TWO Day Course

Top-Down Proteomics AND Mass Spectrometry Data Analysis and Visualization
This two-day course combines both the Top-Down Proteomics (Saturday one-day course) AND the Mass Spectrometry Data Analysis and Visualization (Sunday one-day course).

Top-Down Proteomics (Saturday one-day course)
Instructors

Ying Ge
University of Wisconsin, Madison

Joseph Loo
University of California, Los Angeles

Neil Kelleher
Northwestern University

Ljiljana Pasa-Tolic
Pacific Northwest National Laboratory

Top-down Mass Spectrometry Data Analysis and Visualization (Sunday one-day course)
Instructors

Xiaowen (Kevin) Liu
Tulane University

Mowei Zhou
Pacific Northwest National Laboratory

Sean McIlvain
University of Wisconsin Madison
ASMS Short Course
Top-Down Proteomics

Instructors: Ying Ge, Neil Kelleher, Joseph Loo, Ljiljana Paša-Tolić

“Front-end”
Separation of Intact Proteins

“Middle-part”
Mass Spectrometry Analysis of intact Proteins

“Back-end”
Informatics for Identification and Characterization of Intact proteins
Topics to Be Covered

• Introduction
  History of top-down proteomics
  Concept of proteoforms
  Measurement of Proteoforms and their complexes
  Denatured vs. native mode

• Sample preparation

• Intact protein separation

• Instrumentation, activation and dissociation methods for top-down MS

• Comparison of bottom-up and top-down – pros and cons

• Data interpretation and software tools for top-down proteomics

• Top-down quantitative proteomics – including experimental design

• Biomedical and biopharmaceutical applications of top-down MS

• Future outlook
The bottom-up approach is (therefore) suboptimal for determining modifications and alternative splice variants.

In the top-down approach...if a sufficient number of informative fragment ions are observed, this analysis can provide a complete description of the primary structure of the protein and reveal all of its modifications, as well as any correlations that exist between these modifications.
Concept of Proteoforms

Proteoform: a single term describing protein complexity

Sample Preparation in Top-down Proteomics

Best Practices and Benchmarks for Intact protein Analysis for Top-down Mass Spectrometry


Native Top-down Proteomics


Li et al. Nature Chemistry, 2018, 10, 139

Integrated Native Mass Spectrometry and Top-Down Proteomics
-Connect Sequence to Structure and Function of Macromolecular Complexes
Software Tools Available for Top-down Proteomics

ProSight PTM  https://prosightptm2.northwestern.edu/

ProSightPC™ https://www.thermofisher.com/order/catalog/product/PROSIGHTPC10

Mash Suite Pro:  http://crb.wisc.edu/yinglab/software.html

MASH Explorer:  http://ge.crb.wisc.edu/MASH_Explorer/index.htm

MS-Align+:  http://bix.ucsd.edu/projects/msalign/

TopPIC:  http://proteomics.informatics.iupui.edu/software/toppic

MSPathFinder:  https://omics.pnl.gov/software/mspathfinder


Proteoform Suite:  https://github.com/smith-chem-wisc/ProteoformSuite

pTop:  http://pfind.net/software/pTop/index.html

http://www.topdownproteomics.org/resources/software/
Biomedical and Biopharmaceutical Applications

Discover, Characterize & Quantify Proteoforms In Health and Disease


Conclusions & Outlook

- Measuring proteoforms directly: a major step in the evolution of mass spectrometry-based proteomics
- Top-down proteomics closes knowledge gaps by providing complete molecular specificity for proteins in wellness and disease
- Proteoform-resolved biology will increase efficiency of basic and translational research

It takes a village!
ASMS Short Course

Top-down Mass Spectrometry Data Analysis and Visualization

Xiaowen (Kevin) Liu, Mowei Zhou, and Sean McIlwain
Protein sequences

Top-Down MS Data Analysis workflow

Top-down spectra
Deconvolution
Mass list
Quantitation

Protein sequences
PTMs

Database search
Proteoform identification
Visualization & confirmation
Proteoform abundance
Top-down MS software

Spectral Deconvolution
- **Thrash** [Horn et al., JASMS 2000]
- **Thrash/Xtract** [Horn et al. JASMS 2000, Zabrousakov et al., JASMS 2005]
- **RAPID** [Park et al., Anal. Chem. 2008]
- **Decon2LS** [Jaitly et al., BMC Bioinformatics, 2009]
- **Hardklör** [Hoopmann et al., Anal. Chem. 2007]
- **MS-Deconv** [Liu et al. MCP, 2010]
- **MS-Deconv+/TopFD** [Kou et al., BMC Bioinformatics 2014]
- **UniDec** [Marty et al, AC, 2015]
- **pParseTD** [Sun et al, AC, 2016]
- **ProMex** [Park et al., Nature Methods 2017]
- **Intact** [ProteinMetrics, 2018]
- **ProteinDeconvolution** [Thermo]
- **FLASHDeconv** [Jeong et al., Cell Systems 2020]

Database Search
- **ProSightPC** [Zamdborg et al., Nucleic Acids Res., 2007]
- **PIITA** [Tsai et al., JASMS, 2009]
- **USTag** [Shen et al., Anal. Chem., 2008]
- **MS-TopDown** [Frank et al., Anal. Chem., 2008]
- **MS-Align+** [Liu et al., MCP 2011]
- **MS-Align-E** [Liu et al., JPR 2013]
- **pTop** [Sun et al., AC, 2016]
- **TopPIC** [Kou et al. Bioinformatics 2016]
- **ProteinGoggle** [Xiao et al. Scientific Reports, 2016]
- **Proteoform Suite** [Shortreed et al., JPR, 2016]
- **TopMG** [Kou et al. Bioinformatics 2017]
- **MSPPathFinder** [Park et al., Nature Methods 2017]
- **TDPPortal** [Northwestern, ~2017]
- **PERCEPTRON** [Khalid et al., Nucleic Acid Res 2021]
Several packages with complete solutions for top-down proteomics applications

- TDPortal
  http://nrtdp.northwestern.edu/tdportal-request/
- ProSightPC™
  https://www.thermofisher.com/order/catalog/product/PROSIGHTPC10
- Mash Explorer
  https://labs.wisc.edu/gelab/MASH_Explorer/index.php
- Informed-Proteomics
- TopPIC
  http://www.toppic.org/
Deconvolution of top-down mass spectra

Top-down mass spectra usually have many peaks and complex patterns of isotopic envelopes. This spectrum has about 19,000 peaks.
Spectral alignment for blind PTM search

Spectrum of prefix ions for $\text{PRS}^{+80}\text{TRING}$
\{0, 97, 253, 420, 521, 677, 790, 904, 961\}

Database protein: PRSTRING

Spectral alignment
Spectral alignment with $F$ modifications is a diagonal path from the top left node to the bottom right node with at most $F$ breaks.

Spectral alignment score
Number of 2-D points $(a_i, b_j)$ that the path passes through.
Ultramodified proteoforms

- Histone H4 has billions of possible proteoforms

- Histone H4 proteoform identified by top-down MS