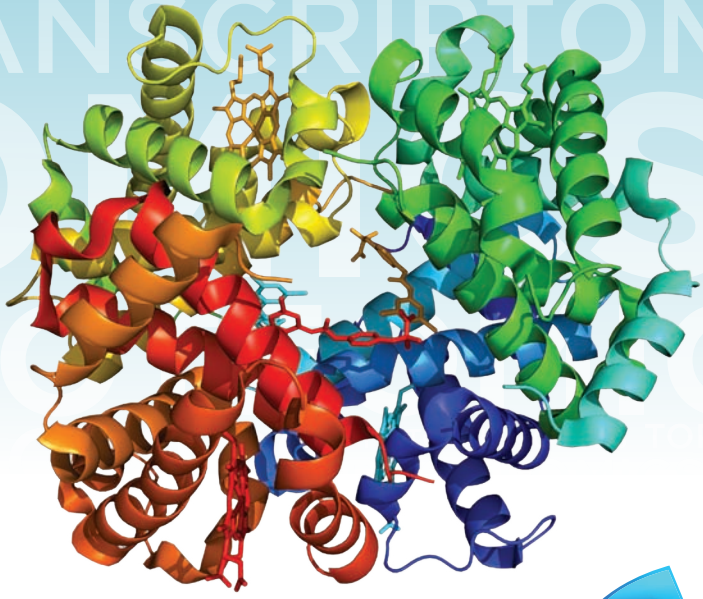
C&EN Supplement
SERVING THE CHEMICAL, LIFE SCIENCES & LABORATORY WORLDS



Trends in OMICS TECHNOLOGIES

December 2015

ACS Chemistry for Life®

American Chemical Society

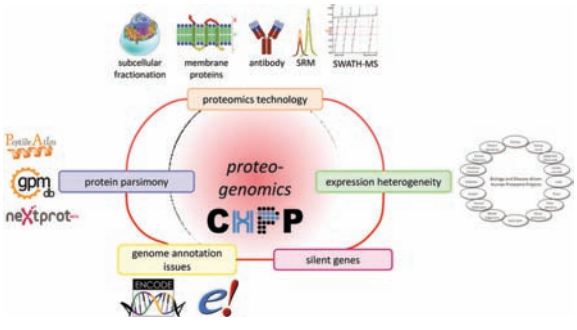
The ACS Scholars Program provides financial assistance, mentoring, and networking to talented students from groups that are underrepresented in the chemical sciences. Please donate now. You'll be unlocking the doors of opportunity for future chemists. **www.acs.org/dreams**

December 2015

A SUPPLEMENT TO C&EN
CHEMICAL & ENGINEERING NEWS

Trends in Omics Technologies

CONTRIBUTOR INDEX



p. 8

A SUPPLEMENT TO
C&EN

This C&EN Supplement is produced by C&EN, the flagship magazine of the American Chemical Society.

American Chemical Society
1155 16th Street, NW
Washington, DC 20036
(202-872-4600; 800-227-5558)

C&EN Supplements Management

Kevin A. Davies
Publisher, C&EN

Bibiana Campos Seijo
Editor-in-Chief, C&EN

Stephanie Holland
Manager, Advertising Sales and Marketing

Renee Zerby
Manager, Digital Production & Imaging

Joseph B. Davis, Jr.
Lead Digital Production Associate

Krystal E. King
Lead Digital Production Associate

Marielyn Cobero
Digital Production Associate

Cesar Sosa
Digital Production Associate

ACS Management

Thomas M. Connelly, Jr.
CEO & Executive Director

Brian D. Crawford
President, ACS Publications

To advertise in C&EN Media Group products, contact Stephanie Holland at (202) 872-4536 (s_holland@acs.org) or go to CENMediaKit.org

Cover: Schematic diagram of the flow of information from DNA to protein. Adapted from *J. Proteome Res*. Vol 13, Iss 1 (January 2014); Marielyn Cobero/C&EN

PUBLISHER'S NOTE

Omics Overload

ONE AFTERNOON DURING my PhD in London back in the mid '80s, the chair of the biochemistry department came into the lab excitedly waiving a new American journal in his hand. It had a bright red cover and a terrifically cool name: *GENOMICS*

The name was so new, we weren't even sure how to pronounce it—Gen-ohm-ics or Gen-omm-ics? (It's the first, of course, rhymes with gnome.)

In the years before and after the completion of the Human Genome Project, the "'Omics" suffix was not unreasonably extrapolated to a few other complementary fields—proteomics and transcriptomics, of course, followed by metabolomics, lipidomics, and glycomics.

But in recent years, scientists have gone to town, inventing so many "omics" terms it's almost impossible to keep count. UC Davis professor Jonathan Eisen penned a commentary in the journal *GigaScience* taking his colleagues to task for the profusion of "bad-omics" jargon, even prompting a story in the *Wall Street Journal*. There is even a random (but quite amusing) bad omics word generator: http://www.ark-genomics.org/badomics-generator

Clearly things have gone too far: The interactome? The culturome? Eisen has diligently flagged countless cringe-worthy examples: Uniqueome, circomics, sexomics, nutrimetabonomics, and phenomics, to name but a few. I just discovered another: Infonomics. Dan Shine, the president of Thermo's mass spectrometry and chromatography business, uses another omics term I hadn't heard before: Multi-omics. Perhaps we should let that one go, sweeping all before it.

Doubtless some scientists see the creation of new omics disciplines as a means to jazz up a manuscript or appeal to stringy funding organizations. And there is some anecdotal evidence to back this up: Eisen noted on his blog how a Penn State scientist's use of the term museomics had significantly increased interest and funding of his work on ancient DNA and comparative genomics.

But from a technology and instrumentation perspective, omics is big science, big data and big business. The methods and techniques at the center of the application notes in this issue, and in the top ten articles from the leading ACS proteomics journal, the *Journal of Proteome Research,* show how much has to be done. Understanding the complete picture of how proteins, RNA and other biomolecules interact in health and disease remains a daunting task. But to fully understand and to be able to treat infectious and complex diseases, there are no short-cuts.

Thanks to contributing editor Nina Notman and the organizations contributing application notes to this supplement. We look forward to a new series of C&EN supplements in 2016.

Kevin Davies PhD
Publisher, C&EN

(*Publishing trivia: the journal's publisher was one Dr Brian Crawford, now president of ACS Publications.)

For the record: The editorial content in this supplement was created without direct involvement of C&EN reporters or editors.

MAPPING THE HUMAN PROTEOME

Nina Notman

WELCOME TO the final C&EN supplement of 2015, offering a snapshot of the latest trends in omics technologies. As C&EN's Publisher Kevin Davies discusses in his accompanying letter, several dozen different "omics" terms have now been coined (for better or worse). For this introduction, I'm focusing on one of the most active research areas: Proteomics.

