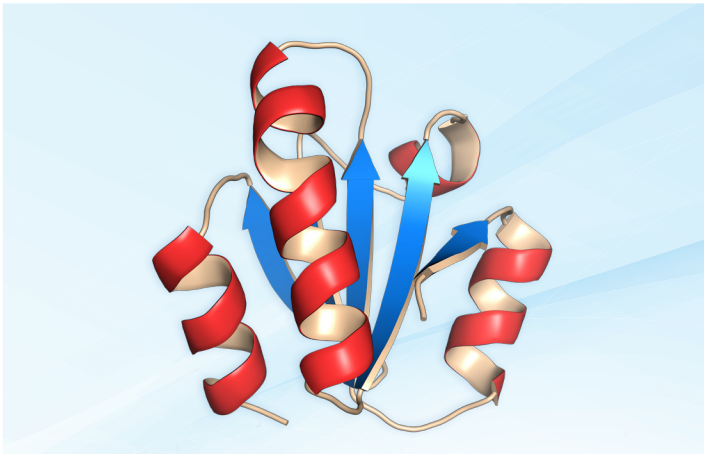
TECHNICAL NOTE 74150

Targeted small protein analysis with a single quadrupole mass detector

Authors: Dan Kalu Appulage[1], Jennifer N. Sutton[2], Martin Samonig[3], Stephan Meding[3]
[1]Thermo Fisher Scientific, Austin, TX, USA
[2]Thermo Fisher Scientific, San Jose, CA, USA
[3]Thermo Fisher Scientific, Germering, Germany

Keywords: Single quadrupole mass spectrometer, ISQ EM, Vanquish Flex Binary UHPLC, intact protein deconvolution, intact mass analysis, small protein



Goal

Demonstrate the benefits of single quadrupole mass detection for small protein analysis with the Thermo Scientific™ Chromeleon™ Chromatography Data System (CDS) and its Intact Protein Deconvolution package.

Introduction

Within regulated industries, the minimum requirements for quality control are continuously increasing. Whereas protein identity and purity confirmation were previously based on retention time and peak quality using ultraviolet (UV) detection, mass confirmation has increasingly become the preferred method. Although high-resolution accurate mass (HRAM) instruments, including Thermo Scientific™ Orbitrap™-based mass spectrometers (MS), provide the most detailed insights, certain analyses require only unit mass resolution. In such cases, single quadrupole mass detection in combination with UV detection offers a more cost-efficient solution.

This technical note demonstrates the capabilities of the Thermo Scientific™ ISQ™ EM single quadrupole mass detector for protein analysis. The ISQ EM is fully integrated into Chromeleon CDS, which allows straightforward LC-MS data analysis by chromatographers even with limited MS experience. Moreover, Chromeleon CDS allows GxP compliant operation, including data analysis with protein deconvolution, making it an ideal setup for regulated environments.

Experimental

A mixture of six proteins (Table 1) was used for the analysis. The sample was dissolved in 100 µL of LC-MS grade water, resulting in a final concentration of 0.76 µg/µL.

Table 1. Thermo Scientific™ Pierce™ Intact Protein Standard Mix

Protein name	Protein accession	Average molecular weight	Monoisotopic mass [M]
Human IGF-I LR3*	P05019 (40–118)	9,111.47	9,105.35
Human Thioredoxin	Q99757 (60–166)	11,865.52	11,858.04
Streptococcus dysgalactiae Protein G	P06654 (223–413)	21,442.61	21,429.76
Bovine Carbonic Anhydrase II*	P00921	28,981.29	28,963.69
Streptococcus Protein AG *(chimeric)*	P02976, P19909	50,459.74	50,429.85
Escherichia coli Exo Klenow	P00582 (324–928)	68,001.15	67,959.43

*Protein may undergo partial deamidation in acidic conditions.

Chromatographic separation was performed using LC-MS grade solvents (Table 2) on a Thermo Scientific™ Vanquish™ Flex Binary UHPLC system (Table 3). For detection of the proteins, both a Vanquish Flex variable wavelength detector and an ISQ EM single quadrupole mass detector were used.

Chromeleon CDS version 7.3 was used for system operation and subsequent data analysis including intact protein deconvolution.

