

Deconvolution of Isotopically Unresolved Multiply Charged States of Intact Proteins and Peptides

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



Overview

Purpose: To deconvolve multiple charges states of different compounds in the sample. To resolve coeluted charge states. 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 stated of peptides and proteins and create possible charge state chains. Based on the maximizing the score of different 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: Algorithm allows not only to determine accurate masses of compounds but also to deconvolve coeluted charge states and identify different modification patterns such as glycosylation. 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 has become a popular method in biochemical laboratories. Even with high resolution instruments, high molecular weight intact proteins can only be partially isotopically resolved. If multiply charged states are present in the spectrum and have sufficient statistical data, then deconvolution of unresolved spectra can be done and neutral masses identified.

There is great interest in the identification of different protein modifications, such as glycosylation, oxidation, etc. Deconvolution of mass spectra will allow characterization of proteins, peptides and their variants. Formulating the deconvolution task as graph theory problem and using appropriate scoring will allow identification of zero charge masses and 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. 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.

Results

Several methods are used for deconvolution of intact proteins and peptides such as Max Entropy. 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.

FIGURE 1. Deconvolved spectrum of Sigma IgG (FTMS spectrum on an Orbitrap™ mass spectrometer. Most of charge state clusters reveal nine glycosylation sites. Deconvolved spectrum reveals the same glycosylation pattern (both initial and deconvolved spectra are shown as raw spectra, without peak modeling or baseline removal).

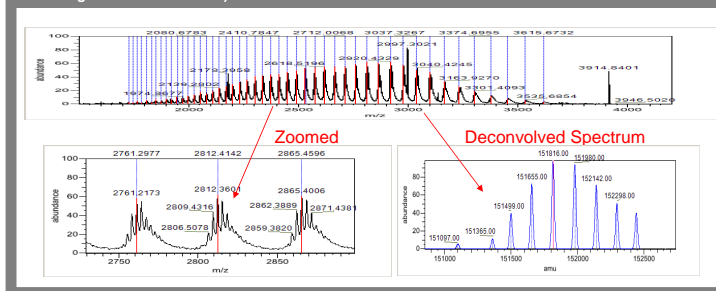


FIGURE 2. Deconvolved spectrum of Sigma IgG on orbitrap. Charge states for identified mass 151,816. Total 47 charges (from 88 to 42) are identified

MZCentroid	MassCalculated	Charge	%
1,726.2460	151,821.0112	88	8%
1,746.1475	151,827.1954	87	7%
1,766.2828	151,813.6956	86	6%
1,787.2252	151,828.5233	85	5%
1,808.3678	151,818.2829	84	4%
1,830.1841	151,821.6735	83	3%
1,852.4572	151,818.8954	82	2%
1,875.3238	151,819.6385	81	1%
1,898.7415	151,818.7347	80	1%
1,922.8554	151,826.0004	79	1%
1,947.3017	151,810.9659	78	1%
1,972.7469	151,823.9900	77	1%
1,999.6761	151,816.0744	76	1%
2,027.2001	151,817.0795	75	1%
2,055.4006	151,812.8442	53	1%
2,920.4329	151,810.1326	52	1%
2,977.8027	151,816.5688	51	1%
3,037.2598	151,812.6251	50	1%
3,099.2968	151,816.1864	49	1%
3,163.7936	151,813.7414	48	1%
3,231.1553	151,816.9566	47	1%
3,301.4093	151,818.4939	46	1%
3,374.8687	151,823.7620	45	1%
3,451.7268	151,831.6574	44	1%
3,531.7165	151,820.4972	43	1%
3,615.8334	151,822.6963	42	1%

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.

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 traps.

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 averages high accuracy spectra from orbitrap instrument.

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

The number of spectra we need to average depends on signal quality and sample. Our experiments showed that successful deconvolution is possible even with averaging of only a few spectra.

One of the problem in performing deconvolution is handling of coeluted charge states from different compounds. For complex mixtures it may lead to false assignments. Our approach allows to resolve coeluted charge states successfully. On Figure 4, a mixture of two proteins (reduced Sigma IgG) eluting at different times with coeluted charge states is shown (let's call protein of mass 23,595 as Small protein and protein of mass 52,323 as Big protein).

FIGURE 3. Deconvolved Spectrum of Sigma IgG (IT spectrum on an Orbitrap instrument). Deconvolved spectrum reveals the same modification pattern (both initial and deconvolved spectra are shown as raw spectra, without peak modeling or baseline removal).