More than a decade after the human genome was mapped; a complete list of the proteins expressed by this genome is still not available. The Human Genome Project was widely touted as the world's largest collaborative biological project, and collaboration is also the name of the game in efforts to map the human proteome.

It has been estimated that humans have around 20,687 protein-coding genes[1], and in May 2014 it was announced that a team of over 70 researchers spanning 20 institutions in six different countries had identified 17,294 (approximately 84%) of these[2]. They used high-resolution mass spectrometry to carry out proteomic profiling—identifying and quantifying the protein components—of 30 normal human tissues and primary cells.

At the same time, a group of 22 researchers based in six different German institutions published mass spectrometry-based protein evidence for 18,097 genes[3]. They had analyzed 60 tissues, 13 body fluids and 147 cell lines, and combined the data with that already available in the literature.

While these two papers were heralded as evidence of great progress, there were also sounds of caution against the suggestion that a map of the human genome was near completion. "The first 90% of projects like these is the easy part, and the last 10% is the really hard and expensive part," John R. Yates III, a proteomics expert at Scripps Research Institute California who was not involved with either study, told Celia Arnold for C&EN when the studies were first published[4].

Another huge international consortium working on mapping the human proteome is the Chromosome-Centric Human Proteome Project (C-HPP)[5]. The primary goal of the C-HPP is to identify representative proteins encoded by every human gene and organize the data by the chromosomes the genes are located on. Another goal is to develop new analytical technologies and bioinformatics to enable the so-called "missing proteins"—proteins for which there was little or no experimental evidence—to be identified. For the past three years, the *Journal of Proteome Research* (an ACS journal) has published an annual special issue describing process thus far.

In a perspective article published in the September 2015 issue[6], which features in our selection of top ten omics abstracts (see p. 4), 54 of the consortium's researchers from 38 institutions in 16 different countries explained why mapping the human proteome is a more significant analytical challenge than the human genome. "Proteins cannot be amplified and are chemically much more heterogeneous than DNA and RNA," they write.

In their article, the researchers state that just three years ago the proportion of missing proteins in the human proteome was 33%. They attribute the recent progress to four main factors: 1) Development of specific sample preparations, 2) the use of advanced spectrometers, 3) the application of different type of assays, and 4) the analysis of unusual human sample types.

We hope you enjoy reading this supplement, and if leaves you wanting to learn more about proteomics the *Journal of Proteome Research* website is an ideal place to start.

References

1. http://www.genome.gov/encode/
2. Kim, M.-S., Pinto, S. M., Getnet, D., et al. A draft map of the human proteome. *Nature*, **2014**, 509, pp. 575–581 (DOI: 10.1038/nature13302)
3. Wilhelm, M., Schlegl, J., Hahne, H., et al. Mass-spectrometry-based draft of the human proteome. *Nature*, **2014**, 509, pp. 582–587 (DOI: 10.1038/nature13319)
4. http://cen.acs.org/articles/92/web/2014/05/Human-Proteome-Map-Assembled.html
5. http://www.c-hpp.org/
6. http://pubs.acs.org/toc/jprobs/14/9

Nina Notman PhD is a contributing editor for this C&EN Supplement. Nina is a freelance science writer and editor based in Salisbury, UK. ■

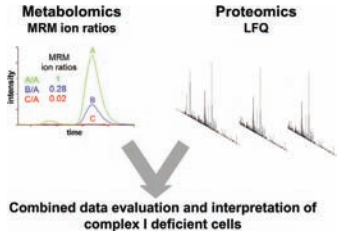
TOP 10 OMICS PAPERS

Ten of the most interesting papers from *Journal of Proteome Research* in 2015

PRESENTING TEN of *Journal of Proteome Research*'s most interesting and read Omics articles over the past 12 months. The abstracts of these key papers are reproduced below. To read these in full and to keep up with the latest developments in proteomics and related fields visit http://pubs.acs.org/journal/jprobs.

Metabolome and Proteome Profiling of Complex I Deficiency Induced by Rotenone

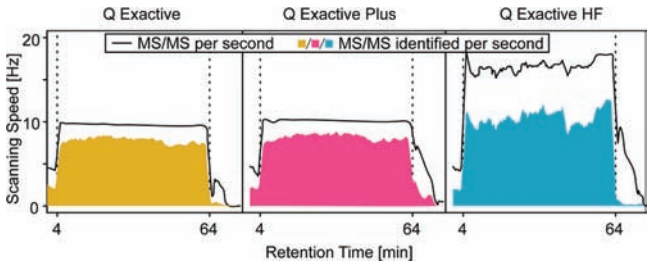
Ina Gielisch and David Meierhofer
Max Planck Institute for Molecular Genetics, Ihnestraße 63–73, 14195 Berlin, Germany
J. Proteome Res., **2015,** 14 (1), pp. 224–235
DOI: 10.1021/pr500894v



Complex I (CI; NADH dehydrogenase) deficiency causes mitochondrial diseases, including Leigh syndrome. A variety of clinical symptoms of CI deficiency are known, including neurodegeneration. Here, we report an integrative study combining liquid chromatography–mass spectrometry (LC–MS)-based metabolome and proteome profiling in CI deficient HeLa cells. We report a rapid LC–MS-based method for the relative quantification of targeted metabolome profiling with an additional layer of confidence by applying multiple reaction monitoring (MRM) ion ratios for further identity confirmation and robustness. The proteome was analyzed by label-free quantification (LFQ). More than 6000 protein groups were identified. Pathway and network analyses revealed that the respiratory chain was highly deregulated, with metabolites such as FMN, FAD, NAD^+, and ADP, direct players of the OXPHOS system, and metabolites of the TCA cycle decreased up to 100-fold. Synthesis of functional iron–sulfur clusters, which are of central importance for the electron transfer chain, and degradation products like bilirubin were also significantly reduced. Glutathione metabolism on the pathway level, as well as individual metabolite components such as NADPH, glutathione (GSH), and oxidized glutathione (GSSG), was downregulated. Overall, metabolome and proteome profiles in CI deficient cells correlated well, supporting our integrated approach. ■

Rapid and Deep Proteomes by Faster Sequencing on a Benchtop Quadrupole Ultra-High-Field Orbitrap Mass Spectrometer

Christian D. Kelstrup†, Rosa R. Jersie-Christensen†, Tanveer S. Batth†, Tabiwang N. Arrey‡, Andreas Kuehn‡, Markus Kellmann‡, and Jesper V. Olsen†
† Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3B, DK-2200 Copenhagen, Denmark
‡ Thermo Fisher Scientific (Bremen) GmbH, Hanna-Kunath-Strasse 11, 28199 Bremen, Germany
J. Proteome Res., **2014,** 13 (12), pp. 6187–6195
DOI: 10.1021/pr500985w

