive, V mode
Capillary voltage:	3200 V
Cone voltage:	40 V
Desolvation temp.:	350 °C
Source temp.:	150 °C
Desolvation gas:	800 L/Hr
Ion guide 1:	5 V
Acquisition range:	600 to 3000 m/z

Peptide mapping LC/MS

Sample preparation

The control and analyte were dissolved in 50 mM ammonium bicarbonate and denatured with 0.025% RapiGest™ SF (Waters), then reduced and alkylated with DTT and iodoacetamide (Sigma) respectively. Finally the protein was digested with porcine trypsin (Promega).

LC conditions

LC system:	Waters ACQUITY UPLC System
Column:	ACQUITY UPLC BEH 300 C_{18}, 2.1 x 150 mm, 1.7 µm 300Å
Flow:	200 µL/min
Mobile phase A:	0.1% Formic acid in water
Mobile phase B:	0.1% Formic acid in acetonitrile
Gradient:	0-50 %B over 120 min
Column temp.:	60 °C
Weak wash:	95 % buffer A / 5 % buffer B
Strong wash:	20 % buffer A / 80 % buffer B

MS conditions

MS system:	Waters SYNAPT™ MS
Ionization mode:	ESI Positive
Capillary voltage:	3.0 kV
Cone voltage:	35 V
Desolvation temp.:	250 °C
Desolvation gas:	350 L/Hr
Source temp.:	120 °C
Acquisition:	50 to 1700 m/z
Lockmass:	100 fmol/µL Glu-Fibrinopeptide B $(M+2H)^{2+}$ =785.8426