thermoscientific

Development of an All-Recombinant Intact Protein Standard for LC & MS Application Development & System Suitability Testing

Rosa Viner[1], Helene L. Cardasis[1], Viktorija Vitkovske[2], Shane Bechler[1], Lukas Taujenis[2], Vikrant Gohil[2], Kay Opperman[1], John Rogers[1], Egle Capkauske[2], , Kestutis Bargaila[2], Agne Alminaite[2], Juozas Siurkus[2], Thermo Fisher Scientific, [1] USA, [2] Lithuania

INTRODUCTION

In recent years, interest in intact protein analysis by HPLC, LC-MS, and MS/MS has increased significantly. This can be attributed to both improvements to LC and MS hardware, instrument control software, and data processing software, as well as conceptual shifts in how we can best address and answer biological questions given these emerging commercially available capabilities. Having witnessed the explosive growth of bottom-up proteomics and the subsequent evolution of high-quality, widely accessible standards to normalize platform performance in time and space, and assist with method development for new applications, we recognize a similar need for the Top-down proteomics field. Here, we describe the development of a multi-purpose intact protein standard for LC, LC-MS, and LC-MS/MS quality control and application development.

MATERIALS AND METHODS

Sample Preparation

A number of protein candidates expected to satisfy the following criteria: 1) evenly covering a MW range of 10kD – 66kD, 2) presenting mostly clean, modification and adduct-free ESI spectra, and 3) having ESI charge state distributions covered a wide m/z range from 500-2,000, were identified and then expressed in *E. coli and B. subtills* and purified. All were screened for ionization efficiency, purity, MW, charge state distribution by LC-MS, and infusion-MS, resulting in selection of the final candidates for the recombinant standard. Mixing ratios were adjusted such that all proteins could be detected simultaneously in a single infusion MS experiment.

Quality and stability of the selected proteins were verified by SDS-PAGE, UV HPLC, infusion-MS and LC-MS, and MS/MS using a Thermo Scientific™ Orbitrap Fusion™ Lumos™ Tribrid™ or a Thermo Scientific™ Q Exactive™ hybrid quadrupole-Orbitrap™ mass spectrometer in intact protein mode. High resolution CID, EThcD and HCD spectra were collected to confirm the sequences assigned to the final protein list.

Data Analysis

Intact protein spectra were deconvoluted with ReSpect™ (for 15K resolution) or Xtract (for 120K resolution) using the sliding window deconvolution algorithm in Thermo Fisher Scientific™ Protein Deconvolution™ 4.0 software. The top down data were analyzed with Thermo Scientific™ ProSightPC™ 4.0 and Thermo Scientific™ Proteome Discoverer™ 2.1 (utilizing the ProSightPD™ node) software packages. All searches were performed against custom databases. Final results were filtered using an E-value cutoff of 1 x10^{-5} and search engine rank 1.

RESULTS

To extend utility of the protein standard mixture to the broader intact protein analysis community, the protein standard mixture should satisfy several additional requirements, such as demonstrating long term stability and the absence of animal origin proteins as per FDA regulations. As such, new non-animal, recombinant protein candidates were synthesized, affinity captured, and purified. Affinity tags were removed and proteins were dialyzed into MS-compatible buffer (Figure 1) Selected candidates (Figure 2) were fully characterized by SDS- PAGE, UV HPLC and LC/MS/MS, and accelerated stability tests were performed on those which met purity, MW, and m/z distribution requirements (Figure 3-7). The new pre-commercial mixture is presented here.

