

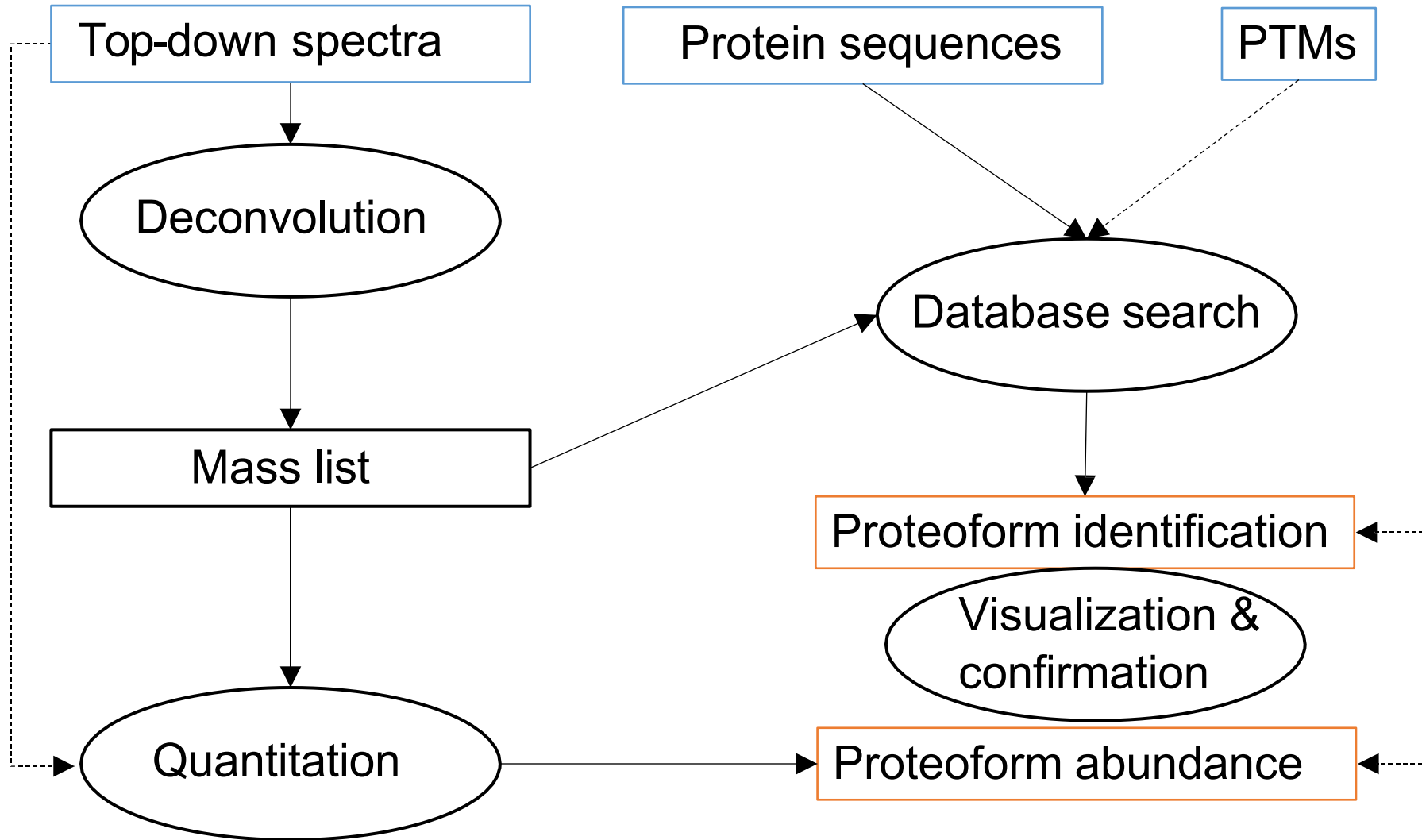
ASMS Short Course



Top-down Mass Spectrometry Data Analysis and Visualization

Xiaowen (Kevin) Liu, Kyowon Jeong, and Boris Krichel

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/>

- Mash Explorer

https://labs.wisc.edu/gelab/MASH_Explorer/index.php

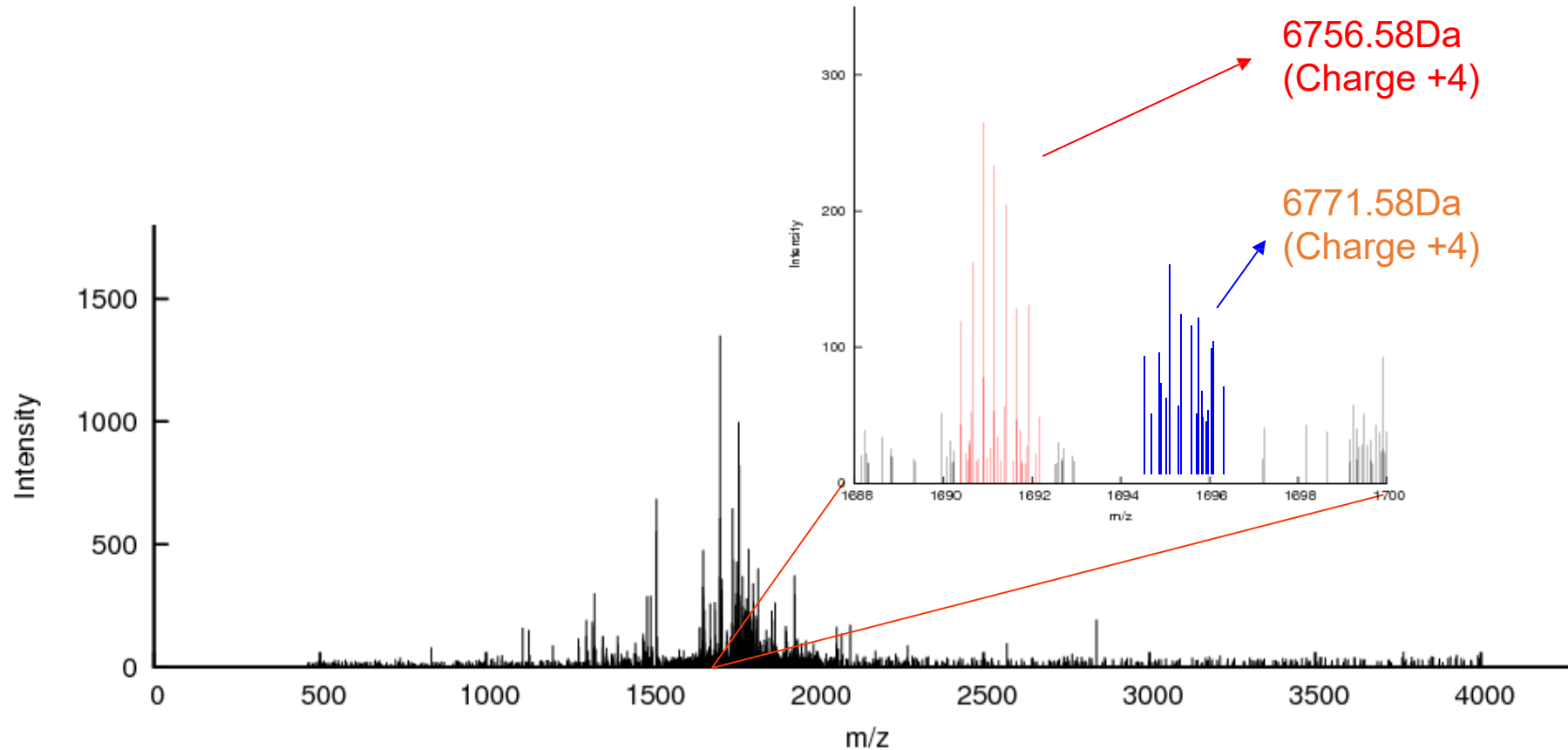
- FlashDeconv

<https://openms.de/FLASHDeconv>