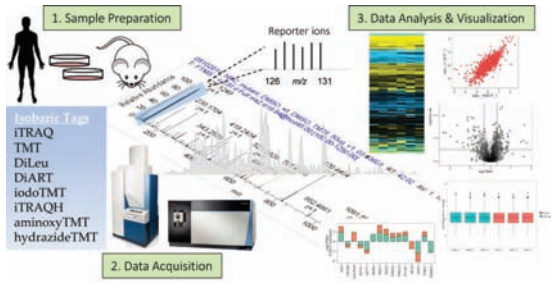


Shotgun proteomics is a powerful technology for global analysis of proteins and their post-translational modifications. Here, we investigate the faster sequencing speed of the latest Q Exactive HF mass spectrometer, which features an ultra-high-field Orbitrap mass analyzer. Proteome coverage is evaluated by four different acquisition methods and benchmarked across three generations of Q Exactive instruments (ProteomeXchange data set PXD001305). We find the ultra-high-field Orbitrap mass analyzer to be capable of attaining a sequencing speed above 20 Hz, and it routinely exceeds 10 peptide spectrum matches

per second or up to 600 new peptides sequenced per gradient minute. We identify 4400 proteins from 1 μg of HeLa digest using a 1 h gradient, which is an approximately 30% improvement compared to that with previous instrumentation. In addition, we show that very deep proteome coverage can be achieved in less than 24 h of analysis time by offline high-pH reversed-phase peptide fractionation, from which we identify more than 140 000 unique peptide sequences. This is comparable to state-of-the-art multiday, multienzyme efforts. Finally, the acquisition methods are evaluated for single-shot phosphoproteomics, where we identify 7600 unique HeLa phosphopeptides in one gradient hour and find the quality of fragmentation spectra to be more important than quantity for accurate site assignment. ■

Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics

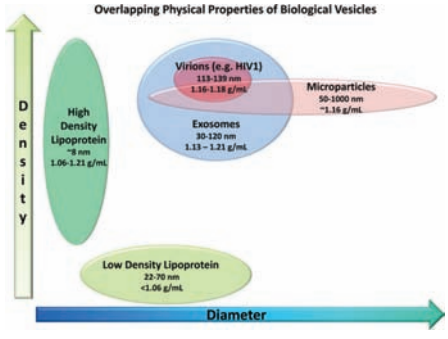
Navin Rauniyar and John R. YatesIII
Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States
J. Proteome Res., **2014,** 13 (12), pp. 5293–5309
DOI: 10.1021/pr500880b



Mass spectrometry plays a key role in relative quantitative comparisons of proteins in order to understand their functional role in biological systems upon perturbation. In this review, we review studies that examine different aspects of isobaric labeling-based relative quantification for shotgun proteomic analysis. In particular, we focus on different types of isobaric reagents and their reaction chemistry (e.g., amine-, carbonyl-, and sulfhydryl-reactive). Various factors, such as ratio compression, reporter ion dynamic range, and others, cause an underestimation of changes in relative abundance of proteins across samples, undermining the ability of the isobaric labeling approach to be truly quantitative. These factors that affect quantification and the suggested combinations of experimental design and optimal data acquisition methods to increase the precision and accuracy of the measurements will be discussed. Finally, the extended application of isobaric labeling-based approach in hyperplexing strategy, targeted quantification, and phosphopeptide analysis are also examined. ■

Mass-Spectrometry-Based Molecular Characterization of Extracellular Vesicles: Lipidomics and Proteomics

Simion Kreimer†‡, Arseniy M. Belov†‡, Ionita Ghiran‖, Shashi K. Murthy†§, David A. Frank⊥#, and Alexander R. Ivanov†‡
†Barnett Institute of Chemical and Biological Analysis, ‡Department of Chemistry and Chemical Biology, and §Department of Chemical Engineering, Northeastern University, Boston, Massachusetts 02115, United States
‖ Beth Israel Deaconess Medical Center, Boston, Massachusetts 02215, United States
⊥ Department of Medical Oncology, Dana−Farber Cancer Institute, Boston, Massachusetts 02115, United States
Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, United States
J. Proteome Res., **2015,** 14 (6), pp. 2367–2384
DOI: 10.1021/pr501279t

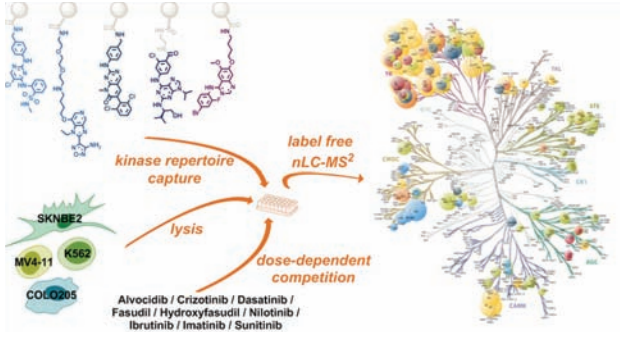


This review discusses extracellular vesicles (EVs), which are submicron-scale, anuclear, phospholipid bilayer membrane enclosed vesicles that contain lipids, metabolites, proteins, and RNA (microRNA and messenger). They are shed from many, if not all, cell types and are present in biological fluids and conditioned cell culture media. The term EV, as coined by the International Society of Extracellular Vesicles (ISEV), encompasses exosomes (30–100 nm in diameter), microparticles (100–1000 nm), apoptotic blebs, and other EV subsets. EVs have been implicated in cell–cell communication, coagulation, inflammation, immune response modulation, and disease progression. Multiple studies report that EV secretion from disease-affected cells contributes to disease progression, e.g., tumor niche formation and cancer metastasis. EVs are attractive sources of biomarkers due to their biological relevance and relatively noninvasive accessibility from a range of physiological fluids. This review is focused on the

molecular profiling of the protein and lipid constituents of EVs, with emphasis on mass-spectrometry-based "omic" analytical techniques. The challenges in the purification and molecular characterization of EVs, including contamination of isolates and limitations in sample quantities, are discussed along with possible solutions. Finally, the review discusses the limited but growing investigation of post-translational modifications of EV proteins and potential strategies for future in-depth molecular characterization of EVs. ■

Optimized Chemical Proteomics Assay for Kinase Inhibitor Profiling

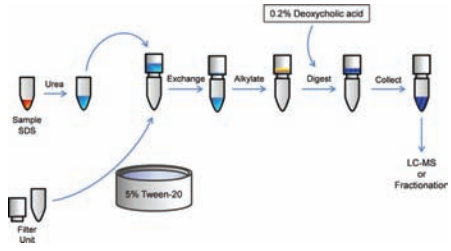
Guillaume Médard†, Fiona Pachl†, Benjamin Ruprecht†, Susan Klaeger†‡, Stephanie Heinzlmeir†‡, Dominic Helm†, Huichao Qiao†, Xin Ku†, Mathias Wilhelm†, Thomas Kuehne†, Zhixiang Wu†, Antje Dittmann§, Carsten Hopf§, Karl Kramer†, and Bernhard Kuster†‡‖
† Chair of Proteomics and Bioanalytics, Technische Universität München, Freising, Germany
‡ German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), Heidelberg, Germany
§ Cellzome GmbH, Heidelberg, Germany
‖ Center for Integrated Protein Science (CIPSM), Munich, Germany
J. Proteome Res., **2015,** 14 (3), pp. 1574–1586
DOI: 10.1021/pr5012608



