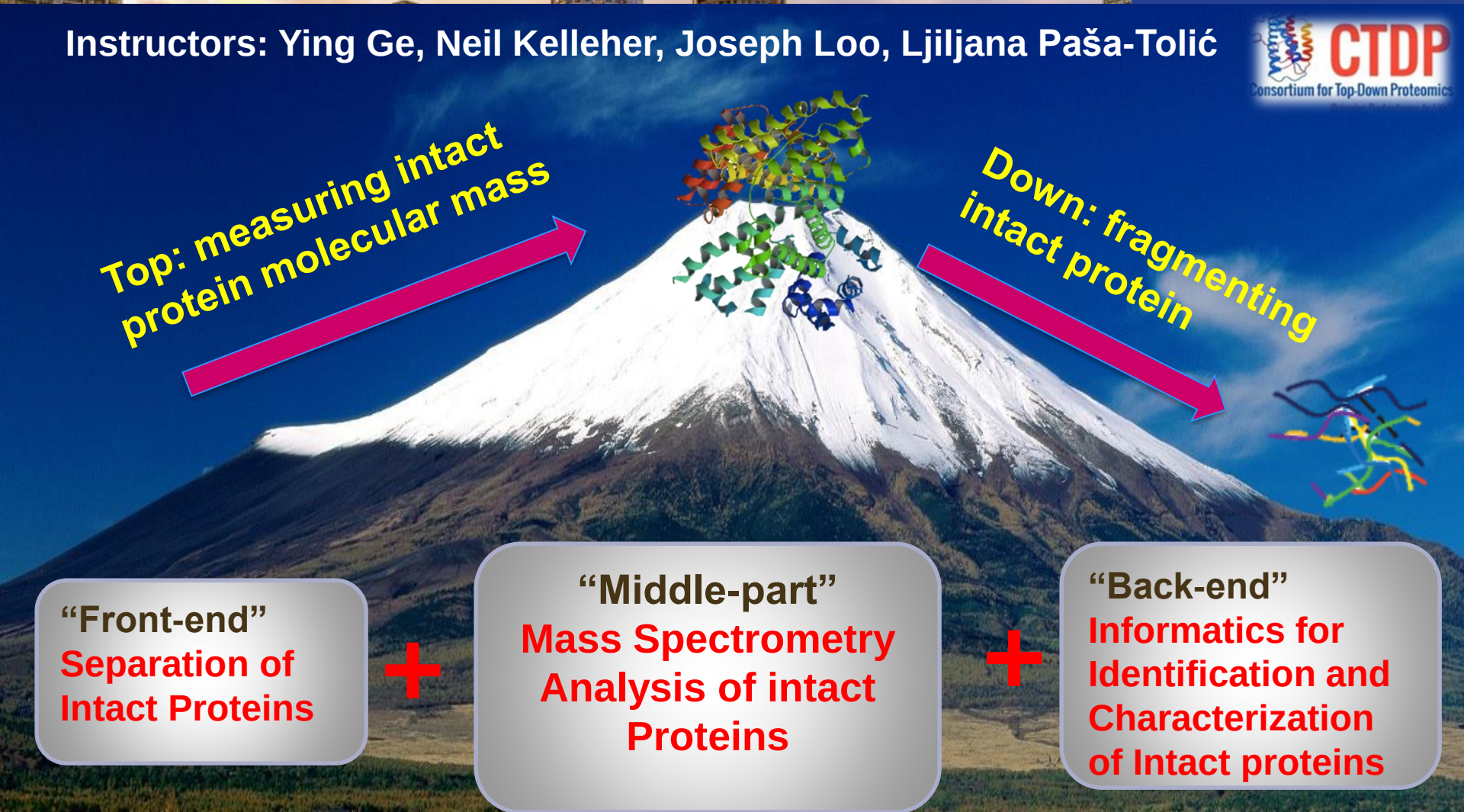
# ASMS Short Course

# Top-Down Proteomics

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

CTDP
Consortium for Top-Down Proteomics




"Front-end"
Separation of Intact Proteins

+

"Middle-part"
Mass Spectrometry Analysis of intact Proteins

+

"Back-end"
Informatics for Identification and Characterization of Intact proteins