Table 2. Solvents and additives

Reagent	Grade	Supplier	Part Number
Pierce™ Intact Protein Mix	N/A	Thermo Scientific™	A33526
Acetonitrile	Optima™ LC-MS	Fisher Chemical™	A955-212
Formic acid	Optima™ LC-MS	Fisher Chemical™	A117-50
Water, purified using Barnstead™ GenPure™ xCAD Plus Ultrapure Water Purification System	Ultra-Pure, 18.2 MΩ at 25 °C	Thermo Scientific™	N/A

Table 3. Vanquish Flex Binary UHPLC system modules

Module	Part number
Vanquish System Base Horizon / Flex	VF-S01-A-02
Vanquish Binary Pump F (with 35 µL mixer set)	VF-P10-A-01 (6044.3870)
Vanquish Split Sampler FT	VF-A10-A-02
Vanquish Column Compartment H	VH-C10-A-03
Vanquish Variable Wavelength Detector F (2.5 µL SST flow cell)	VF-D40-A (6077.0360)

Table 4. Thermo Scientific™ column used for chromatographic separation

Stationary phase	Particle size	Dimension	Part number
MAbPac™-RP	4 µm	2.1 × 50 mm	088648

LC-MS conditions for the analysis

Parameter	Value
Mobile phase	A—Water with 0.1% formic acid B—Acetonitrile with 0.1% formic acid
Gradient	0.0 min: 10% B 0.0–25.0 min: 10–55% B 25.0–26.0 min: 55–95% B 26.0–27.0 min: 95% B 27.1–31.0 min: 10% B
Flow rate	0.6 mL/min
Column temperature	75 °C, forced air mode, passive pre-heater
Injection volume	1.0 µL
UV detection	214 nm, 100 Hz
Source settings	Easy mode with default settings (3) except for sweep gas pressure, which is set to most sensitivity (1)
Full Scan	
Mode	Heated Electrospray Ionization (HESI)
Time	0–31 min
Mass range	500–2000 *m/z*
Dwell time	1.0 s
Polarity	Positive
Source CID	10 V

Table 5. Intact Protein Deconvolution settings

Parameter	Value or setting
Peak retention window	# (see below)
Algorithm	ReSpect™
Output mass range	8,000–70,000 Da
Deconvoluted spectra display mode	Isotopic profile
Model mass range	8,000–70,000
Peak model	Intact protein
Resolution	Raw file specific
Deconvolution mass	100 ppm
Charge carrier	H^+
Charge high	100
Charge low	5
High number adjacent charges	Varies by the protein (6 or 4)
Low number adjacent charges	4
Intensity threshold scale	0.01
Min peak significance	1
Negative charge	False
Noise compensation	True
Noise rejection	95
Number of peak models	1
Peak model with scale	1
Quality score threshold	0
Relative abundance threshold	0
Target peak mass	* (see below)
Target peak shape left	2
Target peak shape right	2

The peak retention window depends on the source spectra parameter selection. If the sliding window method is selected, it will be the entire retention time window where peaks are eluted. If the average over selected retention time is selected, then it will be the retention time window for a particular peak.

*The target peak mass varies depending on the target protein theoretical mass. For example, if the target protein is human thioredoxin, a similar target peak mass should be entered, e.g., 12,000 Da.

Results and discussion

For determining ISQ EM suitability for mass confirmation of proteins, a protein standard comprising six proteins with molecular masses ranging from approximately 9 to 70 kDa was chosen.

The protein mixture was analyzed with a LC-UV-MS setup. Protein absorbance was measured using a single wavelength UV detector operated at 214 nm, while mass detection was performed using the ISQ EM in full scan mode with mass range 500–2000 *m/z*. Automatic source settings were used without further optimization since they resulted in good quality spectra (Table 5). The resulting chromatograms are shown in the Figure 1. The proteins were clearly detected by UV and MS, and peak resolution in MS was sufficient for further spectral analysis.