- TopPIC

<http://www.toppic.org/>

Deconvolution of top-down mass spectra

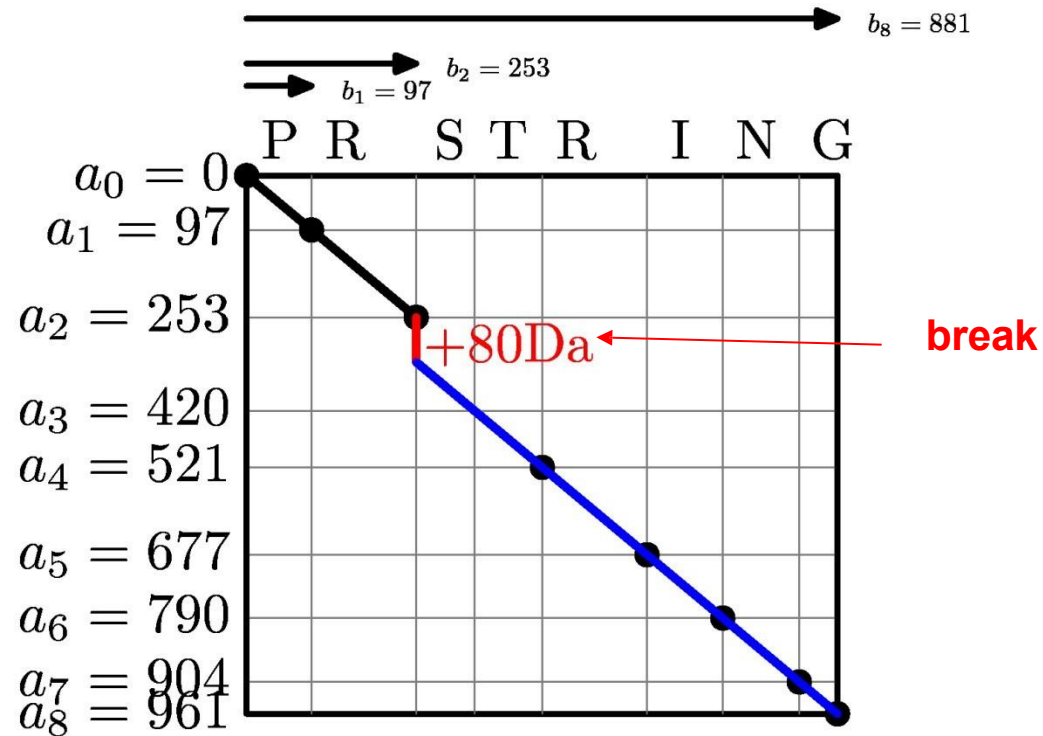


Top-down mass spectra usually have many peaks and complex patterns of **isotopic envelopes**. This spectrum has about 19,000 peaks.

Spectral alignment for blind PTM search

Spectrum of prefix ions for **PR^S+80TRING**
{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