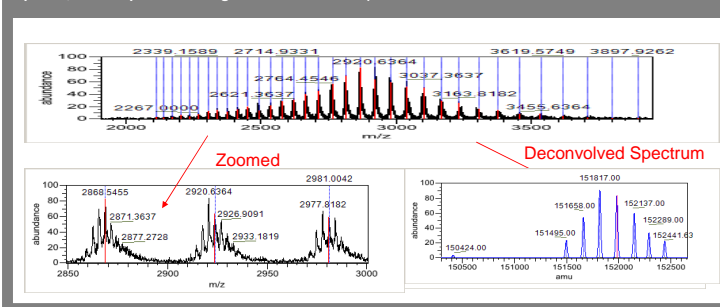
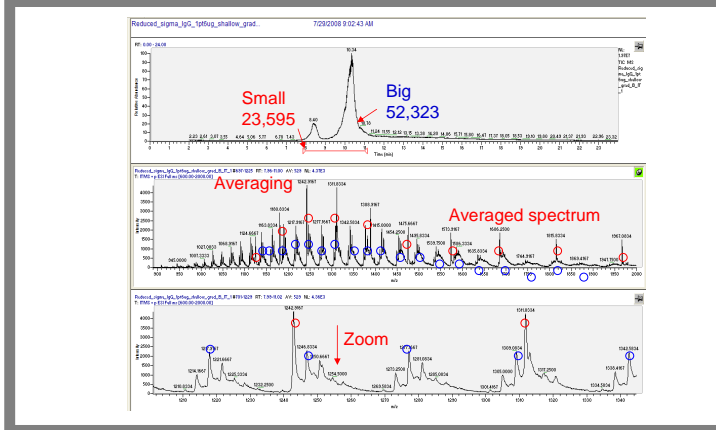


FIGURE 4. After averaging spectra over retention time resulting spectrum will contain coeluted charge states. Red and blue labels correspond to correct charge states assignment.



A spectrum for performing deconvolution is averaged so to include signal from both proteins (total 529 spectra – between 8 and 11 min.).

On Figure 5, the deconvolved mass from Small protein with correspondent charge states assignment is shown. There are 23 different charge states identified, and mass calculated based on identified charge states has standard deviation 9.35 amu.

On Figure 6 the deconvolved mass from Big protein with correspondent charge states assignment is shown. There are 41 different charge states identified, and mass calculated based on identified charge states has standard deviation 3.9 amu. For big protein additional modifications are also identified.

FIGURE 5. Deconvolved spectrum with correct charge state assignment for small protein (red labels)

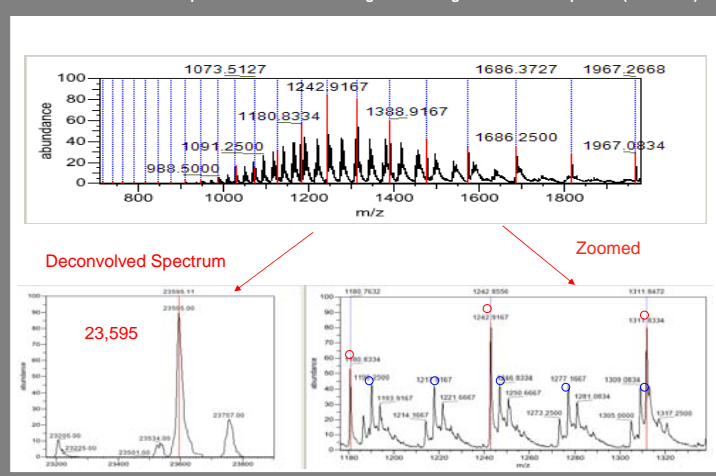
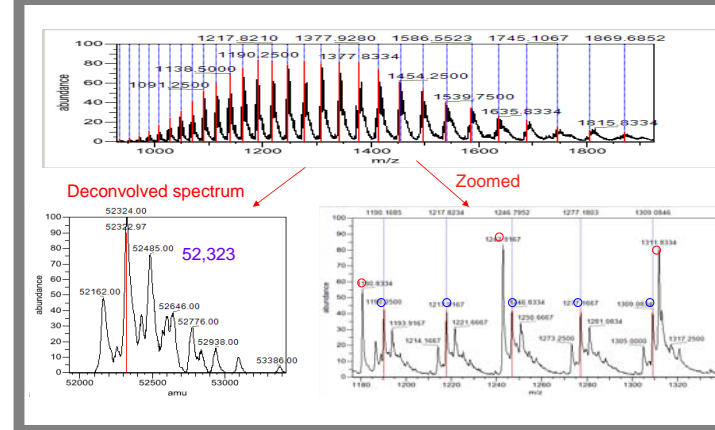


FIGURE 6. Big protein (mass 52,323) is deconvolved successfully. All charge states are restored (blue labels)



Conclusions

Identification and quantification of intact proteins and peptide (for which MS spectra are isotopically unresolved) using electrospray and high-resolution mass spectrometers 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 work well for high-resolution and low-resolution data (LTQ and LTQ Orbitrap instruments). Deconvolution of the algorithm resolves coeluted charge states successfully.

With intelligent automatic averaging, automatic processing of raw files can be done fast.

References

- Mann, M.; Meng, C. K.; Fenn, J. B., Interpreting Mass Spectra of Multiply Charged Ions, *Anal. Chem.* 1989, 61, 1702-1708.
- Michael W. Senko, Steven C. Beu, and Fred W. McLafferty, Automated Assignment of Charge States from Resolved Isotopic Peaks for Multiply Charged Ions, *J. Am. Soc. Mass Spectrom.* 1995, 6, 52-56