Initial data processing, including defining the peak integration window, the baseline noise area, and UV peak integration, was done in the Data Processing tab of Chromeleon CDS. Protein deconvolution was performed in the Intact Protein Deconvolution (IPD) tab, a component of Thermo Scientific™ BioPharma Finder™ software, which is fully integrated into Chromeleon CDS. IPD provides complete data processing within a single software suite, and therefore, can be set up in a GxP compliant way. Within IPD there are two different ways to process unit mass resolution data. Either a peak window can be defined when there is only a single protein of interest or, for multiple proteins, a sliding window with a defined width is recommended. IPD has multiple parameters that need to be chosen appropriately in order to obtain good quality data. Parameters selected for this analysis are shown in Table 5. Further information on the parameters and which values should be used can be found in the user guide of Biopharma Finder 4.0 software.[1]

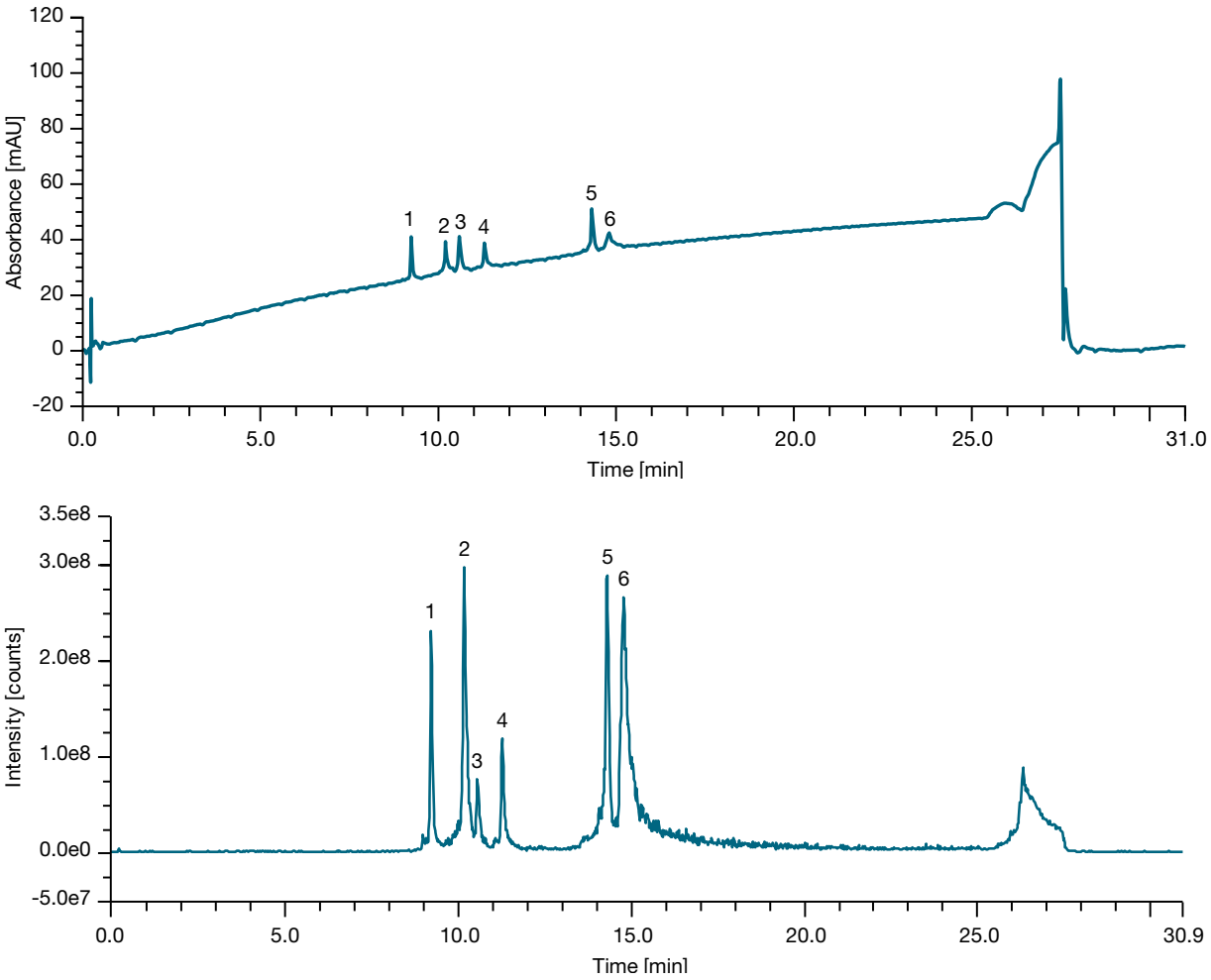


Figure 1. UV trace of the separation of the protein mix (top); mass detector total ion count trace for the protein mix (bottom).
1—*Streptococcus dysgalactiae* Protein G, 2—*Streptococcus* Protein AG *(chimeric)*, 3—Human IGF-I LR3, 4—Human Thioredoxin, 5—Bovine Carbonic Anhydrase II, 6—*Escherichia coli* Exo Klenow