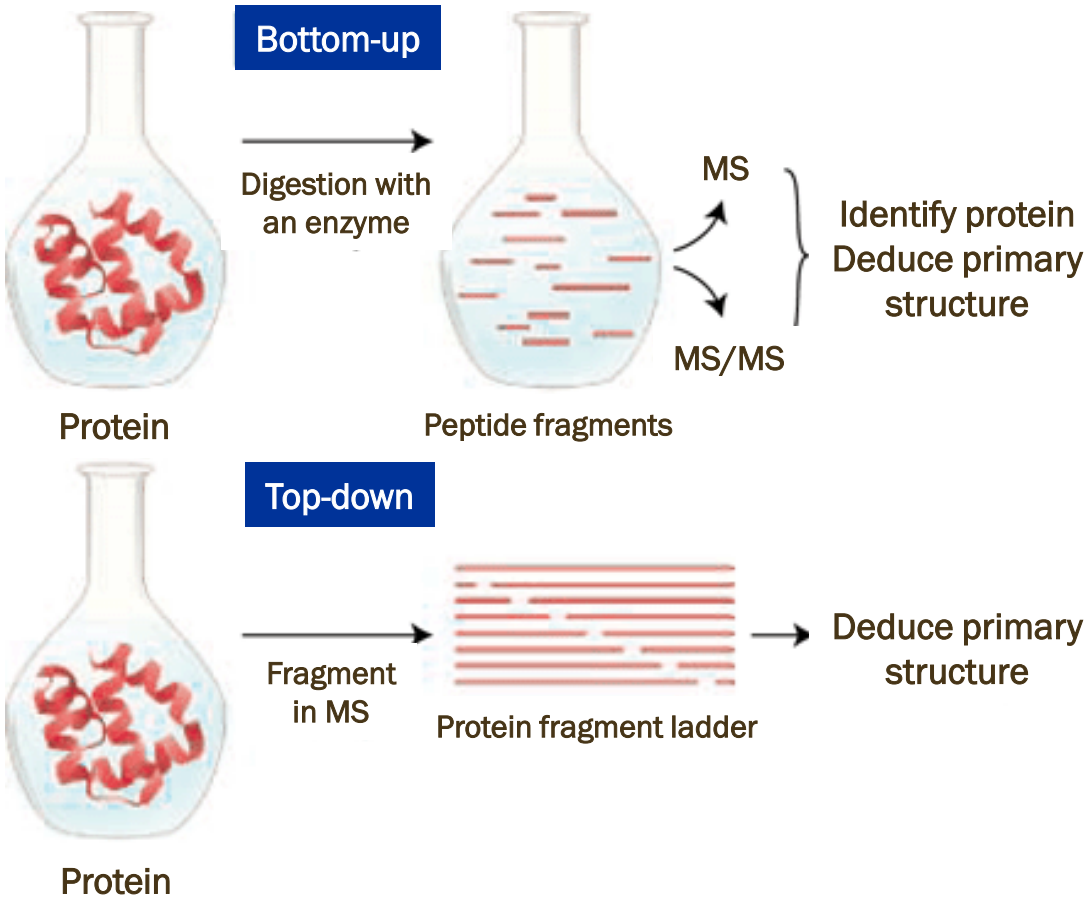
# Topics to Be Covered

- 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

SCIENCE VOL 314 6 OCTOBER 2006

# Mass Spectrometry: Bottom-Up or Top-Down?

Brian T. Chait



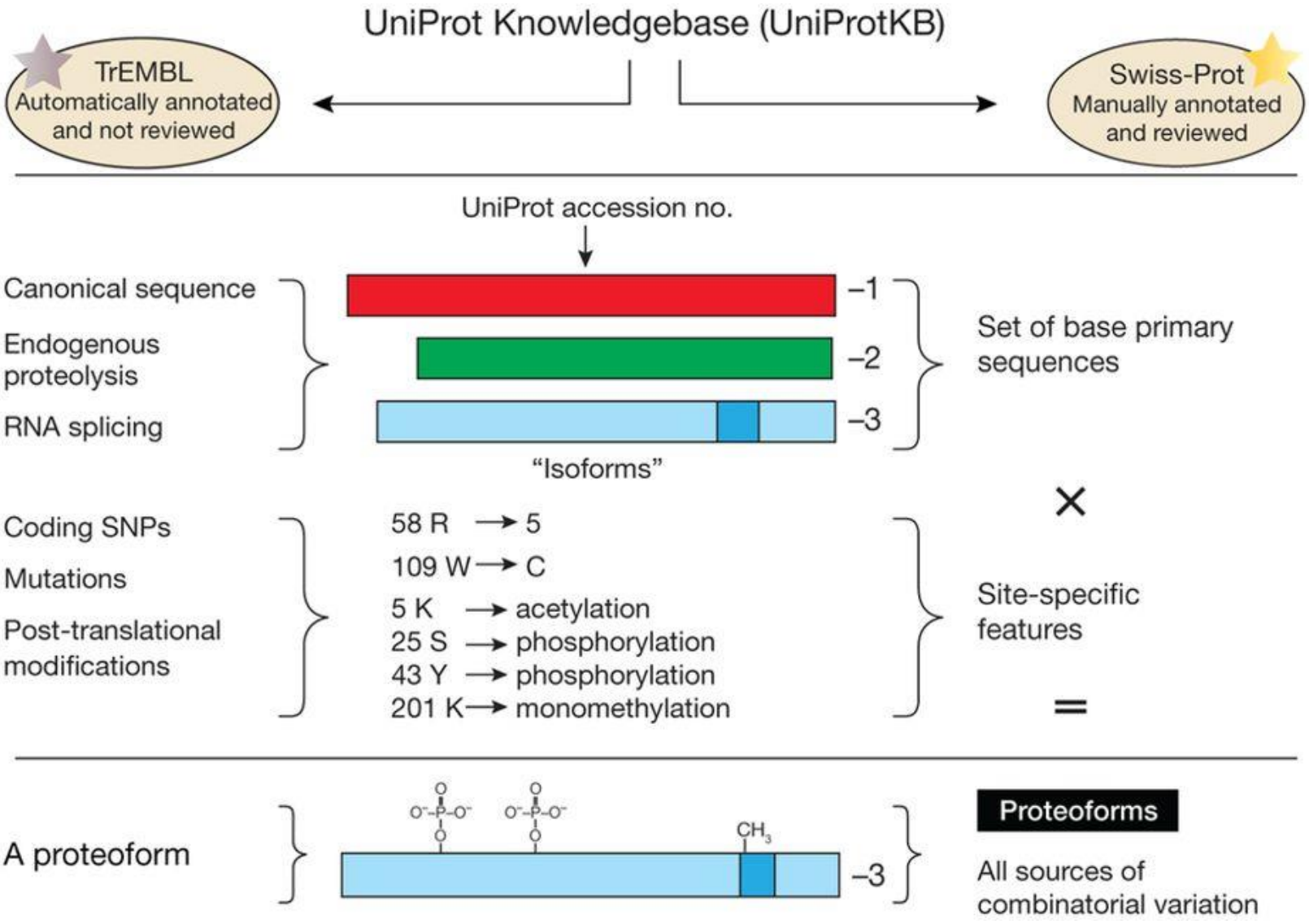

The bottom-up approach is (therefore) *suboptimal for determining modifications* and alternative splice variants.