Figure1. Workflow for recombinant protein mixture generation

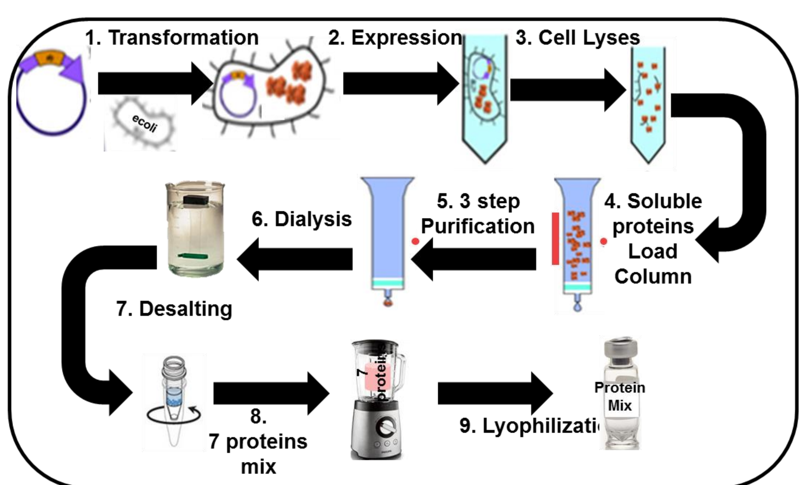


Figure 2. Prototype of recombinant protein mixture standard

No.	~Molecular weight (kDa)	Recombinant Proteins	Amount in a vial µg
1	9 kDa	IGF-I LR3 (Peprotech)	12
2	12 kDa	Thioredoxin	6
3	21 kDa	Protein G	12
4	29 kDa	Carbonic Anhydrase II	13
5	51 kDa	Protein AG	15
6	68 kDa	Exo Klenow	18
		Total protein amount	76 µg

Mix is lyophilized in 10 mM ammonium acetate MS - grade buffer

Figure 3. MS Analysis of selected candidates. Left panel shows the full scan spectrum of each candidate protein collected on a Thermo Scientific™ Q Exactive™ HF MS at a resolution of 15K (Klenow and Protein AG) or 120K at *m/z* 200. The center and right panels show HCD spectra of each protein collected on a Q Exactive HF along with ProSightPC 4.0 software results .The ProSight results from the largest protein, Klenow, were derived from a high resolution CID spectra from the Orbitrap Fusion Lumos MS.

