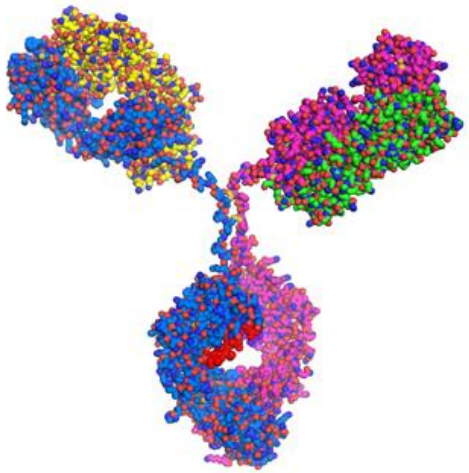
Applications of Mass Spectral Libraries to Therapeutic Protein Analysis:

From Simultaneous Assessment of Post-Translational Modification Sites to Multi-Attribute Quality Control



NISTmAb

Qian Dong, Yuxue Liang, Xinjian Yan, and Stephen E. Stein
Mass Spectrometry Data Center
National Institute of Standards and Technology (NIST)
Gaithersburg, MD USA

September 18th, 2019
MASS SPEC 2019, CASSS

OUTLINE

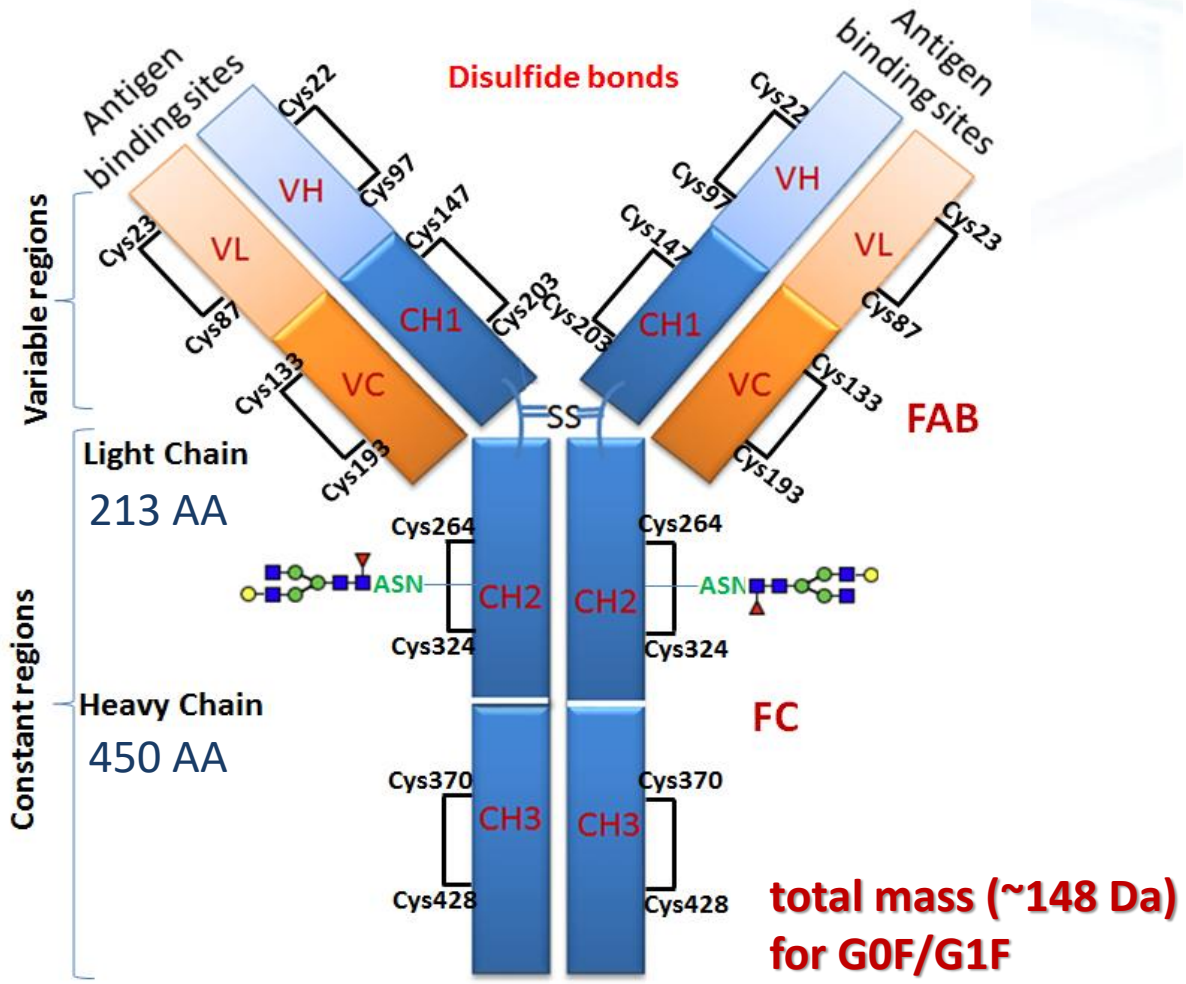
1. Challenges in MS bioanalysis: highly complex low abundance PTMs
2. Methods to create spectral libraries with 1D and 2D high mass accuracy/high resolution MS data for deep coverage of major PTMs
3. Examples of Applications of the NISTmAb libraries to simultaneous detection of various peptides and modifications
4. Advantages of the spectral library search relative to conventional methods