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

CTDP
Consortium for Top-Down Proteomics

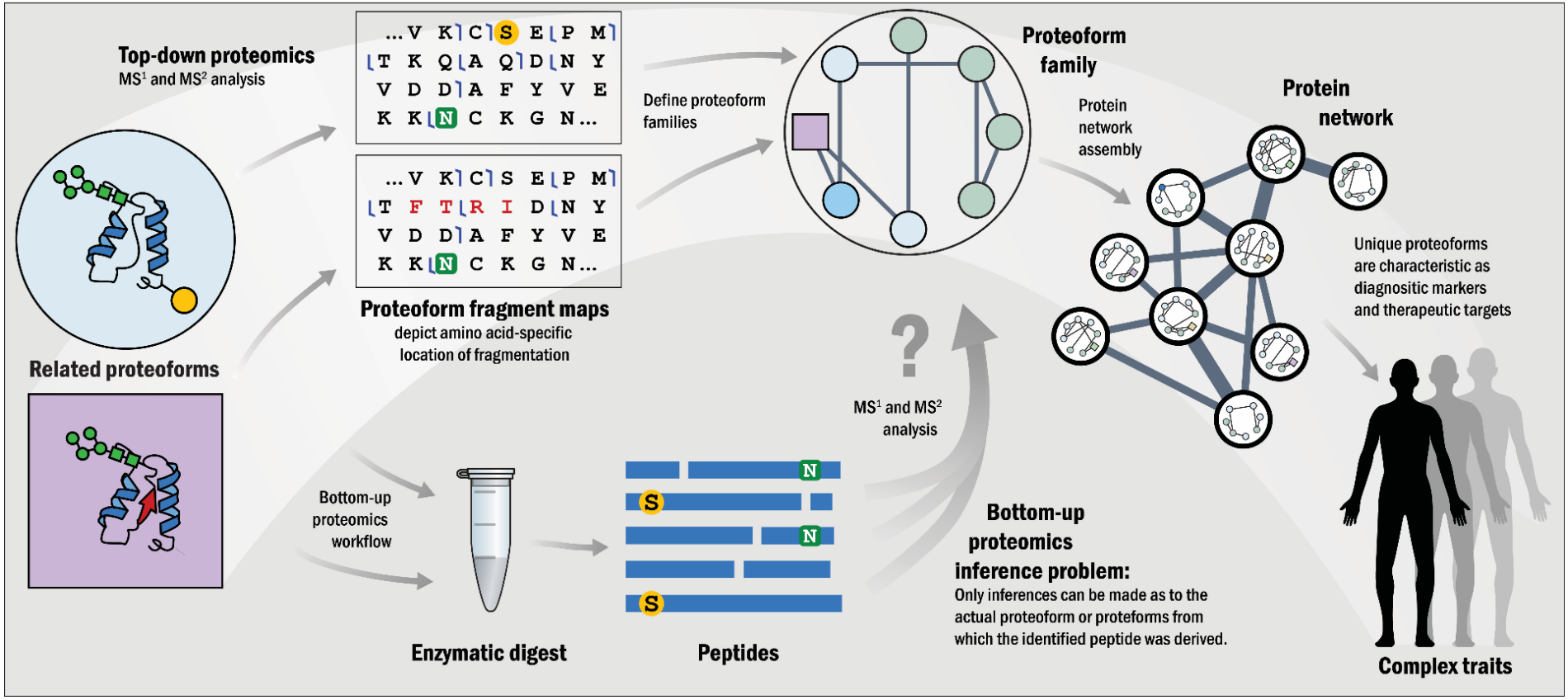
## Proteoform: a single term describing protein complexity




Smith, Kelleher & Consortium for Top-Down Proteomics. *Nat Methods* 2013, *10* (3), 186-7.

# Proteoforms as a New "Currency" in Proteomics

CTDP
Consortium for Top-Down Proteomics

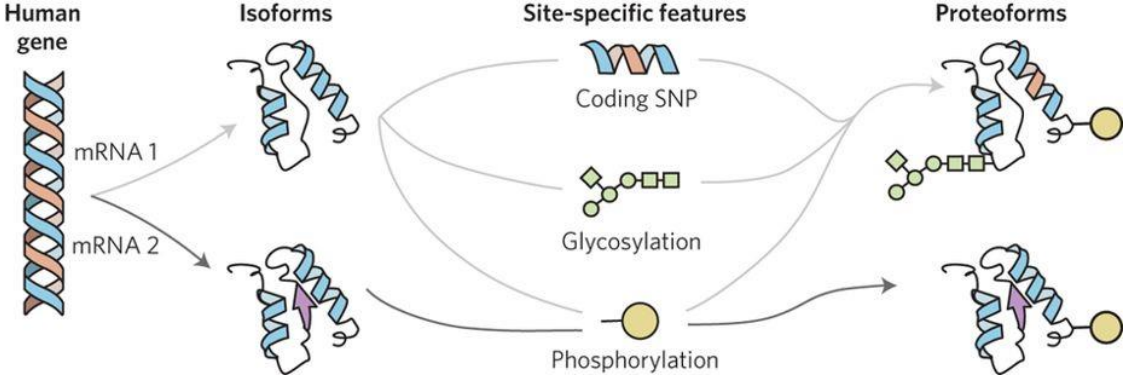



PERSPECTIVE
PUBLISHED ONLINE: 14 FEBRUARY 2018 | DOI: 10.1038/NCHEMBIO.2576
nature chemical biology

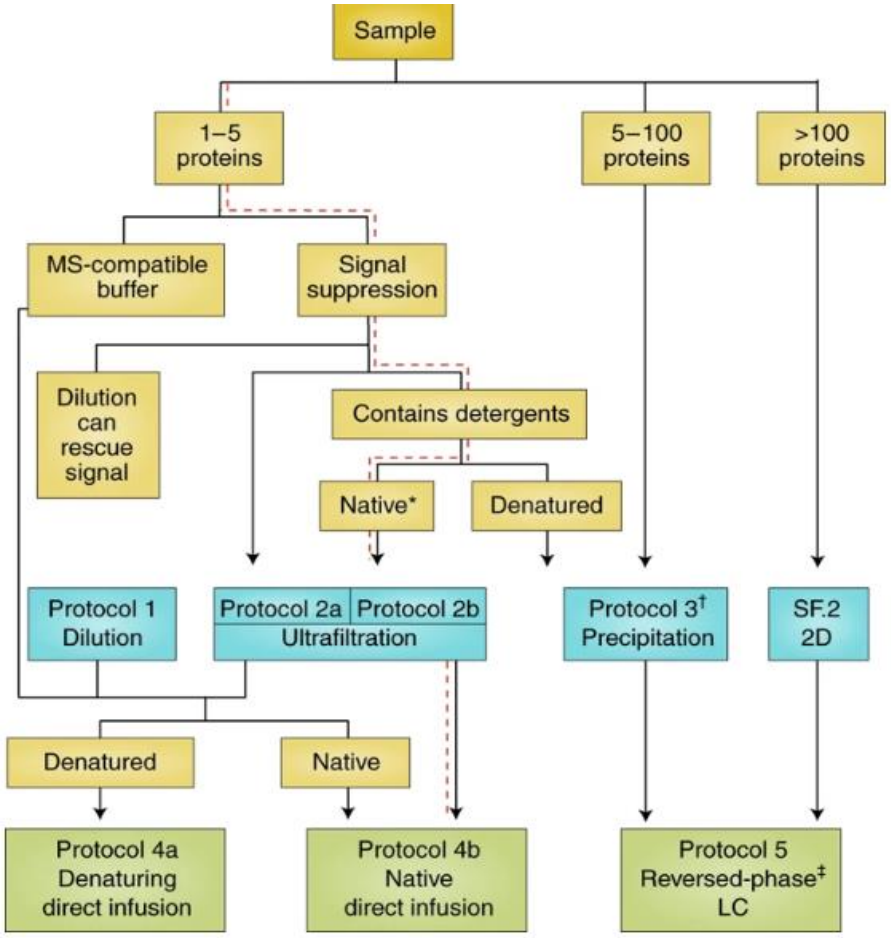
## How many human proteoforms are there?