Solid supported probes have proven to be an efficient tool for chemical proteomics. The kinobeads technology features kinase inhibitors covalently attached to Sepharose for affinity enrichment of kinomes from cell or tissue lysates. This technology, combined with quantitative mass spectrometry, is of particular interest for the profiling of kinase inhibitors. It often leads to the identification of new targets for medicinal chemistry campaigns where it allows a two-in-one binding and selectivity assay. The assay can also uncover resistance mechanisms and molecular sources of toxicity. Here we report on the optimization of the kinobead assay resulting in the combination of five chemical probes and four cell lines to cover half the human kinome in a single assay (~260 kinases). We show the utility and large-scale applicability of the new version of kinobeads by reprofiling the small molecule kinase inhibitors Alvocidib, Crizotinib, Dasatinib, Fasudil, Hydroxyfasudil, Nilotinib, Ibrutinib, Imatinib, and Sunitinib. ■

Enhanced FASP (eFASP) to Increase Proteome Coverage and Sample Recovery for Quantitative Proteomic Experiments

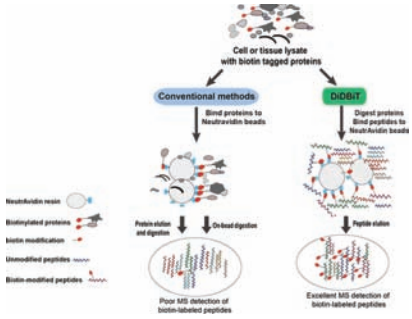
Jonathan Erde†, Rachel R. Ogorzalek Loo‡, and Joseph A. Loo†‡
†Department of Chemistry and Biochemistry and ‡Department of Biological Chemistry, University of California-Los Angeles, Los Angeles, California 90095, United States
J. Proteome Res., **2014,** 13 (4), pp. 1885–1895
DOI: 10.1021/pr4010019



The integrity of quantitative proteomic experiments depends on the reliability and the robustness of the protein extraction, solubilization, and digestion methods utilized. Combinations of detergents, chaotropes, and mechanical disruption can yield successful protein preparations; however, the methods subsequently required to eliminate these added contaminants, in addition to the salts, nucleic acids, and lipids already in the sample, can result in significant sample losses and incomplete contaminant removal. A recently introduced method for proteomic sample preparation, filter-aided sample preparation (FASP), cleverly circumvents many of the challenges associated with traditional protein purification methods but is associated with significant sample loss. Presented here is an enhanced FASP (eFASP) approach that incorporates alternative reagents to those of traditional FASP, improving sensitivity, recovery, and proteomic coverage for processed samples. The substitution of 0.2% deoxycholic acid for urea during eFASP digestion increases tryptic digestion efficiency for both cytosolic and membrane proteins yet obviates needed cleanup steps associated with use of the deoxycholate sodium salt. For classic FASP, prepassivating Microcon filter surfaces with 5% TWEEN-20 reduces peptide loss by 300%. An express eFASP method uses tris(2-carboxyethyl)phosphine and 4-vinylpyridine to alkylate proteins prior to deposition on the Microcon filter, increasing alkylation specificity and speeding processing. ■

Direct Detection of Biotinylated Proteins by Mass Spectrometry

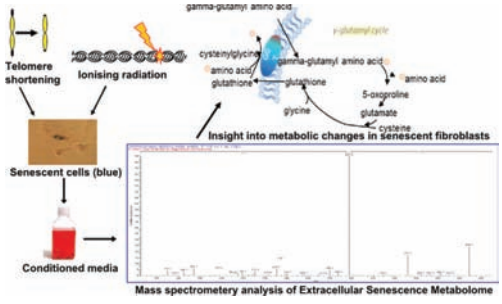
Lucio Matias Schiapparelli†, Daniel B. McClatchy‡, Han-Hsuan Liu†§, Pranav Sharma†, John R. YatesIII‡, and Hollis T. Cline†‡
†The Dorris Neuroscience Center, Department of Molecular and Cellular Neuroscience, ‡Department of Chemical Physiology, and §Kellogg School of Science and Technology, The Scripps Research Institute, La Jolla, California 92037, United States
J. Proteome Res., **2014,** 13 (9), pp. 3966–3978
DOI: 10.1021/pr5002862



Mass spectrometric strategies to identify protein subpopulations involved in specific biological functions rely on covalently tagging biotin to proteins using various chemical modification methods. The biotin tag is primarily used for enrichment of the targeted subpopulation for subsequent mass spectrometry (MS) analysis. A limitation of these strategies is that MS analysis does not easily discriminate unlabeled contaminants from the labeled protein subpopulation under study. To solve this problem, we developed a flexible method that only relies on direct MS detection of biotin-tagged proteins called "Direct Detection of Biotin-containing Tags" (DiDBiT). Compared with conventional targeted proteomic strategies, DiDBiT improves direct detection of biotinylated proteins ~200 fold. We show that DiDBiT is applicable to several protein labeling protocols in cell culture and in vivo using cell permeable NHS-biotin and incorporation of the noncanonical amino acid, azidohomoalanine (AHA), into newly synthesized proteins, followed by click chemistry tagging with biotin. We demonstrate that DiDBiT improves the direct detection of biotin-tagged newly synthesized peptides more than 20-fold compared to conventional methods. With the increased sensitivity afforded by DiDBiT, we demonstrate the MS detection of newly synthesized proteins labeled in vivo in the rodent nervous system with unprecedented temporal resolution as short as 3 h. ■

Senescent Human Fibroblasts Show Increased Glycolysis and Redox Homeostasis with Extracellular Metabolomes That Overlap with Those of Irreparable DNA Damage, Aging, and Disease

Emma L. James†, Ryan D. Michalek‡, Gayani N. Pitiyage†, Alice M. de Castro†, Katie S. Vignola‡, Janice Jones‡, Robert P. Mohney‡, Edward D. Karoly‡, Stephen S. Prime†, and Eric Kenneth Parkinson†
† Centre for Clinical & Diagnostic Oral Sciences, Institute of Dentistry, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, Turner Street, London E1 2AD, United Kingdom
‡ Metabolon, Inc., 617 Davis Drive, Suite 400, Durham, North Carolina 27713, United States
J. Proteome Res., **2015,** 14 (4), pp. 1854–1871
DOI: 10.1021/pr501221g