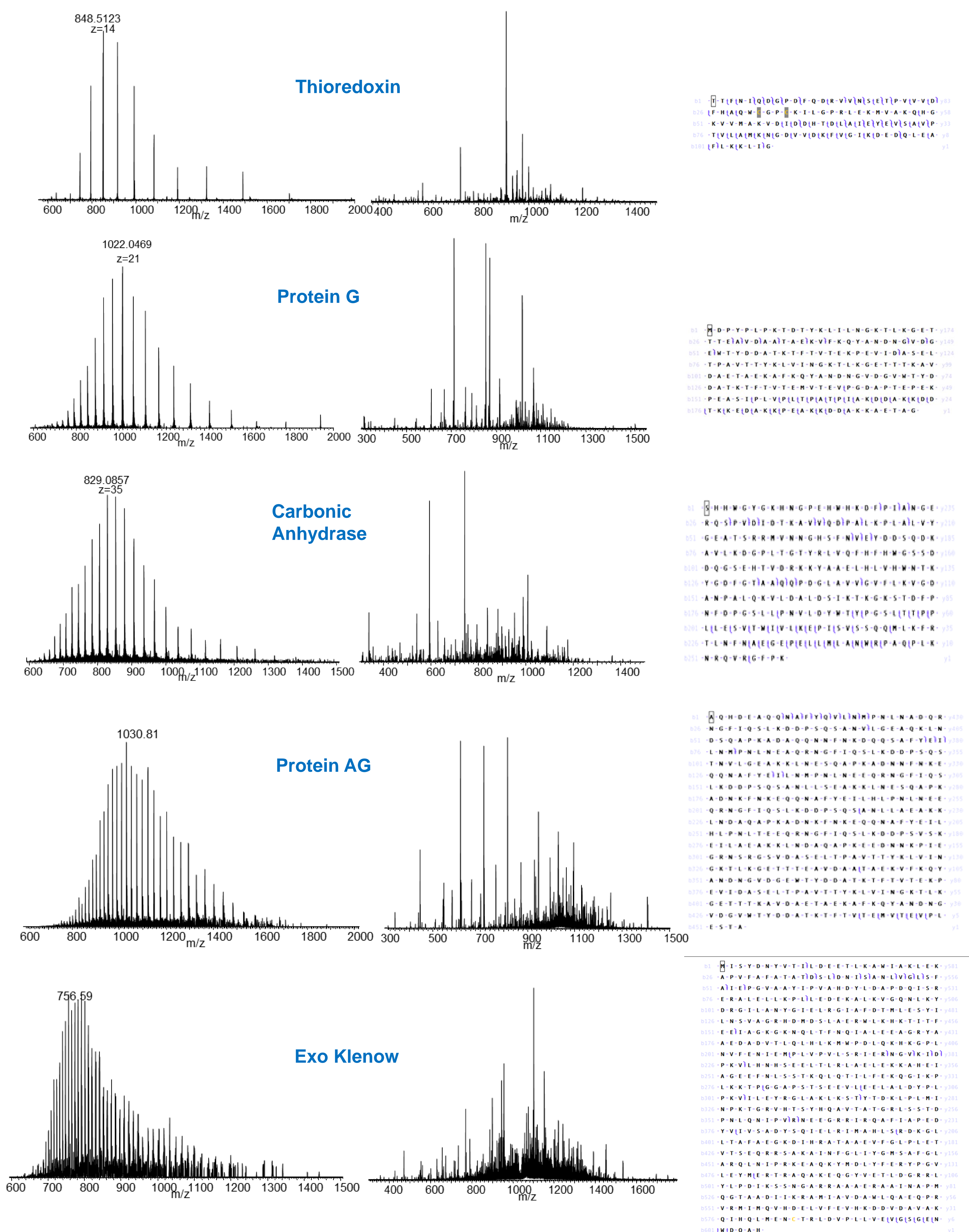


Figure 4. Final recombinant protein mixture infusion-MS spectra. Top panels – infusion spectral of new candidate mixture at 15k resolution and corresponding deconvolution results from Protein Deconvolution 4.0 software. All proteins typically detected, including larger proteins, Protein AG and Klenow. Bottom panels – infusion spectra of the optimized mixture using new recombinant protein candidates at 120k resolution and corresponding deconvolution results from Protein Deconvolution 4.0 software. IGF-ILR3, TRX, Protein G, and Carbonic Anhydrase are detected at this resolution setting.