Ruedi Aebersold[1], Jeffrey N Agar[2], I Jonathan Amster[3], Mark S Baker[4], Carolyn R Bertozzi[5], Emily S Boja[6], Catherine E Costello[7], Benjamin F Cravatt[8], Catherine Fenselau[9], Benjamin A Garcia[10], Ying Ge[11,12], Jeremy Gunawardena[13], Ronald C Hendrickson[14], Paul J Hergenrother[15], Christian G Huber[16], Alexander R Ivanov[2], Ole N Jensen[17], Michael C Jewett[18], Neil L Kelleher[19]*, Laura L Kiessling[20], Nevan J Krogan[21], Martin R Larsen[17], Joseph A Loo[22], Rachel R Ogorzalek Loo[22], Emma Lundberg[23,24], Michael J MacCoss[25], Parag Mallick[5], Vamsi K Mootha[13], Milan Mrksich[18], Tom W Muir[26], Steven M Patrie[19], James J Pesavento[27], Sharon J Pitteri[5], Henry Rodriguez[6], Alan Saghatelian[28], Wendy Sandoval[29], Hartmut Schlüter[30], Salvatore Sechi[31], Sarah A Slavoff[32], Lloyd M Smith[12,33], Michael P Snyder[24], Paul M Thomas[19], Mathias Uhlén[34], Jennifer E Van Eyk[35], Marc Vidal[36], David R Walt[37], Forest M White[38], Evan R Williams[39], Therese Wohlschlager[16], Vicki H Wysocki[40], Nathan A Yates[41], Nicolas L Young[42] & Bing Zhang[42]

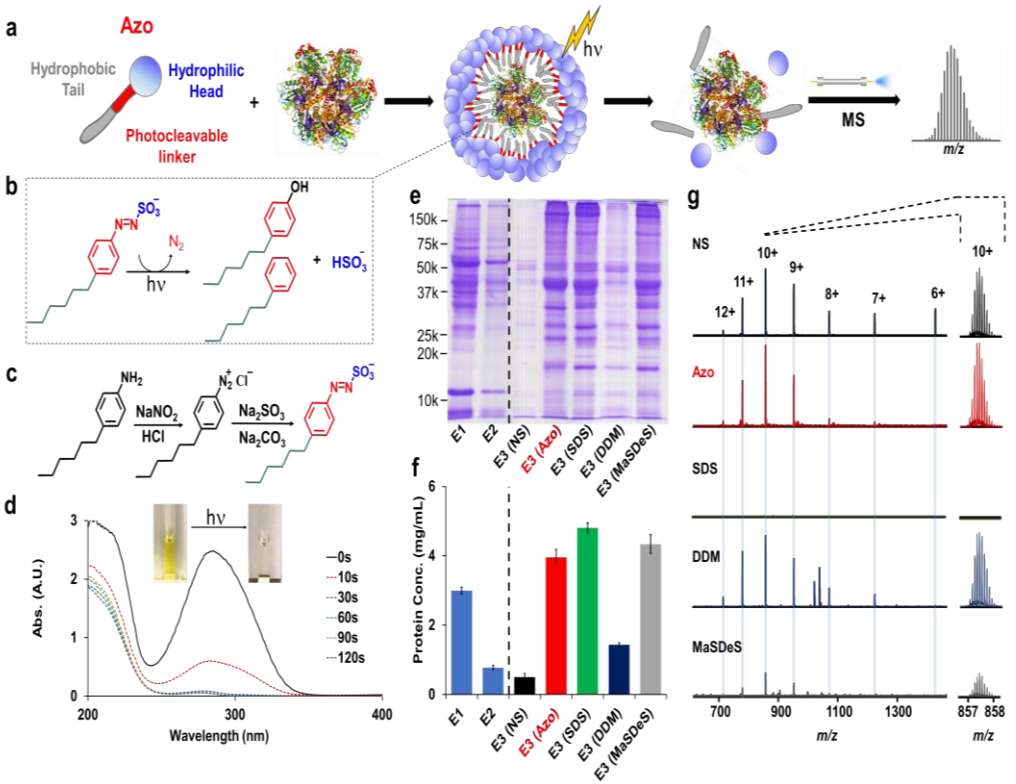
*Smith and Kelleher, Science* 2018, *359* (6380), 1106-1107.