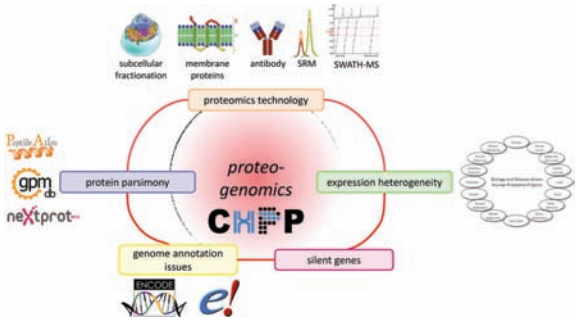


Cellular senescence can modulate various pathologies and is associated with irreparable DNA double-strand breaks (IrrDSBs). Extracellular senescence metabolomes (ESMs) were generated from fibroblasts rendered senescent by proliferative exhaustion (PEsen) or 20 Gy of γ rays (IrrDSBsen) and compared with those of young proliferating cells, confluent cells, quiescent cells, and cells exposed to repairable levels of DNA damage to identify novel noninvasive markers of senescent cells. ESMs of PEsen and IrrDSBsen overlapped and showed increased levels of citrate, molecules involved in oxidative stress, a sterol, monohydroxylipids, tryptophan metabolism, phospholipid, and nucleotide catabolism, as well as reduced levels of dipeptides containing branched chain amino acids. The ESM overlaps with the aging and disease body fluid metabolomes, supporting their utility in the noninvasive detection of human senescent cells in vivo and by implication the detection of a variety of human pathologies. Intracellular metabolites of senescent cells showed a relative increase in glycolysis, gluconeogenesis, the pentose-phosphate pathway, and, consistent with this, pyruvate dehydrogenase kinase transcripts. In contrast, tricarboxylic acid cycle enzyme transcript levels were unchanged and their metabolites were depleted. These results are surprising because glycolysis antagonizes senescence entry but are consistent with established senescent cells entering a state of low oxidative stress. ■

Quest for Missing Proteins: Update 2015 on Chromosome-Centric Human Proteome Project

Péter Horvatovich†, Emma K. Lundberg‡, Yu-Ju Chen§, et al.
† Analytical Biochemistry, Department of Pharmacy, University of Groningen, A. Deusinglaan 1, 9713 AV Groningen, The Netherlands
‡ Science for Life Laboratory, KTH - Royal Institute of Technology, SE-171 21 Stockholm, Sweden
§ Institute of Chemistry, Academia Sinica, 128 Academia Road Sec. 2, Taipei 115, Taiwan
J. Proteome Res., **2015,** 14 (9), pp. 3415–3431
DOI: 10.1021/pr5013009

This paper summarizes the recent activities of the Chromosome-Centric Human Proteome Project (C-HPP) consortium, which develops new technologies to identify yet-to-be annotated proteins (termed "missing proteins") in biological samples that lack sufficient experimental evidence at the protein level for confident protein identification. The C-HPP also aims to identify new protein forms that may be caused by genetic variability, post-translational modifications, and alternative splicing. Proteogenomic data

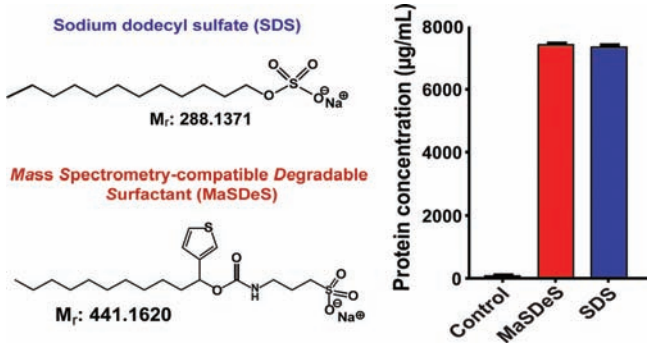


integration forms the basis of the C-HPP's activities; therefore, we have summarized some of the key approaches and their roles in the project. We present new analytical technologies that improve the chemical space and lower detection limits coupled to bioinformatics tools and some publicly available resources that can be used to improve data analysis or support the development of analytical assays. Most of this paper's content has been compiled from posters, slides, and discussions presented in the series of C-HPP workshops held during 2014. All data (posters, presentations) used are available at the C-HPP Wiki (http://c-hpp.webhosting.rug.nl/) and in the Supporting Information. ■

New Mass-Spectrometry-Compatible Degradable Surfactant for Tissue Proteomics

Ying-Hua Chang†, Zachery R. Gregorich†‡, Albert J. Chen†, Leekyoung Hwang§, Huseyin Guner‖, Deyang Yu†⊥, Jianyi Zhang#, and Ying Ge†‡§‖⊥
†Department of Cell and Regenerative Biology, ‡Molecular and Cellular Pharmacology Program, §Department of Chemistry, ‖Human Proteomics Program, and ⊥Molecular and Environmental Toxicology Program, University of Wisconsin-Madison, 1300 University Avenue, Madison 53706, Wisconsin, United States
Division of Cardiology, Department of Medicine, Department of Biomedical Engineering, Stem Cell Institute, University of Minnesota, 2001 6th Street SE, Minneapolis, Minnesota 55455, United States
J. Proteome Res., **2015,** 14 (3), pp. 1587–1599
DOI: 10.1021/pr5012679

Tissue proteomics is increasingly recognized for its role in biomarker discovery and disease mechanism investigation. However, protein solubility remains a significant challenge in mass spectrometry (MS)-based tissue proteomics. Conventional surfactants such as sodium dodecyl sulfate (SDS), the preferred surfactant for protein solubilization, are not compatible with MS. Herein, we have screened a library of surfactant-like compounds and discovered an MS-compatible degradable surfactant (MaSDeS) for tissue proteomics that solubilizes all categories of proteins with performance comparable to SDS. The use of MaSDeS in the



tissue extraction significantly improves the total number of protein identifications from commonly used tissues, including tissue from the heart, liver, and lung. Notably, MaSDeS significantly enriches membrane proteins, which are often under-represented in proteomics studies. The acid degradable nature of MaSDeS makes it amenable for high-throughput MS-based proteomics. In addition, the thermostability of MaSDeS allows for its use in experiments requiring high temperature to facilitate protein extraction and solubilization. Furthermore, we have shown that MaSDeS outperforms the other MS-compatible surfactants in terms of overall protein solubility and the total number of identified proteins in tissue proteomics. Thus, the use of MaSDeS will greatly advance tissue proteomics and realize its potential in basic biomedical and clinical research. MaSDeS could be utilized in a variety of proteomics studies as well as general biochemical and biological experiments that employ surfactants for protein solubilization. ■

NEW SPECIAL ISSUE

The Chromosome-Centric Human Proteome Project

GUEST EDITORS

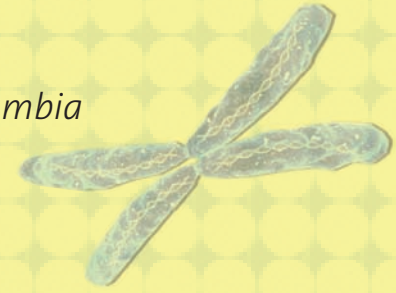
Young-Ki Paik
Yonsei University

Gilbert S. Omenn
University of Michigan

CO-EDITORS

Chris Overall
University of British Columbia

William S. Hancock
Northeastern University



tinyurl.com/JPRChromosomeIssue

Journal of proteome research

LIPID VISUALIZATION AND IDENTIFICATION THROUGH CCS AIDED CORRELATION OF MS IMAGING

Mark Towers And Emmanuelle Claude WATERS CORPORATION

Abstract

Collisional cross section (CCS) areas can be utilized to correlate mass spectrometry (MS) imaging data with ex situ MS data followed by MS-MS identification, for increased lipid identification.