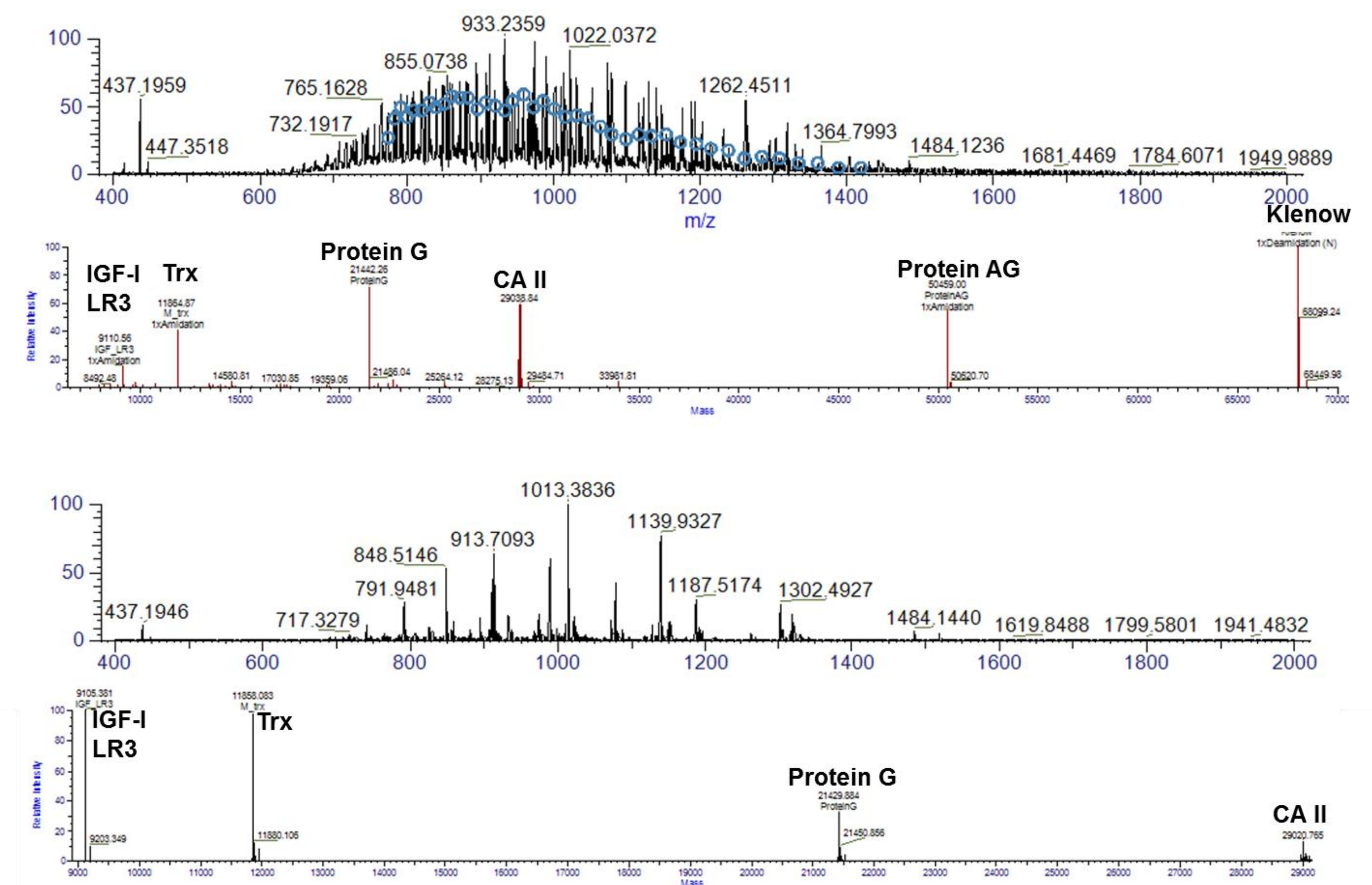


Figure 5. UV-HPLC analysis of recombinant proteins. Thermo Scientific™ Dionex™ UltiMate™ 3000 RSLCnano system; Column: prototype Thermo Scientific™ MAbPac™ RP 75µm x 25cm; Mobile Phase A: Water + 0.1% Formic acid or TFA, B: 80/20 (v/v) ACN/ Water + 0.1% Formic acid or TFA; Flow Rate: 2 µL/min; Temperature: 60° C; Sample: 100 ng protein mix (see chromatogram); Injection: 0.1 µL. Detection: UV 214 nm