# Sample Preparation in Top-down Proteomics




## A Photo-cleavable Surfactant for Top-down Proteomics



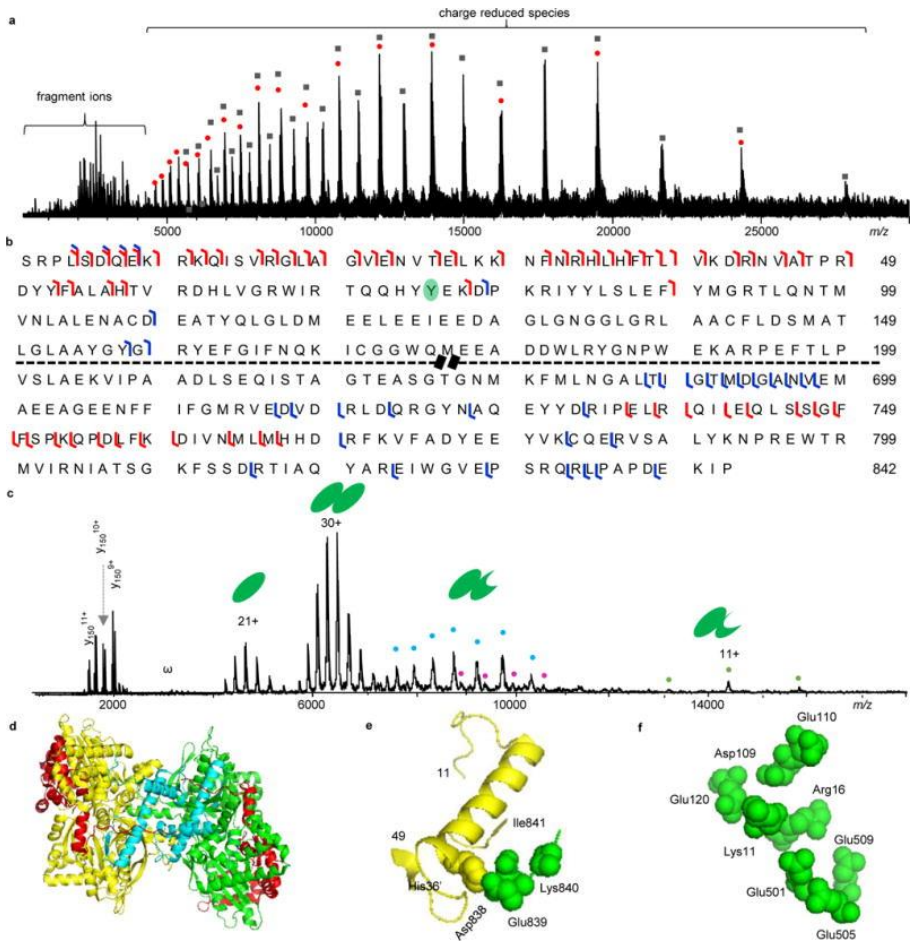

**Best Practices and Benchmarks for Intact protein Analysis for Top-down Mass Spectrometry**

Donnelly et al. *Nature Methods.* 2019, 16, 587-594

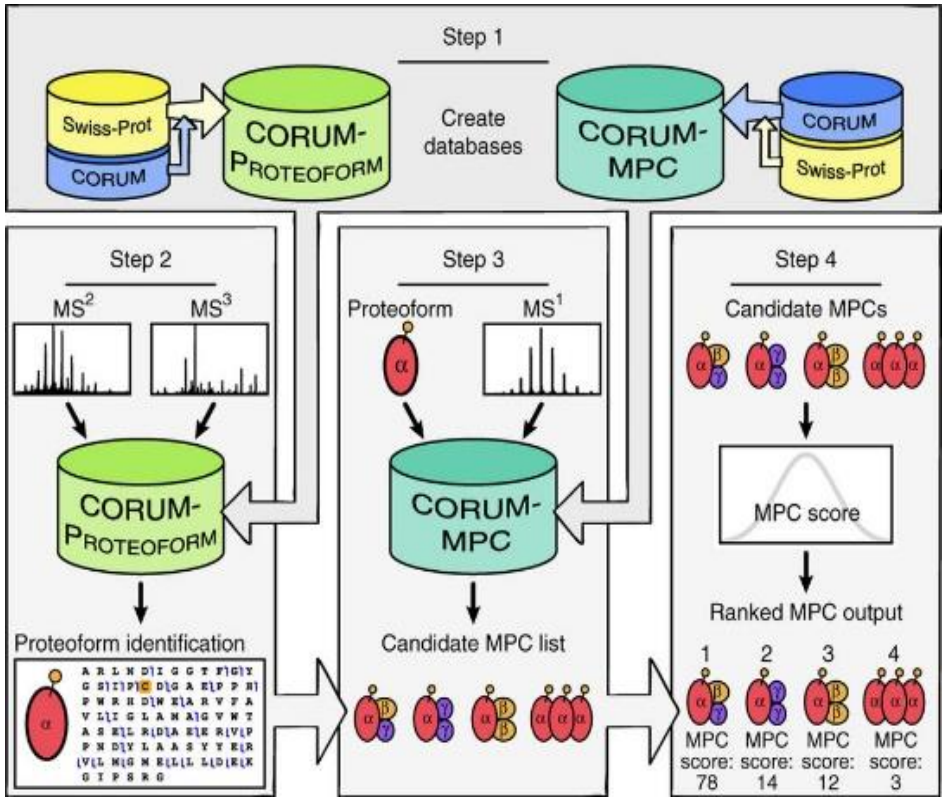
Brown et al. *Nature Methods* 2019, 16, 417-420

CTDP
Consortium for Top-Down Proteomics

# Native Top-down Proteomics




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




Skinner et al. Nat. Methods 2016, 13, 237

Integrated Native Mass Spectrometry and Top-Down Proteomics
-Connect Sequence to Structure and Function of Macromolecular Complexes

CTDP
Consortium for Top-Down Proteomics

# 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

**Informed Proteomics:** https://github.com/PNNL-Comp-Mass-Spec/Informed-Proteomics

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

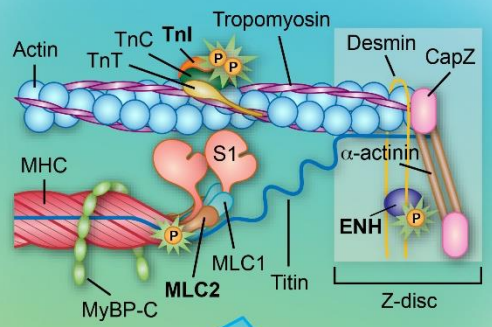
**pTop**: http://pfind.net/software/pTop/index.html