Introduction

Using MS imaging, a broad range and number of lipid species can be visualized within a tissue section. Lipids can be identified by extracting them from the same, or a consecutive, tissue section that was used for MS imaging analysis and subsequently performing MS-MS experiments on those extracted lipids. However, when correlating MS imaging and ex situ MS data, definitive molecular identifications can be complicated due to the lack of certainty that the lipids extracted and identified, relate to the m/z peaks seen in the MS imaging data. This can be especially challenging when relying on accurate mass alone to identify an extracted lipid species. The CCS area of a peak of interest can be used as an additional qualifying factor to ensure that the peaks identified in the two data sets represent the same molecular species.

Experimental Methods

The ion mobility cell of the SYNAPT® G2-Si HDMS® Mass Spectrometer is calibrated via IntelliStart™. For the calibration, polyalanine was deposited on a MALDI target plate, mixed with the negative mode matrix 9-aminoacridine (9-AA). A thin section of mouse brain was produced using a cryotome and deposited on a non-conductive glass slide. The 9-AA matrix was applied evenly to the sample by spray coating. The MS imaging data was acquired at a spatial resolution of 45 μm in negative ion mode across a mass range of m/z 50-1,150, utilizing TriWave® ion guide optics. The MS imaging dataset was subsequently processed using High Definition Imaging (HDI®) Software. Lipid peaks present in the imaging data were compared to the Lipid Maps database (www.lipidmaps.org). In this case, 168 lipid candidates were identified based on mass accuracy better than +/-3 ppm. The CCS areas of the lipid candidates were calculated using the polyalanine calibration constants.

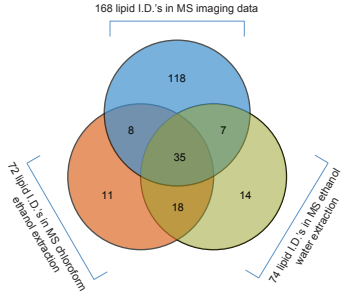


Figure 1

Graphical representation of the results of a database search on lipids identified by MALDI imaging and by MALDI MS of extracted lipids using two different extraction solvents, showing the number of peaks unique and common to both data sets.

Results and discussion

Lipids were extracted from two consecutive tissue sections by carefully depositing droplets of a solvent mixture onto the tissue section. Then, the extracted lipids were withdrawn by removing the extraction solvent. A 2:1 chloroform:ethanol solvent mixture was used for lipid extraction on one of the tissue section, while 4:1 ethanol:water was used on the second serial tissue section to be consistent with the solvent mixture used in the MS imaging experiment. The two lipid extracts were spotted separately onto a standard MALDI target plate using 9-AA as the matrix. After MS analysis, the peaks present were compared to the lipid maps database and the CCS areas calculated. The resulting lipid candidate lists were then compared to the list of 168 lipid candidates from the MALDI MS imaging dataset. The CCS areas of the peaks were cross validated (+/-0.5%) between the three datasets. The average CCS difference for the matches was found to be +/- 0.11%. 50 candidate lipid peaks were identified as being common to the MALDI imaging data set and one of the two lipid extract MS datasets (Figure 1). In total, 34 of 36 potential glycerophospholipid species were successfully identified. 18 fragment peaks matched the 21 possible peaks listed in the Lipid Maps database for phosphatidylinositol (18:0_20:4).

Summary

The use of CCS areas provided an additional level of confidence regarding the molecular identification. It supplemented and confirmed the results obtained from the accurate mass analysis, and lead to a more definitive identification of the lipid species of interest. ■

Additional information: http://www.waters.com/waters/library.htm?cid=511436&lid=134841560

ACS
Chemistry for Life®

American Chemical Society

STAY AHEAD OF THE GAME

Gain a competitive advantage with the **ACS Career Navigator™**. Visit ***www.acs.org/StayAhead*** to get started.

Brought to you by the ACS Career Navigator™ www.acs.org/CareerNavigator

NEW HORIZONS IN INTACT PROTEIN ANALYSIS: OPTIMIZATION OF TOP-DOWN PROTEIN ANALYSIS

Seema Sharma,[1] Parag Mallick,[2] Tanya Stoyanova,[3] Christopher Mullen,[1] Chad Weisbrod,[1] Jesse Canterbury,[1] David Horn,[1] Vlad Zabrouskov[1]

[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]Stanford University, Stanford, CA, USA; [3]University of California, Los Angeles, Los Angeles, CA, USA THERMO FISHER SCIENTIFIC

Overview

In this work we optimize top-down analysis of intact proteins with the goal of maximizing sequence coverage on a Thermo Scientific™ Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer[1]. Intact proteins were characterized using electron-transfer dissociation high dynamic range (ETD HD)[2], higher energy collisional dissociation (HCD) and collision-induced dissociation (CID), either by direct infusion into the mass spectrometer or online LC-MS. We demonstrated high sequence coverage for a range of intact proteins.

Introduction

Advancements in mass spectrometry have equipped researchers to explore new frontiers in biological science by enabling some of the most difficult analyses. Top-down proteomics involves identifying proteins in complex mixtures without prior digestion into their corresponding peptide species. Analysis allows the measurement of intact protein masses and provides information on post-translational modifications (PTMs), proteoforms as well as the protein sequence via fragmentation of the intact proteins in the mass spectrometer. In this work, we present top-down analysis of intact proteins on a Orbitrap Fusion Lumos Tribrid MS, with enhanced sensitivity facilitated by advanced vacuum technology for improved transmission of ions and an ETD HD fragmentation mode, providing significantly increased fragment ion coverage at faster acquisition rates. Results for intact protein fragmentation are presented for the standard proteins ubiquitin, carbonic anhydrase, enolase and IgG. We also present top-down analysis of the extra cellular domain of Trop-2, a trans-membrane glycoprotein highly expressed in epithelial cancers. Its function is regulated by intra-membrane cleavage. Aberrant cleavage may drive tumorigenesis as accumulation of the intracellular domain in the nucleus can drive proliferation, transformation and self-renewal.