Challenge in Mapping PTMs in mAbs by LC-MS/MS

1. PTMs are diverse and divergent (>30)
2. PTMs are low abundance species
3. Extensive distribution sites in mAbs
   - 55 degradation sites
   - 51 glycation sites
   - 33 oxidation sites
   - 13 cysteine variant sites
   - 66 amino acid substitution site
4. Heterogeneous glycosylation at a single site
   - 60 detectable glycans
5. Complex and unexpected fragmentation patterns
   - glycation, glycosylation, and disulfide linkage
6. Experiment-induced modifications
   - oxidation, deamidation, formylation

NISTmAb: First Publicly Available Reference Material for Biological Drug Manufacturers and Regulators

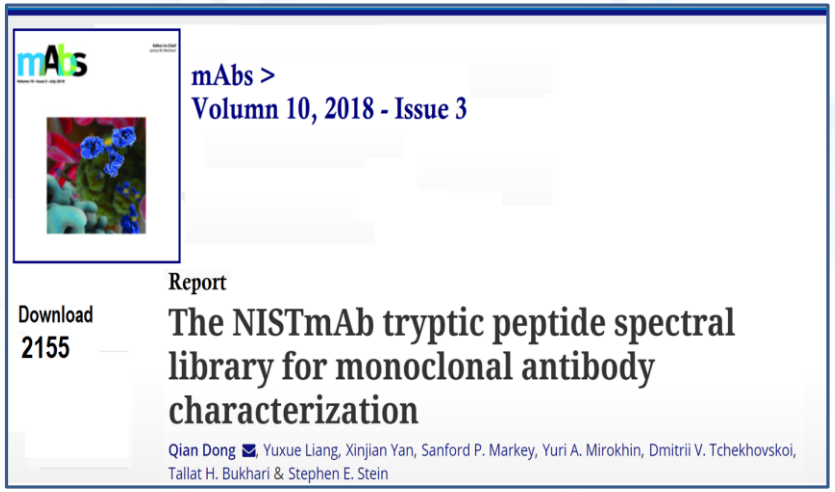
https://www-s.nist.gov/srmors/view_detail.cfm?srm=8671