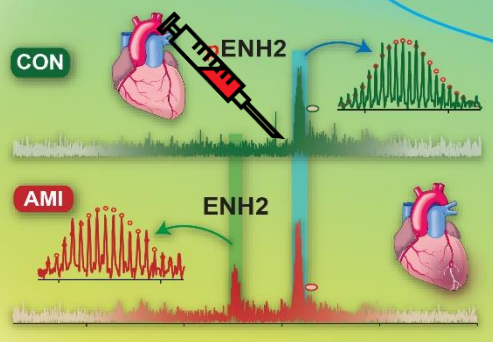
CTDP
Consortium for Top-Down Proteomics

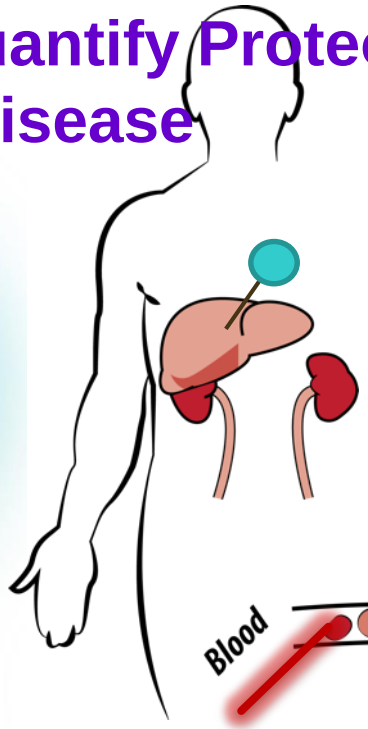
http://www.topdownproteomics.org/resources/software/

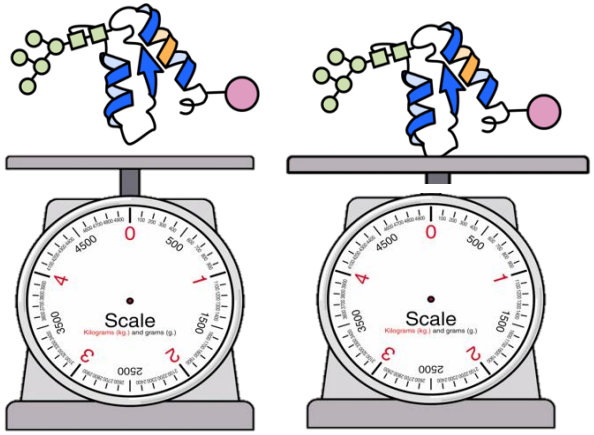
# Biomedical and Biopharmaceutical Applications

## Discover, Characterize & Quantify Proteoforms In Health and Disease

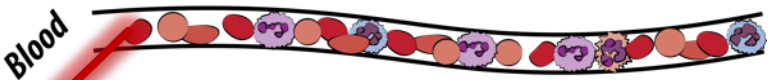


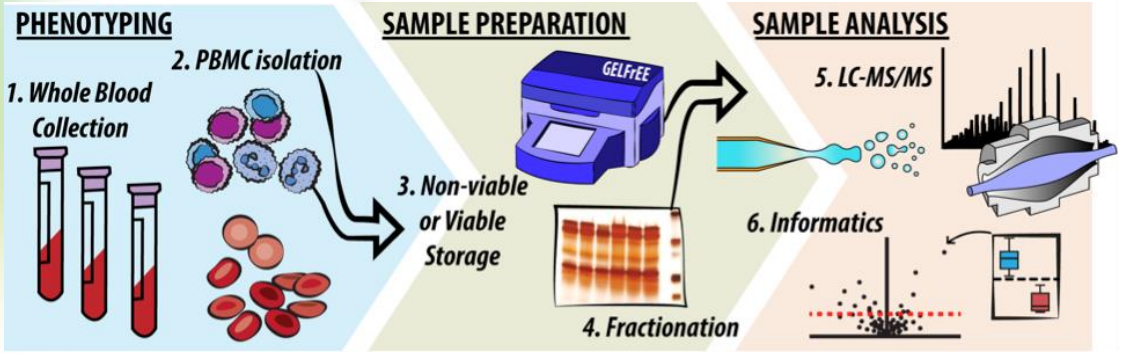





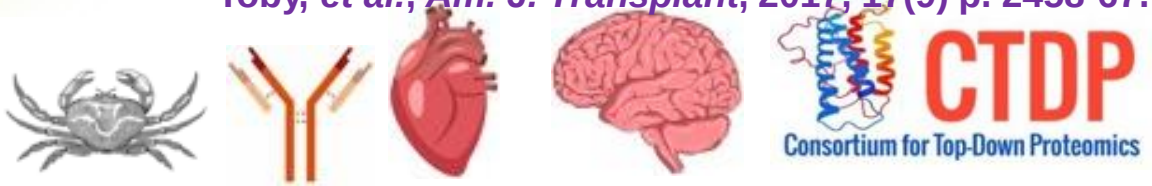

**Northwestern | Proteomics**



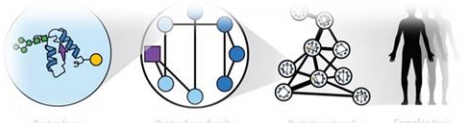

Toby, *et al.*, *Am. J. Transplant*, 2017, 17(9) p. 2458-67.

Peng, et al. Mol. Cell. Proteomics, 2014, 13, 2752-2764.