Methods

Sample preparation and liquid chromatography

Protein standards (ubiquitin, carbonic anhydrase and enolase) were purchased from Sigma Aldrich. Proteins were infused at 1 pmol/μl at a flow rate of 5 μl/min for direct infusion experiments. Recombinant human Trop-2 protein (extracellular domain) was purchased from Sino Biological Inc. Intact IgG mass check standard and Trop-2 were reduced and denatured with TCEP at 60°C for 1 hour followed by PNGase incubation at 37°C for deglycosylation. Concentrations of proteins on column were 2 pmol for carbonic anhydrase, 4 pmol for enolase, 100 ng for IgG mass check standard and 1 μg for recombinant Trop-2.

A Thermo Scientific UltiMate™ 3000 RSLCnano system, operating in microflow mode was used for RP-LC analysis of 1 μl of protein loaded onto a Thermo Scientific ProSwift™ monolithic capillary column (200 um x 25 cm). Proteins were eluted with a gradient 25–65% acetonitrile in 0.1% formic acid over 6 min at a flow rate of 10 μl/min.

Mass spectrometry

Intact proteins were characterized using the higher capacity ETD HD fragmentation mode as well as HCD and CID under

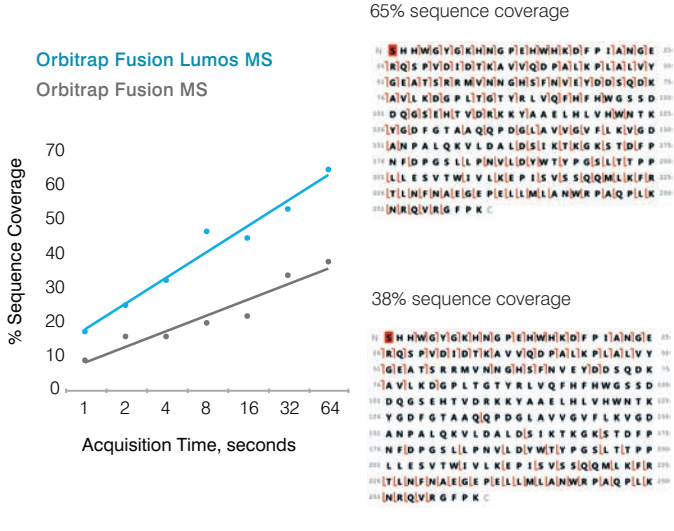


Figure 1

Sequence coverage for carbonic anhydrase as a function of acquisition time for Orbitrap Fusion MS vs. Orbitrap Fusion Lumos Tribrid MS. Spectra were acquired by isolating and fragmenting a single charge state at 240K in direct infusion mode.

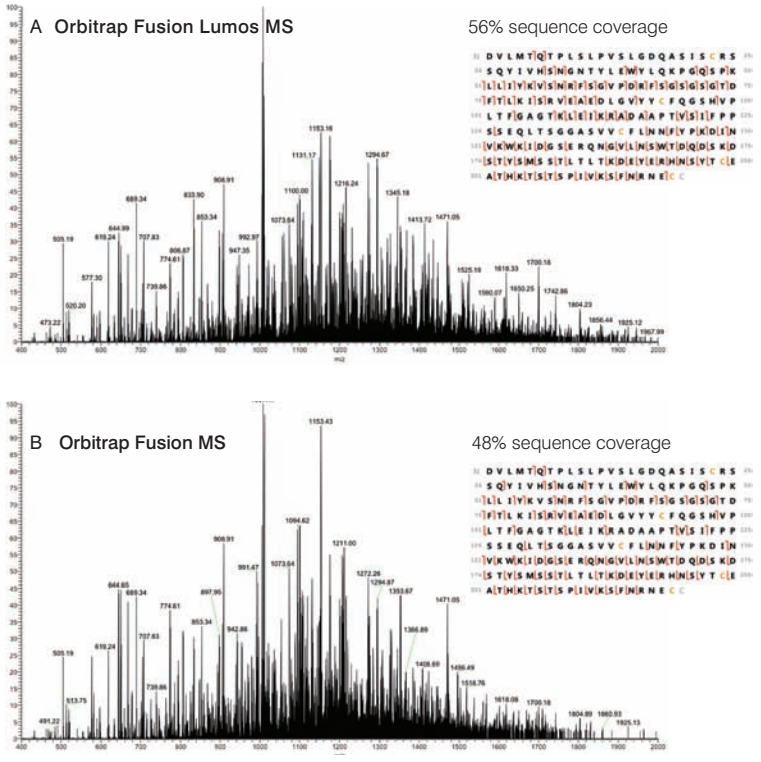


Figure 2

Comparison of the higher capacity ETD HD fragmentation of the IgG light chain on the Orbitrap Fusion Lumos Tribid MS (A) with ETD fragmentation on the Orbitrap Fusion MS (B). MS/MS spectra were acquired using direct infusion by isolating and fragmenting charge state 24+ at 240K and averaging 500 micro scans. ETD reaction times were optimized to allow maximum sequence coverage.

the improved vacuum conditions afforded by advanced vacuum technology[3] on the Orbitrap Fusion Lumos Tribrid MS. The instrument was operated in intact protein mode with an ion routing multipole pressure of 1–2 mtorr.

Fragmentation spectra were acquired at 240K for direct infusion experiments and 120K for online LC-MS experiments. Results from multiple experiments, isolating different charge states, ETD reaction times and different HCD and CID collision energy values, were combined in order to maximize the sequence coverage. Top-down data was processed using Thermo Scientific ProSightPC™ 3.0 sp1 software with 15 ppm mass tolerance for the fragment ion searches.

Results

Enhanced ETD performance

Improved precursor storage capacity in the ETD HD mode leads to a larger fragment ion population, thereby increasing the dynamic range of the ETD acquisition scan. The Advanced Vacuum Technology unique to the Orbitrap Fusion Lumos MS provides optimized conditions for improved performance with intact protein analysis. Advanced vacuum technology improves the transmission of ions, particularly those having a larger cross-section, into the Orbitrap analyzer by allowing a higher pressure in the ion routing multipole while maintaining an ultra-low vacuum in the Orbitrap analyzer. Both of these improvements result in significant increase in sequence coverage of intact proteins. Higher sequence coverage (50–100%) for carbonic anhydrase is obtained in a much shorter amount of time (4–8x improvement) on the Orbitrap Fusion Lumos Tribrid MS as compared to the Orbitrap Fusion MS (Figure 1). Increased sequence coverage obtained for IgG light chain (56%) using the higher capacity ETD HD as compared to ETD (48%) may also be obtained (Figure 2). The AGC (automatic gain control) target value was set to 1e6 for precursor ions in the ETD HD mode on an Orbitrap Fusion Lumos Tribrid MS and 3e5 for ETD on the Orbitrap Fusion MS. The reagent target was set at 7e5 in both cases.

