

# Deconvolution of Multiply Charged States of Intact Proteins

Natalia Belyaeva, Tonya Pekar Second, Jim Shofstahl  
Thermo Fisher Scientific, San Jose, California, U.S.A.

## Overview

**Purpose:** To deconvolve multiple charges states of different compounds in the sample into neutral masses. To resolve coeluted charge states of different compounds. To do accurate relative quantification of different modifications and isoforms of the same compound.

**Methods:** Use signal processing and statistical analysis to determine possible charge state of peptides and proteins and create possible charge state chains. Based on the maximizing the score of different possible charge state chains, choose the optimal charge-state chain and deconvolve charge states from  $m/z$  domain to mass domain. Based on identified masses, do deconvolving signal with correspondent peak shape and resolve a signal from coeluted charge states.

**Results:** The algorithm allows not only determination of accurate masses of compounds but also to deconvolve coeluted charge states and identify and characterize different modification patterns such as glycosylation or drug conjugation. With peak shape modeling, accurate relative quantification of different modifications is being achieved.

## Introduction

Analysis and identification of intact proteins and peptides through LC/MS is widely used in biochemical laboratories. Even with high-resolution instruments, high molecular weight intact proteins are usually isotopically unresolved. There is great interest in the identification of different protein modifications, such as glycosylation or oxidation. Deconvolution based on graph theory and using appropriate scoring will allow identification of zero-charge masses and relative quantification of their modification with high accuracy and reliability.

## Methods

Typical individual spectra from LC/MS do not produce enough charge states to perform deconvolution. The first step is to average enough spectra to provide good statistics for different charge states and decrease random noise. On a formal graph model, all unknown charge states are presented as possible states. The relation between different states is formalized as the probability of belonging to the same mass. Then all charge states belonging to the same mass present a charge state chain. The problem of deconvolution can be formalized as the problem of search of all optimal paths in the graph. The scoring function assigned to the optimal charge state reflects the reliability of identification. From all possible charge state chains the chain with best score is assigned to a mass. The algorithm works for both high resolution and low resolution instruments.

## Results

Several methods are used for deconvolution of intact proteins and peptides, such as maximum entropy. The maximum entropy method demands the evaluation of peak width. It is not always possible to identify peak width with enough accuracy. An additional problem is that peak width is changed along  $m/z$  region, even for the same compound. An alternative method based on graph search of all optimal paths is free from this problem. It also works well for co-eluting charge states. It is demonstrated for IgG that a deconvolved spectrum, after appropriate averaging to get enough different charge states and improve charge state envelope, reveals multiple glycosylation sites. Other antibody variants are revealed as well. Analysis of separate charge states with clear variant patterns confirms the validity of modifications in the deconvolved signal.

For complex spectra there is a possibility of randomly assigning to a wrong charge state. A score is assigned to each determined charge state chain based on the quality of peaks, matching to the mass, and to the charge state envelope shape. It is shown that high accuracy deconvolution provides not only correct measurement of intact protein molecular weights but also allows accurate quantification.

FIGURE 1. Deconvolved spectrum of Sigma IgG (FTMS spectrum on an LTQ Orbitrap XL mass spectrometer). Most of charge-state clusters reveal nine glycosylation sites. Deconvolved spectrum reveals the same glycosylation pattern as individual modification clusters correspondent to the same charge.

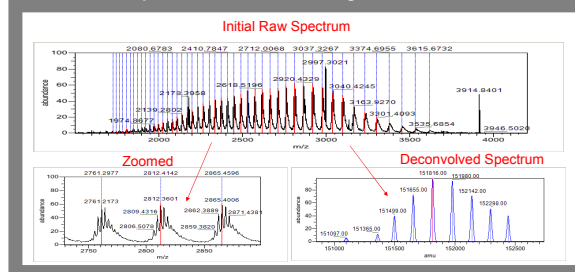


FIGURE 2. Deconvolved spectrum of Sigma IgG on Orbitrap instrument. Charge states for identified mass 151,816 Total 47 charges (from 88 to 42) are identified