# 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**

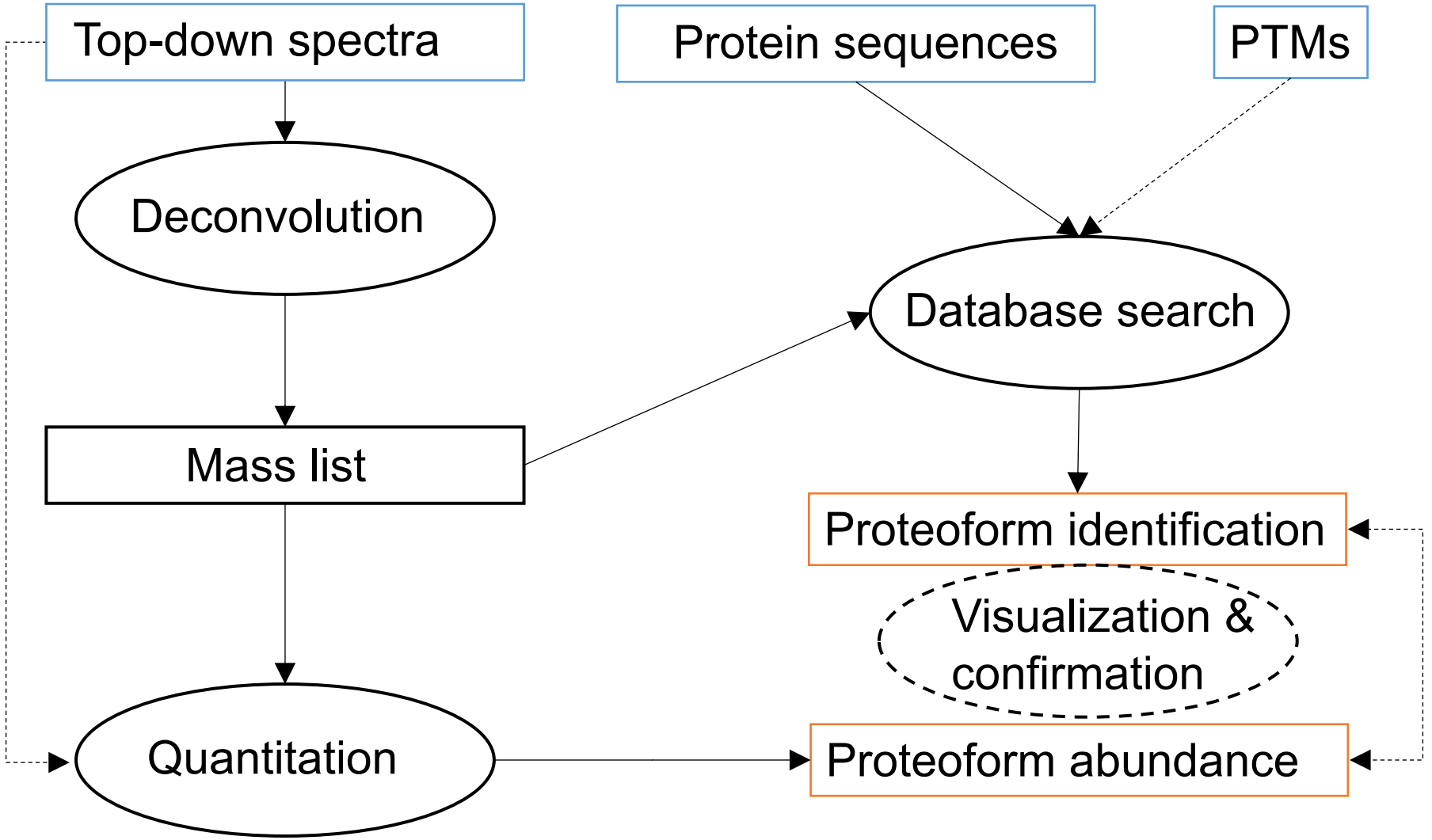
- **Proteoform-resolved biology will increase efficiency of basic and translational research**

# ASMS Short Course

## Top-down Mass Spectrometry Data Analysis and Visualization

Xiaowen (Kevin) Liu, Kyowon Jeong, Eli Larson and Ryan Fellers

# Top-Down MS Data Analysis workflow



# Top-down MS software

## Spectral Deconvolution

- **Thrash** [Horn *et al*., JASMS 2000]
- **Thrash/Xtract** [Horn *et al*. JASMS 2000,Zabrouskov 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]
- **MSPathFinder** [Park et al., Nature Methods 2017]
- **TDPortal** [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™

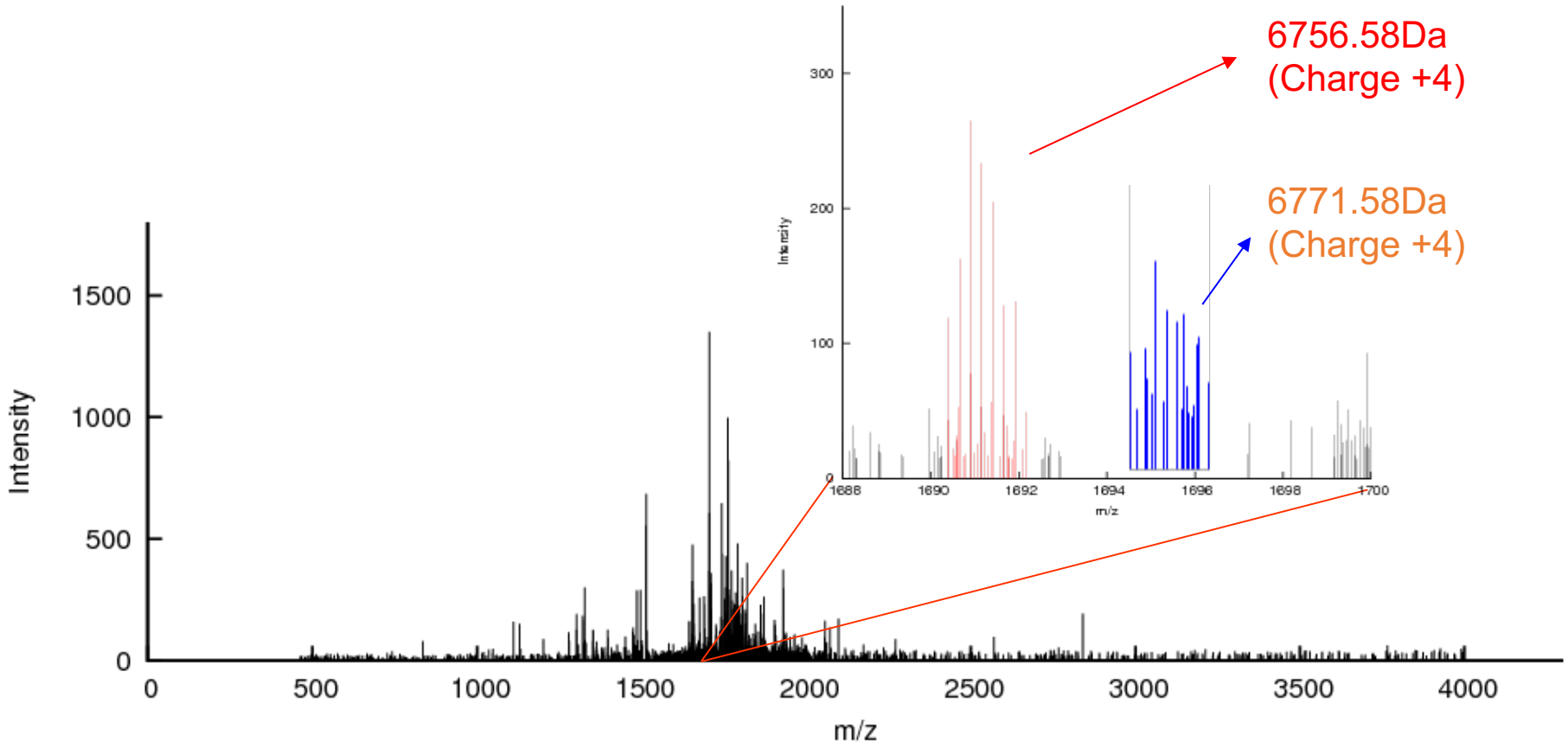
  https://www.thermofisher.com/order/catalog/product/PROSIGHTPC10
- Mash Explorer