Improved top-down sequence coverage

The Orbitrap Fusion Lumos Tribrid MS allows for increased gas pressure in the ion routing multipole while maintaining ultra-high vacuum in the Orbitrap region. This results in improved ion transfer and detection in the Orbitrap analyzer, especially when working with ions with large collisional cross-section, such as for large intact proteins. The combination of the various fragmentation techniques, the improved detection limits and dynamic range provided by ETD HD, and Advanced Vacuum Technology results in high sequence coverage for both IgG light and heavy chains using direct infusion (Figure 3). Different charge states were isolated in the mass selecting quadrupole with isolation window 5 *m/z* and fragmented using ETD HD with reaction times up to 15 ms, CID collision energy values 25–55%, and HCD collision energy values from 8–15%.

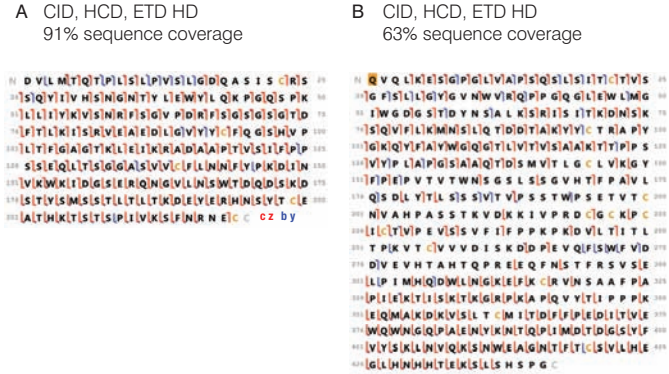


Figure 3

(A) 91% sequence coverage for the light chain and (B) 63% sequence coverage for the deglycosylated heavy chain was obtained from top-down analysis. Results from isolation of different charge states and ETD HD reaction times, HCD and CID collision energy values were combined in order to generate the sequence maps.

Top-down LC-MS analysis of intact proteins

Top down analysis of intact proteins was performed using online LC-MS chromatography as described in the methods section. The ion routing multipole pressure was set at 2 mtorr and targeted MS/MS spectra were acquired at 120K resolution. Precursor isolation widths were set at 5 *m/z* for isolating single charge states and 200 *m/z* for co-isolating and fragmenting multiple charge states of the protein. 68% sequence coverage for carbonic anhydrase and 41% sequence coverage for enolase was obtained in the LC-MS experiments by combining results from ETD HD, CID, and HCD (data not shown).

LC-MS based top-down analysis of Trop-2 protein

The experimental mass measurement for the deglycosylated Trop-2 recombinant protein suggests that the N-terminal truncation extends four amino acids further (1–30) than the predicted signal peptide sequence (1–26) for the extracellular domain. This is validated by sequence verification from the top-down fragmentation experiments (data not shown). Furthermore, the fragmentation results also suggest modification of the N-terminal glutamine to pyroglutamate (data not shown).

Conclusion

- The higher dynamic range ETD HD fragmentation mode provides increased sequence coverage at faster acquisition rates and time scales (increase from 38% to 65% for about 1 minute FT acquisition time demonstrated for carbonic anhydrase).
- Results from top-down analysis via multiple fragmentation modes for reduced IgG were combined to achieve 91% sequence coverage for the light chain and 63% sequence coverage for the heavy chain.
- Top-down LC-MS analysis of carbonic anhydrase and enolase via multiple fragmentation modes provided 68% and 41% sequence coverage respectively.
- Trop-2 sequence verification by top-down analysis shows truncation of the N-terminal (1–30) that extends further than the predicted signal peptide sequence (1–26) for the extracellular domain. The results also indicate pyroglutamte conversion at the N terminus and deamidation of three asparagine residues following the reaction with PNGase F.
- We demonstrate 18% sequence coverage for IgG light chain and 20% sequence coverage for the heavy chain based on LC-MS ETD HD analysis of intact IgG. ■

References

1. Thermo Scientific Orbitrap Fusion Lumos Tribrid MS, product specifications.
2. Mullen, C., et al., "Considerations for attaining improved ETD performance for top down applications," ASMS 2015 Poster 48, Tuesday 6/2.
3. Canterbury, J., et al., "Improvements for high-resolution analysis on a modified Tribrid mass spectrometer," ASMS 2015 Poster 124, Tuesday 6/2.

Thermo SCIENTIFIC

Thermo Fisher Scientific
355 River Oaks Parkway
San Jose, CA, 95134 USA
Website: www.thermoscientific.com

ACS
Chemistry for Life®

American Chemical Society

DARE TO BE DIFFERENT

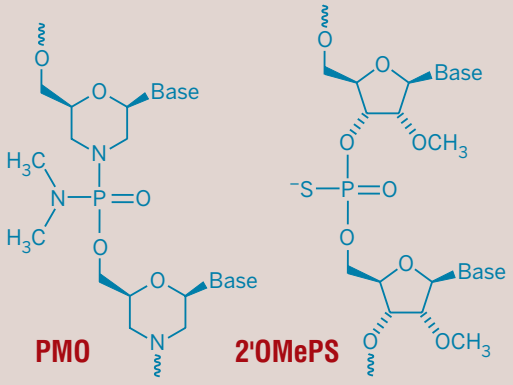
Learn how to break the mold with the *Fostering Innovation* course from the **ACS Leadership Development System®**. Visit ***acs.org/DareToBeDifferent*** to get started.

Brought to you by the ACS Career Navigator™ www.acs.org/CareerNavigator

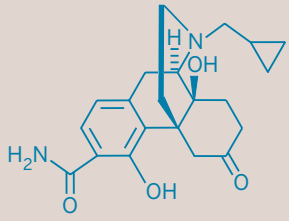
Correction

C&EN would like to issue the following corrections in the September 2015 "Top 20 Drugs in the Pipeline" supplement:

- In the article "Drisapersen, Eteplirsen, and Translarna" (p. 18–19), the labels of the sugar molecules on the diagram on p. 19 were transposed. The correct figure is here:



- The structure of Samidorphan (p. 22) was incorrect. The correct structure is shown here:



Samidorphan

- Finally, Teixobactin (p. 29) was discovered, developed, and is owned exclusively by NovoBiotic Pharmaceuticals LLC—not Northeastern University as stated. According to NovoBiotic president Dallas E. Hughes, Northeastern and the University of Bonn were collaborators on the project. Most of the work on teixobactin was funded through a Small Business Innovation Research grant from NIH to NovoBiotic and through private investment in the company. Drs Kim Lewis and Slava Epstein, both professors at Northeastern, are founders and paid consultants of NovoBiotic.

This new, monthly online journal is the FIRST journal to focus on the role of chemistry in infectious diseases and provides critical insights into this fast-growing, globally important research area.

pubs.acs.org/infectious

ACS Publications
Most Trusted. Most Cited. Most Read.

ACS PUBLICATIONS
CONGRATULATES THE 2015

NOBEL LAUREATES IN CHEMISTRY

TOMAS LINDAHL

PAUL MODRICH

AZIZ SANCAR

READ MORE ABOUT THE WINNERS AT

axial.acs.org