Each of the six proteins exhibited average mass spectra showing multiple charge states. The smaller proteins showed fewer *m/z* peaks. The total number of observed peaks scaled with increasing molecular mass (Figure 2). Within each spectrum, *m/z* peaks represent different charge states of each protein. The higher the charge state, the smaller the *m/z* value. The respective charge state is calculated using Equation 1, with n being the number of protons (z).

Equation 1

$$m/z = \frac{[M+nH]^{n+}}{n}$$

Peaks are selected during deconvolution and the theoretical mass of the protein is inferred based on the *m/z* peak pattern (Figure 3). Further details on how protein deconvolution with the ReSpect algorithm is performed can be found in the BioPharma Finder software user guide.[1]

The theoretical masses and determined masses are shown in Table 6. The observed mass deviation increases with molecular weight. For the proteins around 10 kDa, it is 1–1.5 Da and increases above 20 Da for the largest protein, nearly 70 kDa in size. Observed values and the increase in mass deviation are in line with the expected results. The mass deviation is a result of the mass accuracy of the ISQ EM and the charge states of the *m/z* peaks used for deconvolution. The charge states used for deconvolution of the proteins are shown in Table 6. For example, charge states from 5 to 10 were used for deconvolution of the Human IGF-1 spectrum (Table 6). The ISQ EM has a mass accuracy of 0.1 *m/z*. For singly charged species, the mass accuracy equals the mass deviation. For proteins carrying multiple charges, the mass accuracy is the product of charge state and mass accuracy values. For Human IGF-1 a mass deviation between 0.5 Da (z = 5) and 1.0 Da (z = 10)

is expected, which matches the observed value (1.0 Da). For thioredoxin, the charge states 7 to 15 were used. Therefore, a mass deviation of less than 2.0 Da is expected (observed 1.5 Da). For all other proteins, the mass deviation was also within the expected range. Therefore, for small proteins up to ~15 kDa the molecular mass can be confirmed with a 1 to 2 Da mass deviation, while for larger proteins mass confirmation becomes less accurate.

Table 6. Theoretical masses and the observed mass using the intact protein deconvolution with mass deviation

Protein	Average molecular weight /Da	Observed molecular weight /Da	Mass deviation /Da	Charge states used for deconvolution
Human IGF-I LR3*	9,111.5	9,112.5	1.0	5–10
Human Thioredoxin	11,865.5	11,866.9	1.4	7–15
Streptococcus dysgalactiae Protein G	21,442.6	21,445.3	2.7	12–29
Bovine Carbonic Anhydrase II*	28,981.3	28,989.1	7.8	22–38
Streptococcus Protein AG *(chimeric)*	50,459.7	50,470.4	10.7	42–65
Escherichia coli Exo Klenow	68,001.6	68,022.5	20.9	63–86