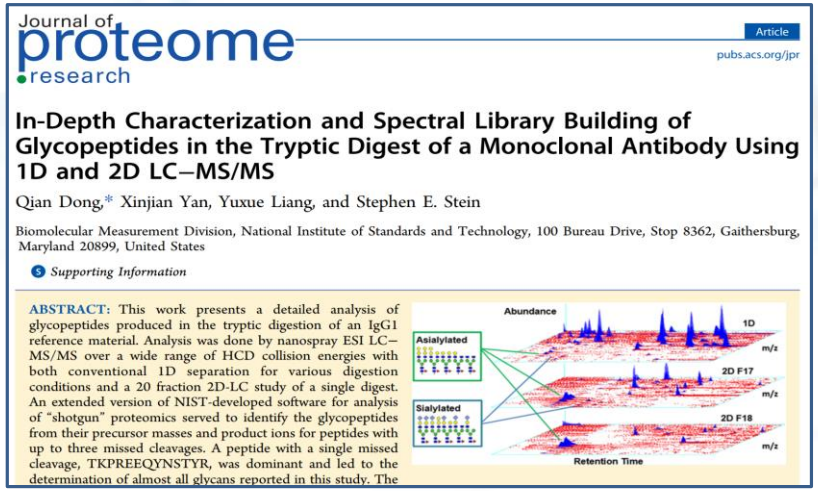


- Toward the goal of improving quality bioanalysis measurements
  - NIST released a humanized IgG1k reference mAb, **NISTmAb**, in 2016
  - It is a common control reference mAb for originator and biosimilar alike
- John Schiel and his team at NIST

  Schiel JE, Davis DL, Borisov OV, editors. State-of-the-art and emerging technologies for therapeutic monoclonal antibody characterization. **Three volumes**. Washington, D.C.: American Chemical Society; 2015
- It is the most widely characterized publicly available monoclonal antibody. Around 30 papers have been published since 2016

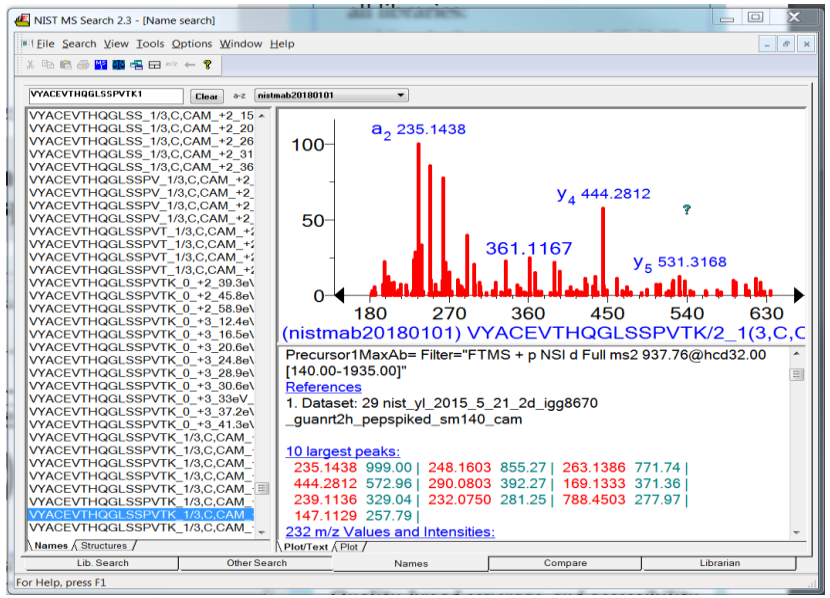
Spectral Libraries of NISTmAb for Simultaneously Analyzing Multiple Posttranslational Modifications