MZCentroid	MassCalculated	Charge
1,726.2460	151,821.0112	86
1,746.1475	151,827.1954	87
1,766.2828	151,813.6956	88
1,787.2252	151,828.5233	88
1,808.3678	151,818.2829	84
1,830.1841	151,821.6735	85
1,852.4572	151,818.8954	82
1,875.3238	151,819.6385	81
1,898.7415	151,818.7347	80
1,922.8554	151,826.0064	78
1,947.3017	151,810.9659	78
1,972.3468	151,823.5599	77
2,000.0000	151,818.0000	76
2,028.4006	151,812.8442	53
2,050.4329	151,810.1326	52
2,077.8027	151,816.5688	51
3,037.2598	151,812.6251	50
3,089.2968	151,816.1854	49
3,163.7936	151,813.7414	48
3,231.1553	151,816.9556	47
3,301.4093	151,818.4939	46
3,374.8687	151,823.7620	45
3,451.7268	151,831.6574	44
3,531.7165	151,820.4972	43
3,615.8334	151,822.6963	42

The method works well for high-resolution spectra and for low-resolution spectra. It can be applied to isotopically resolved data also. The deconvolution algorithm was applied for to data acquired using both Orbitrap and ion trap mass spectrometers.

On Figure 1, an averaged spectrum (450 averaged FTMS spectra from orbitrap instrument) of Sigma IgG is shown. Averaging before deconvolution gives sufficient ion statistics for different charge states, decreases noise, and random spectrum-to-spectrum variation. ESI provides multiple highly charged ions. On Figure 2, identified charge states for the Sigma IgG for one particular identified mass (151.816 amu) from 88 to 42 are shown. Nine different glycosylation modifications are identified for this sample. For this particular sample masses of different modifications calculated based on multiple charge states have standard variation about 3-15 Da (3 amu for strong modifications and 15 amu for weaker modifications), which gives CV about 0.002-0.01% respectively. It is possible to identify up to 50-60 different charge states for appropriately averaged high accuracy spectra from orbitrap instrument.

The same approach gives good results for low-resolution data as well. On Figure 3, an averaged spectrum of Sigma IgG (60 averaged spectra from a linear ion trap) and deconvolved spectrum are shown. Between 45 and 50 different charge states are identified for strong modifications and between 25 and 30 for weak modifications (glycosylation). STDs of identified masses vary between 8 and 20 amu. Several modifications are identified for this sample (including glycosylation).

For intact IgG, due to its large mass, glycosylation modifications are practically the smallest main modifications that can be identified. By chemical reduction of disulfide bonds and producing light and heavy chains, richer modification structure can be further identified and quantified. In Figure 4, an averaged spectrum of heavy chain Fab of reduced IgG (24 averaged FTMS spectra from a Thermo Scientific LTQ Orbitrap XL hybrid mass spectrometer) is shown. In addition to main glycosylated sites, a lot of smaller modifications can be identified and quantified.

Also, an averaged spectrum (20 averaged FTMS spectra) of light chain Fc and its deconvolved spectrum are shown. Three main components (with zero, one, and two galactose residues respectively).

If a sample contains a lot of agents that form a non-covalent adducts or contain a lot of small modifications such as an oxidation or a methylation, its spectrum will have skewed long tail peaks. With intelligent centroiding based on the quality of the data, deconvolving of such spectrum will give correct results.

In Figure 5, an averaged spectrum (150 averaged ITMS spectra from an LTQ Orbitrap XL™ instrument) of enolase is shown. Individual charge states have very long tails; and only a few charge states (in high-mass range) have partially resolved modifications. Deconvolved raw spectrum (blue dashed line) would hide these modifications. Correctly identified masses (black solid lines) are shown on deconvolved spectrum. The main mass (46,681.45) has an STD of 13.5, which is very good for this quality of data.

FIGURE 3. Deconvolved spectrum of Sigma IgG (ITMS spectrum on an LTQ Orbitrap XL mass spectrometer). Most of charge-state clusters reveal the same glycosylation sites. Based on similarity of glycosylation patterns of individual modification clusters, weak modifications can be confirmed.