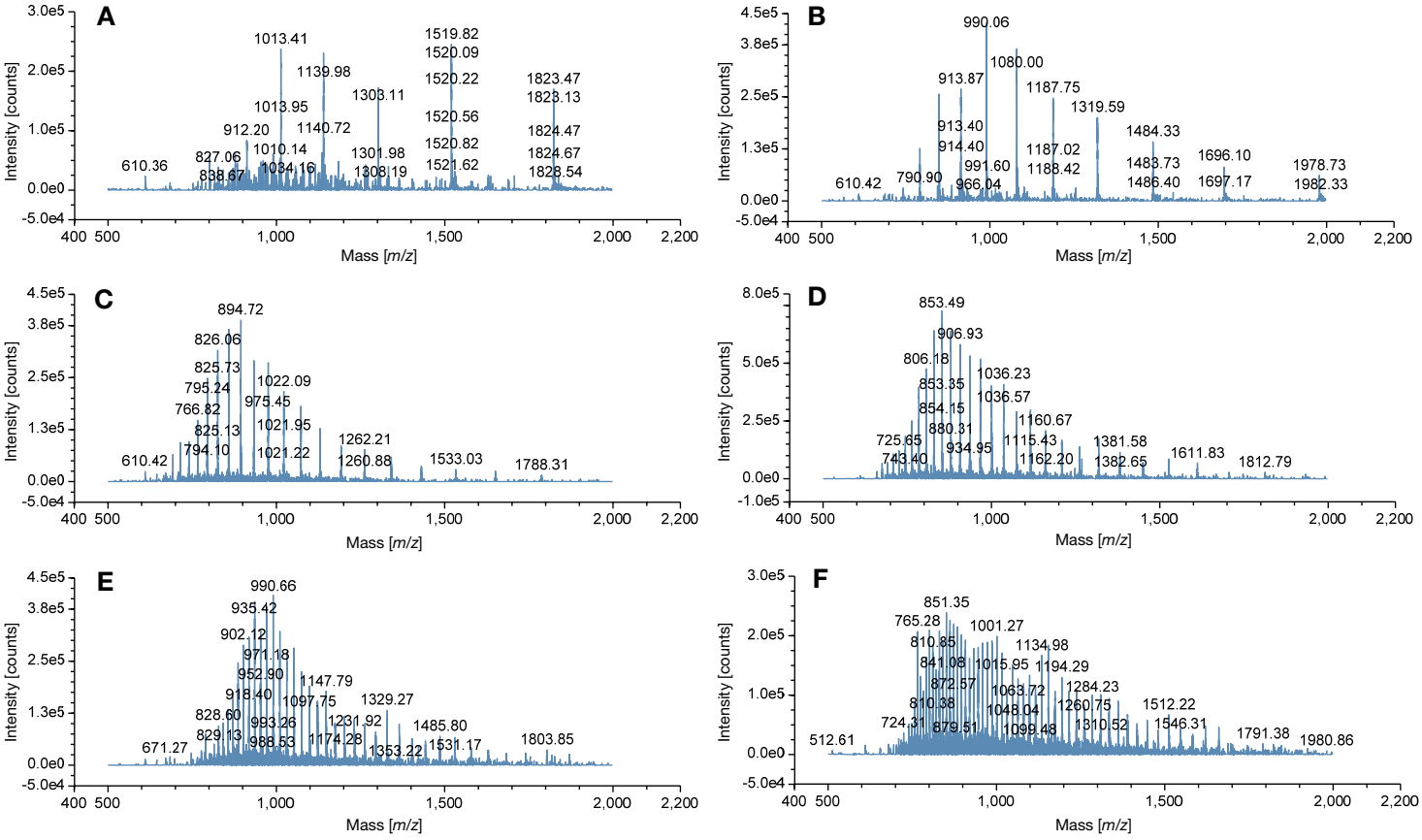


Figure 2. Average mass spectra of proteins. A. Human IGF-I LR3, B. Human Thioredoxin, C. *Streptococcus dysgalactiae* Protein G, D. Bovine Carbonic Anhydrase II, E. *Streptococcus* Protein AG, E. *Escherichia coli* Exo Klenow