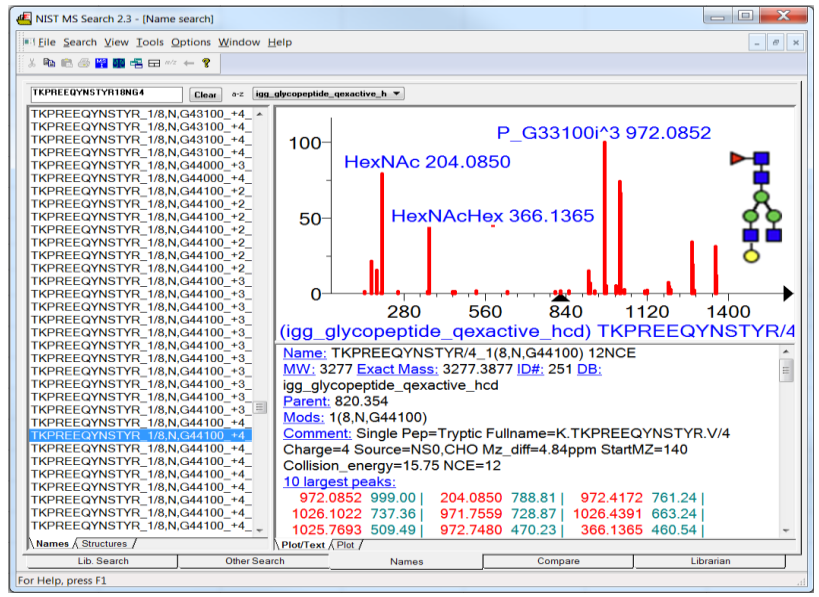




2018:
Tryptic peptide Spectral Library

13,000 spectra
3,300 peptides
50% modified spectra
spectral metadata

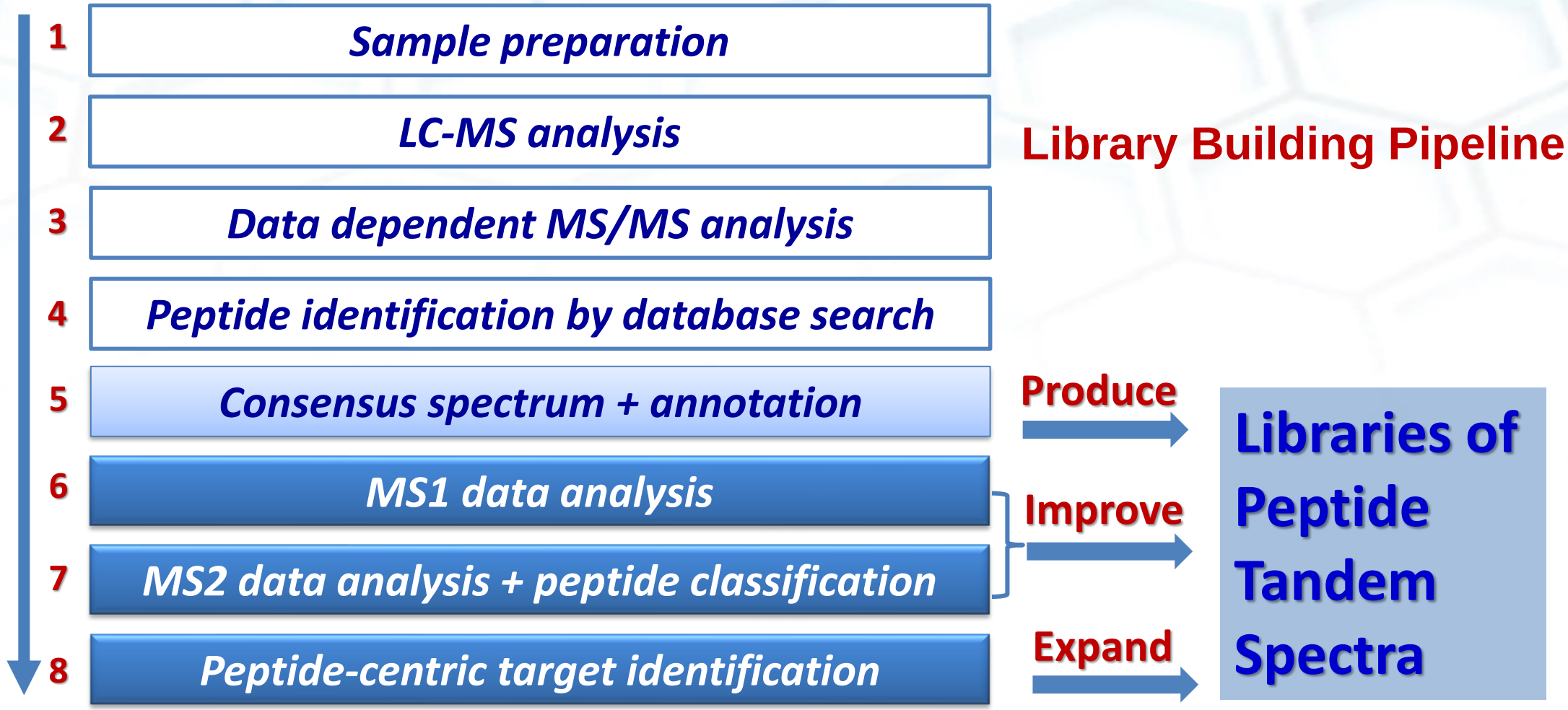