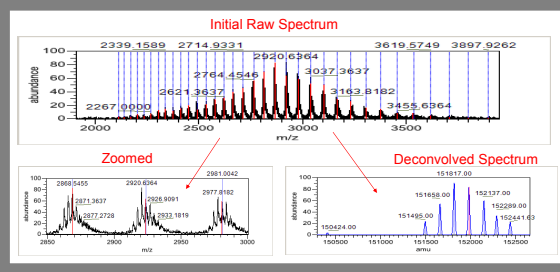


FIGURE 4. For heavy chain of reduced IgG, additional modifications (in addition to main glycosylation sites) can be identified and quantified.

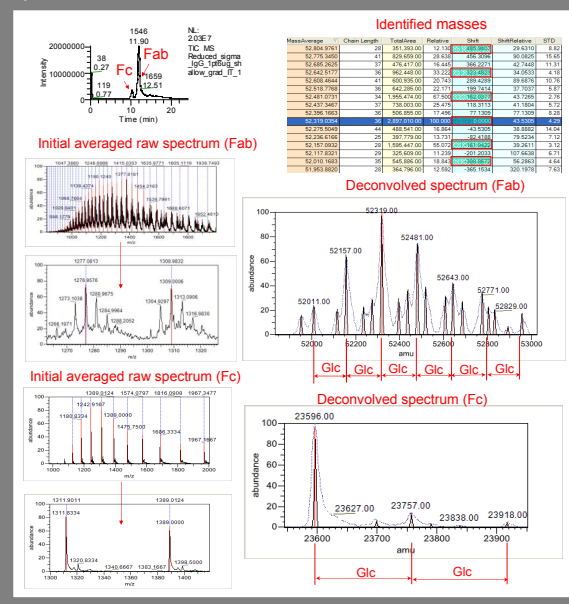
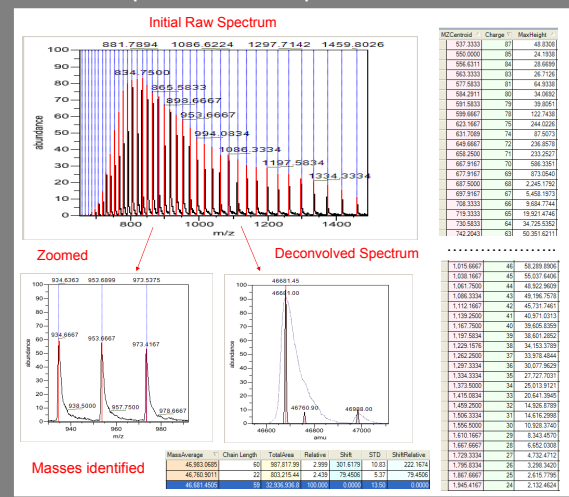


FIGURE 5. With theoretical peak modeling, additional modifications can be identified. Blue line – raw deconvolved spectrum, black line – deconvolved spectrum for modeled peaks



## Conclusion

- Identification and quantification of intact proteins and peptides (for which MS spectra are isotopically unresolved) using electrospray and high-resolution mass spectrometry allows processing of very large proteins (150 K and higher) and provides reliable analysis.
- The relative quantification of different proteins and their modifications can be done very accurately.
- The same approach works well for high-resolution and low-resolution data (LTQ and LTQ Orbitrap XL mass spectrometers).
- Deconvolution of the algorithm resolves coeluted charge states successfully.

## References

- Mann, M. et al, Interpreting Mass Spectra of Multiply Charged Ions, *Anal. Chem.* 1989, 61, 1702-1708.
- Michael W. Senko et al, Automated Assignment of Charge States from Resolved Isotopic Peaks for Multiply Charged Ions, *J. Am. Soc. Mass Spectrom.* 1995, 6, 52-56

All trademarks are the property of Thermo Fisher Scientific and its subsidiaries.