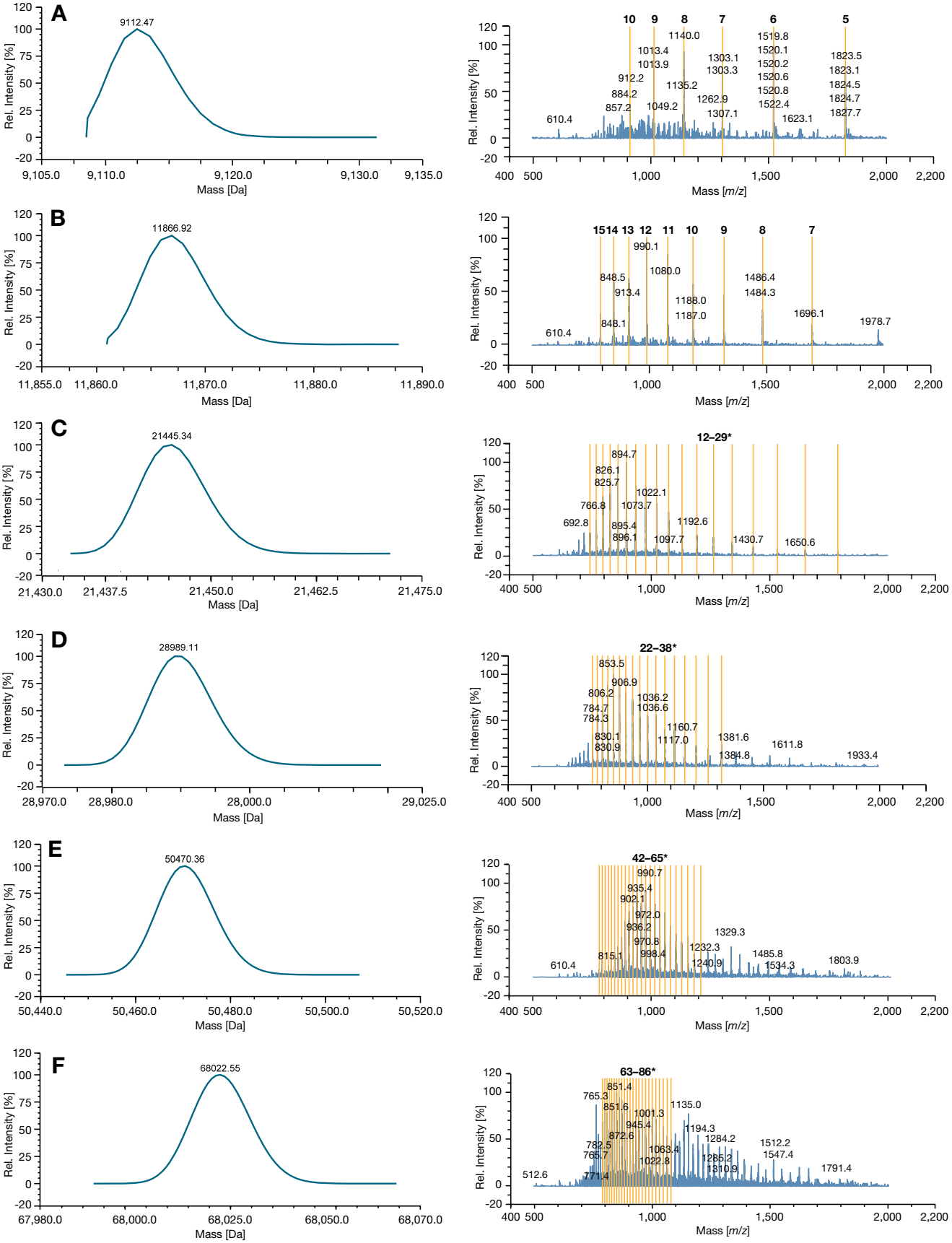


Figure 3. Deconvolution spectra of proteins. (Left—deconvoluted average spectra; right—source spectra used for deconvolution). A. Human IGF-I LR3, B. Human Thioredoxin, C. *Streptococcus dysgalactiae* Protein G, D. Bovine Carbonic Anhydrase II, E. *Streptococcus* Protein AG, E. *Escherichia coli* Exo Klenow.

*charge states used for deconvolution as indicated by orange lines in the mass spectra.

thermo scientific

Conclusion

Small protein analysis for mass confirmation was successfully performed using an ISQ EM and IPD in Chromeleon CDS, delivering mass deviations of less than 2 Da for proteins up to ~15 kDa. This level of accuracy is sufficient for mass confirmation workflows such as in quality control of peptide and protein syntheses and protein purification. Therefore, the ISQ EM and IPD in Chromeleon CDS is a cost-efficient mass spectrometer and software combination for intact mass determination in quality control workflows.

Reference

1. Thermo Scientific BioPharma Finder, User Guide—Software version 4.0. https://assets.thermofisher.com/TFS-Assets/CMD/manuals/man-xcali-98190-biopharma-finder-user-manxcali98190-en.pdf

Find out more at **thermofisher.com/isqem**

ThermoFisher SCIENTIFIC