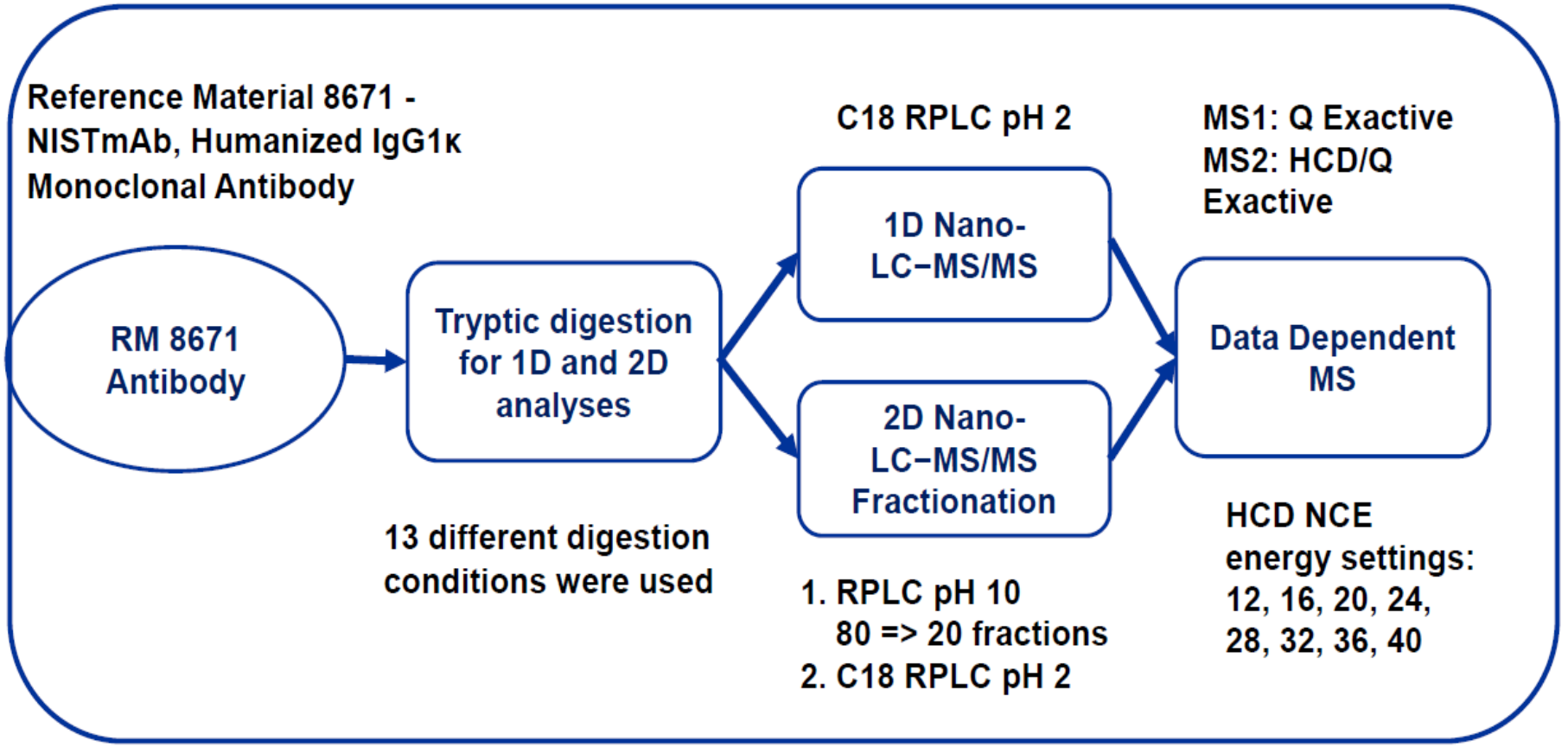
2016:
Glycopeptide Spectral Library

1,700 spectra
247 glycopeptides
(NISTmAb, Avastin, Enbrel, Humira, and Rituxan)