  https://labs.wisc.edu/gelab/MASH_Explorer/index.php
- Informed-Proteomics

  https://github.com/PNNL-Comp-Mass-Spec/Informed-Proteomics
- TopPIC

  http://www.toppic.org/

# Deconvolution of top-down mass spectra

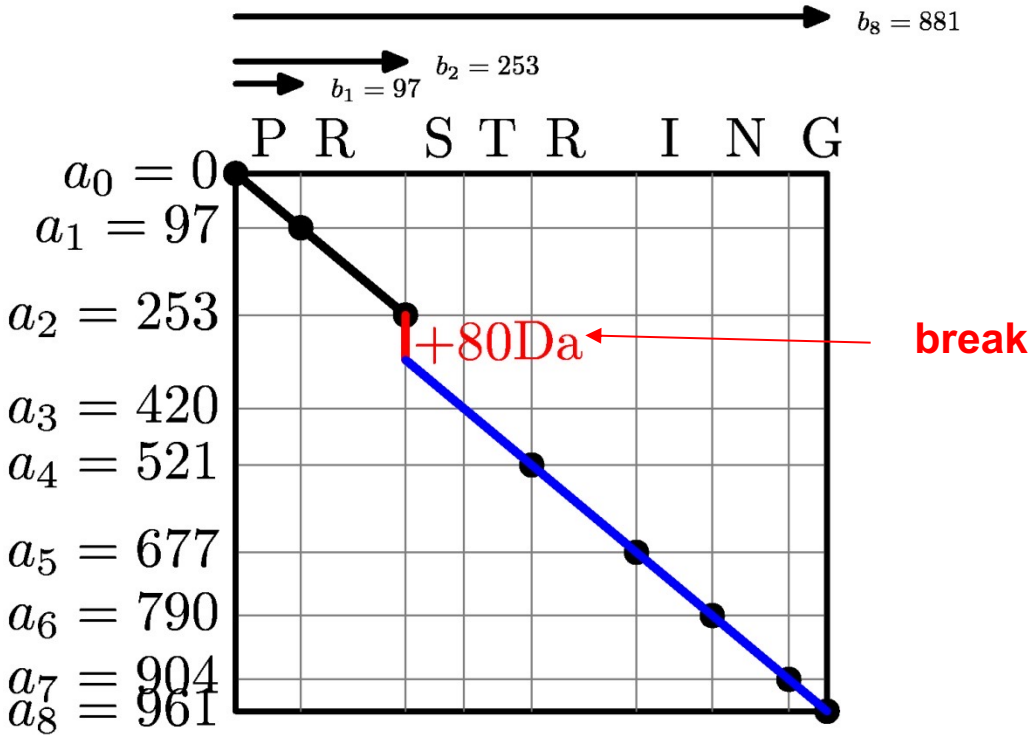



Top-down mass spectra usually have many peaks and complex patterns of **isotopic envelopes.** This spectrum has about 19,000 peaks.

# Spectral alignment for blind PTM search

**S**pectrum of prefix ions for **PRS$^{+80}$TRING**
**{0, 97, 253, 420, 521, 677, 790, 904, 961}**

Database protein: **PRSTRING**




## Spectral alignment

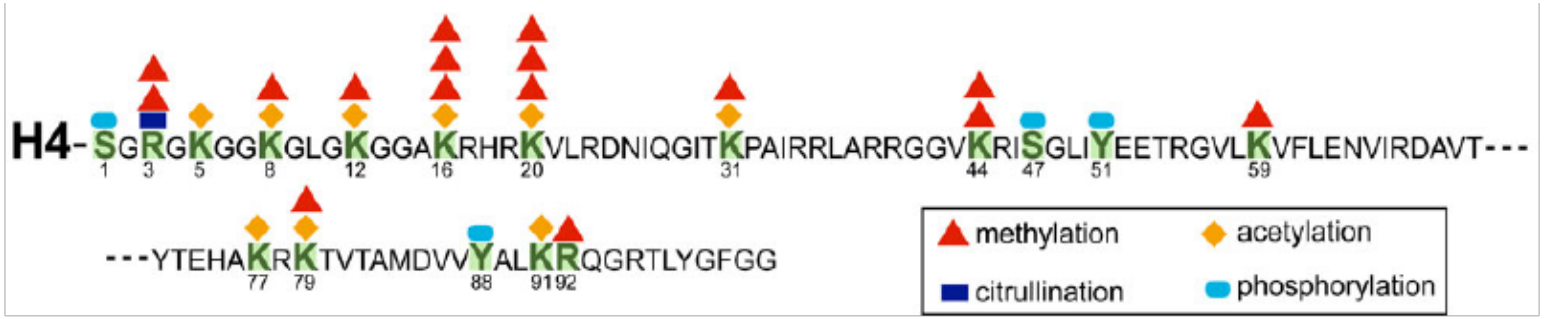
Spectral alignment with *F* modifications is a diagonal path from the top left node to the bottom right node with at most *F* **breaks**.

## 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