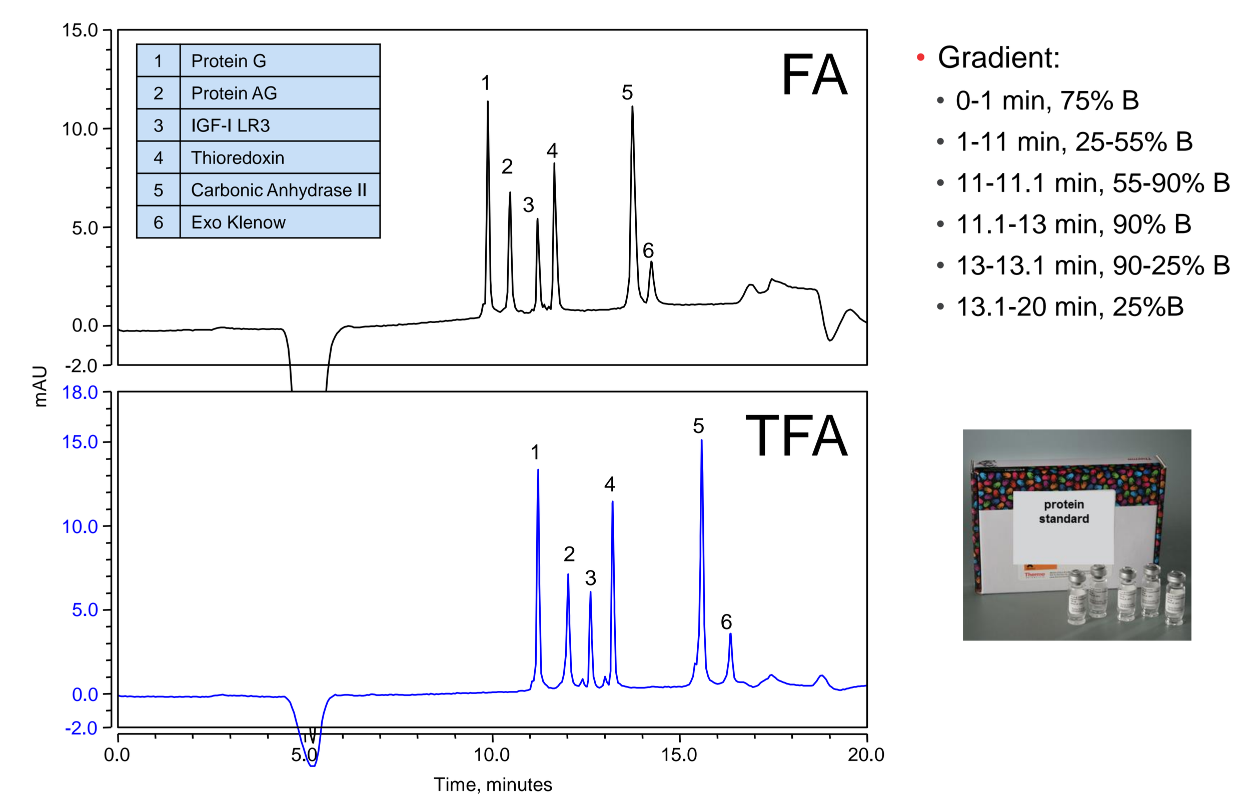
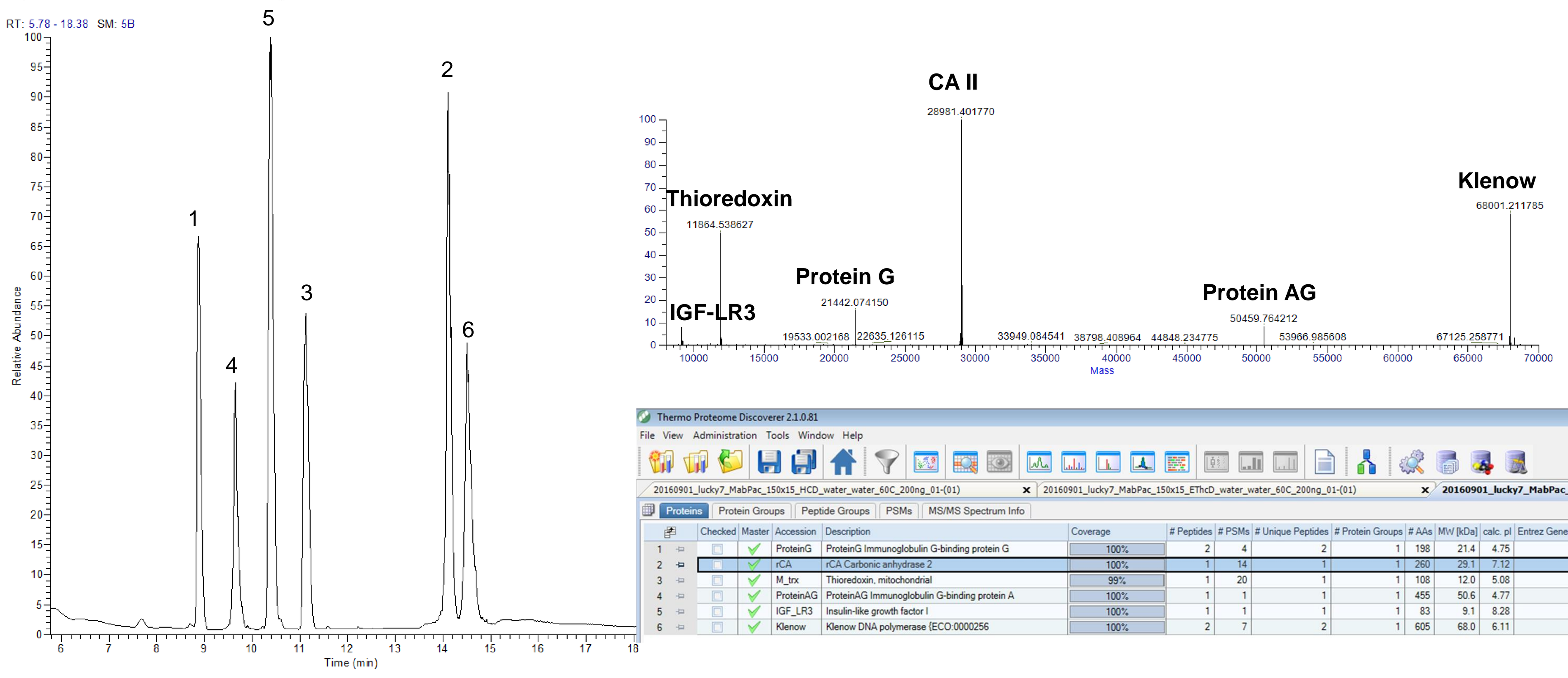