Single Protein Digest Libraries for In-Depth Bioanalysis

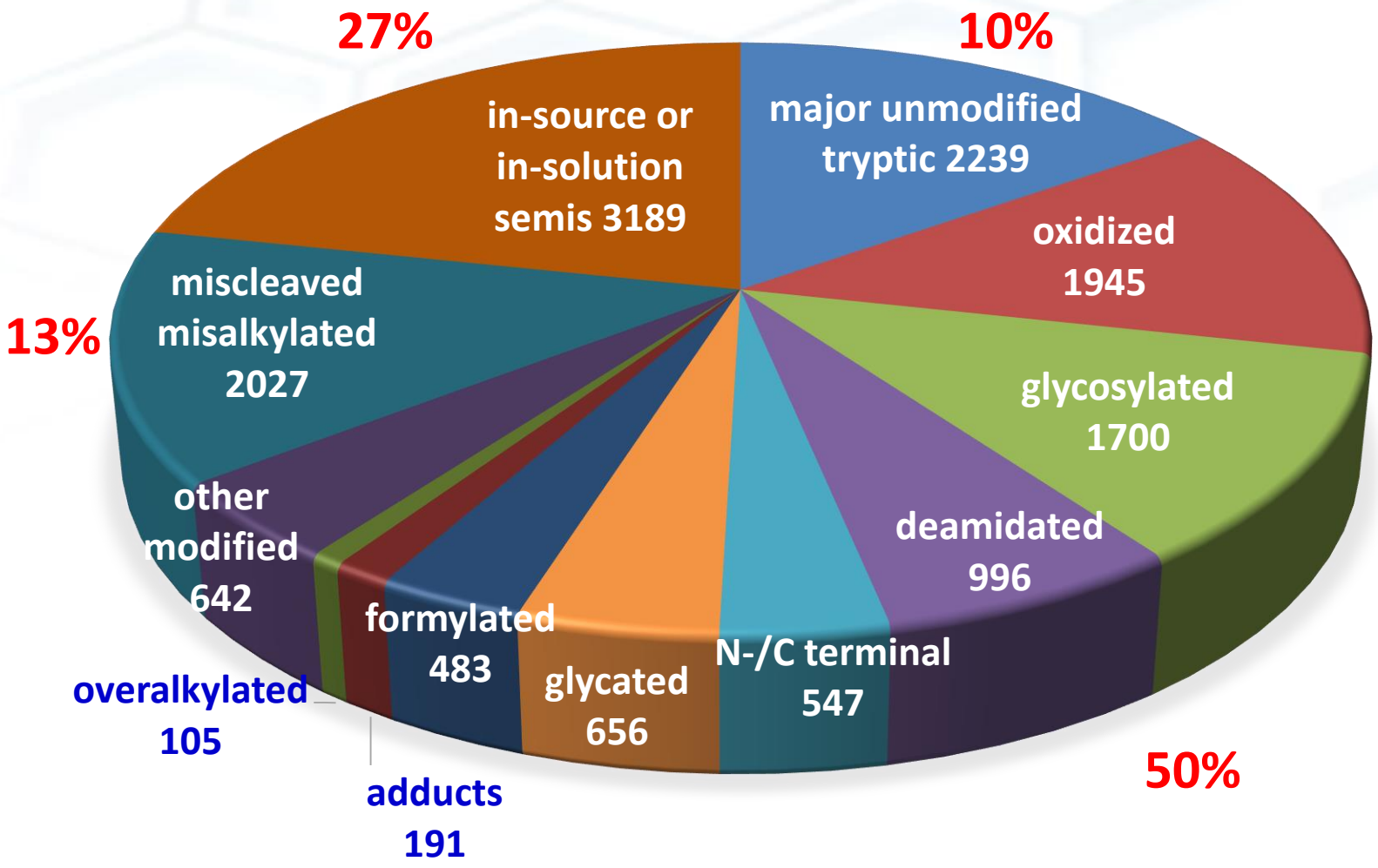


Improve Depth of Coverage in Single Protein Digests



- Many digestion conditions to obtain a broad coverage of peptides
- 2D LC separation to increase the sensitivity and reduce complexity
- Wide collision energy range to cover all fragmentation conditions

A First Comprehensive Spectral Reference Library for Characterization of Therapeutic Antibodies



- Around 13,000 high-quality tandem spectra of > 3,000 peptide ions
- Validation by:
  - accurate mass
  - differential elution
  - peptide classification
- Over 50% spectra for modified peptide ions

Usage of Commonly Observed Analytical Artifacts

Artifacts

- unexpected miscleavage
- misalkylation
- Overalkylation
- method-induced modifications
- semi-tryptic cleavage

1. Analytical analyses of biological samples produce many artifacts
2. Database searches require a significant computational cost to consider all possible artifacts
3. Artifacts represent a source of false positives if mis-identified
4. A spectral reference library readily facilitates the differentiation of true modifications from these artifacts
5. Artifacts can be used for assessing the quality of the sample preparation

25 Modification Categories Identified in the NISTmAb Library

Modification	Mass difference	# Spectra	Origin
Oxidation	15.9949	2482	in-sample, in-source, in-column or in-digestion
N-linked glycosylated	delta mass * 60	1700	in-sample
Deamidated	0.9840	996	in-sample or in-digestion
Formyl	27.9949	483	in-digestion
Hex	162.0528	656	in-sample or in-digestion
Gln->pyro-Glu	−17.0265	291	in-sample or in-digestion
Cation:Fe[III]	52.9115	60	in-source
Glu->pyro-Glu	−18.0106	86	in-digestion
Carbamidomethyl	57.0215	105	in-digestion
Cation:Fe[II]	53.9193	43	in-source
Carbamyl	43.0058	66	in-digestion
Lys (transpeptidation)	128.0950	70	in-digestion
Cation:Na	21.9819	37	in-source
Lys loss (heavy chain)	−128.0950	72	in-sample
Arg (transpeptidation)	156.1011	49	in-digestion
Cation:Ca[II]	37.9469	29	in-source
Cation:2Na	43.9639	22	in-source
Carboxy	43.9898	23	in-digestion
Dehydrated	−18.0106	36	in-source or in-digestion
Pyro-carbamidomethyl	39.9949	17	in-digestion
Ammonia-loss	−17.0265	15	in-sample or in-digestion
Nitro	44.9851	9	in-source
Acetyl	42.0106	5	in-sample or in-digestion
Dioxidation	31.9898	9	in-sample or in-source
Trioxidation	47.9847	1	in-sample or in-source

Determination of N-Linked Glycopeptides by 1D and 2D LC–MS/MS Analyses

Table 1. Identified Glycopeptides in 1D and 2D Studies[a]

peptide sequences of glycosite Asn300	1D studies		2D studies		charge states
	glycans	PepIDs	glycans	PepIDs	
TKP**RE**EQY<u>N</u>STYR (**M1**)	30	73	60	91	2+, 3+, **4+**
EEQY<u>N</u>STYR (**M0**)	21	56	55	88	2+, **3+**
TKP**RE**EQY<u>N</u>STY**RV**VSVLTVLHQDWLNG**KE**YK (**M3**)	10	23	13	27	4+, 5+, **6+**
EEQY<u>N</u>STY**RV**VSVLTVLHQDWLNG**KE**YK (**M2**)	10	15	9	16	3+, 4+, **5+**
EEQY<u>N</u>STY**RV**VSVLTVLHQDWLNGK	7	9	8	10	3+, **4+**
TKP**RE**EQY<u>N</u>STY**RV**VSVLTVLHQDWLNGK	7	12	4	8	4+, **5+**, 6+
P**RE**EQY<u>N</u>STYR	3	3	4	4	**3+**

- Total 1,703 tandem spectra of 247 glycopeptides obtained from NISTmAb and Avastin, Enbrel, Humira, and Rituxan
- Total 81 unique glycan structures (NS0 and CHO), including several in-source glycans

Dong, Q. Xinjian Yan, Yuxue Liang, and Stephen E. Stein. *Journal of Proteome Research* 2016 *15* (5), 1472-1486.