Figure 6. LC/MS/MS analysis of selected recombinant proteins. Left panel - Intact proteins were analyzed on an UltiMate 3000 RSLC system using a MAbPac RP 1 mm x 25 cm prototype column. Mobile Phase A: Water + 0.1% Formic acid, B: 100% ACN + 0.1% Formic acid; 20-55% B in 25 minutes Flow Rate:100 µL/min; Temperature: 60° C; Sample: 500 ng of protein mix (see chromatogram). Right panels- Samples were analyzed on a Q Exactive HF mass spectrometer in intact protein mode. MS/MS spectra were acquired using Top 3-5 DDA methods. OT MS[1] data was acquired at resolution settings of 15 at *m/z* 200 and OTMS[2] at a resolution of 120K at *m/z* 200. Top panel shows deconvolution results from Protein Deconvolution 4.0 software and Bottom panel shows top down results from ProSight PD 1.1 node in Proteome Discoverer 2.1 software.



Fresh NL: 4.44E6

7 days @ 37C NL: 4.51E6

Figure 7. Preliminary stability studies. Shown here is Protein G, reconstituted in water and analyzed immediately or after 7 days incubated at 37C, corresponding to 1 year storage at -20C.

CONCLUSIONS

A high quality intact protein standard was developed that meets needs for HPLC, MS, LC-MS/MS quality control, method development, and optimization. The completely recombinant nature of the sample makes it compatible for clinical research applications. The proteins cover a wide MW range, *m/z* range, and demonstrate good chromatographic separation.

ACKNOWLEDGEMENTS

We would like to acknowledge Jason Neil, Eugen Damoc, Tabi Arrey, Shanhua Lin for their help with early testing of, and feedback on, the R&D mixture. We would like to thank the Kelleher lab for testing, creating custom protein database and feedback on candidate proteins for the final mixture.

TRADEMARKS/LICENSING

© 2016 Thermo Fisher Scientific Inc. All rights reserved. ProSightPC and ProSightPD are trademarks of Proteinaceous, Inc. ReSpect is a trademark of Positive Probability, Ltd. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. This information is not intended to encourage use of these products in any manner that might infringe the intellectual property rights of others.

ThermoFisher SCIENTIFIC