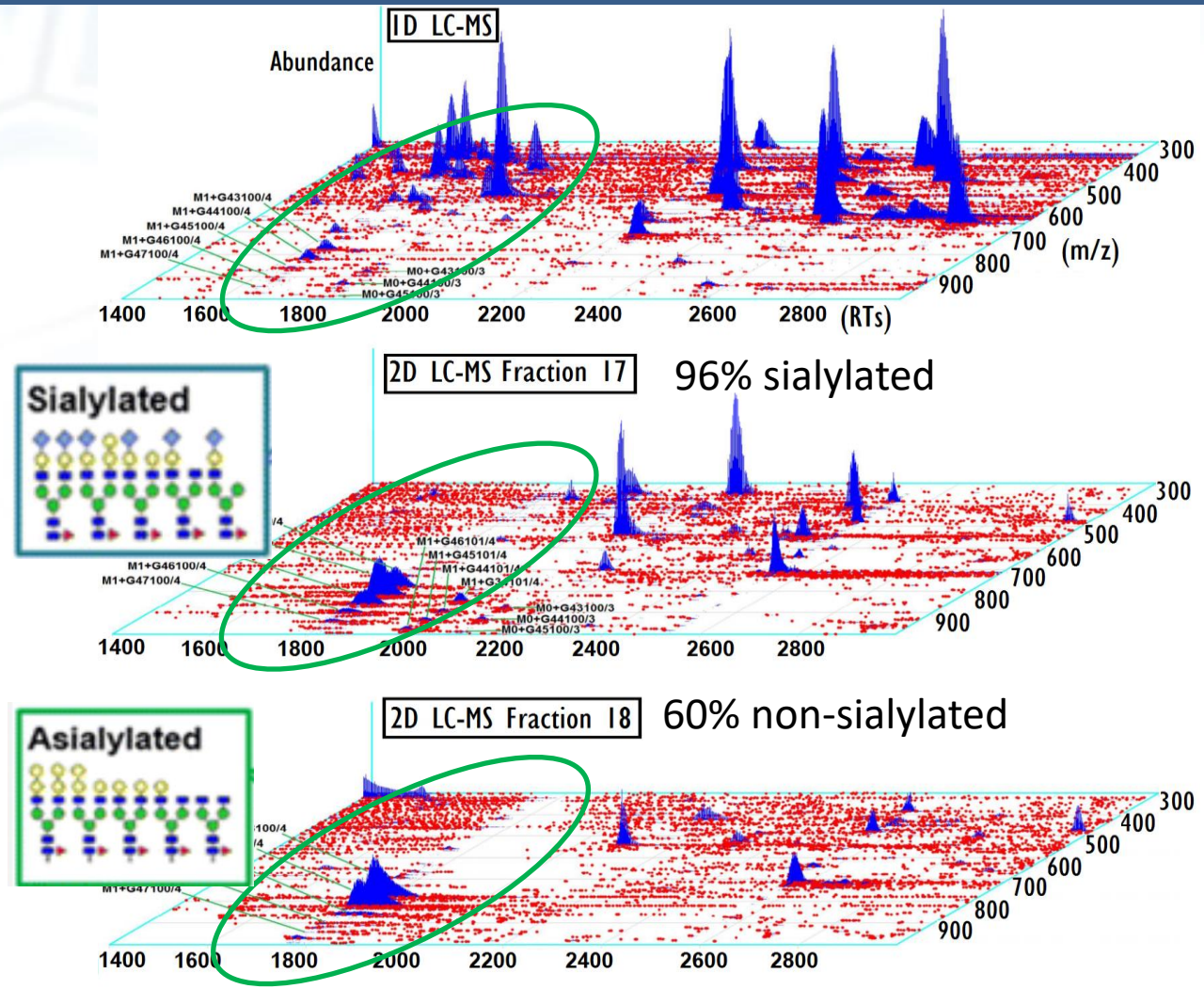
2D Studies Detected More Low Abundance NISTmAb Glycan Structures Than 1D Techniques

Group	Glycan structure classification	Number of unique glycan structures		
		Released glycan[2]	1D LC-MS[1] glycopeptide	2D LC-MS[1] glycopeptide
1	Core-fucosylated biantennary glycans	3	3	3
2	Core-fucosylated monoantennary glycans	4	4	4
3	Core-fucosylated triantennary glycans	4	3	4
4	Sialylated glycans with Neu5Gc	5	9	21
5	Galactosylated glycans	9	5	18
6	Other minor glycans	7	8	26
	Total:	**32**	**32**	**76**

25 new glycans

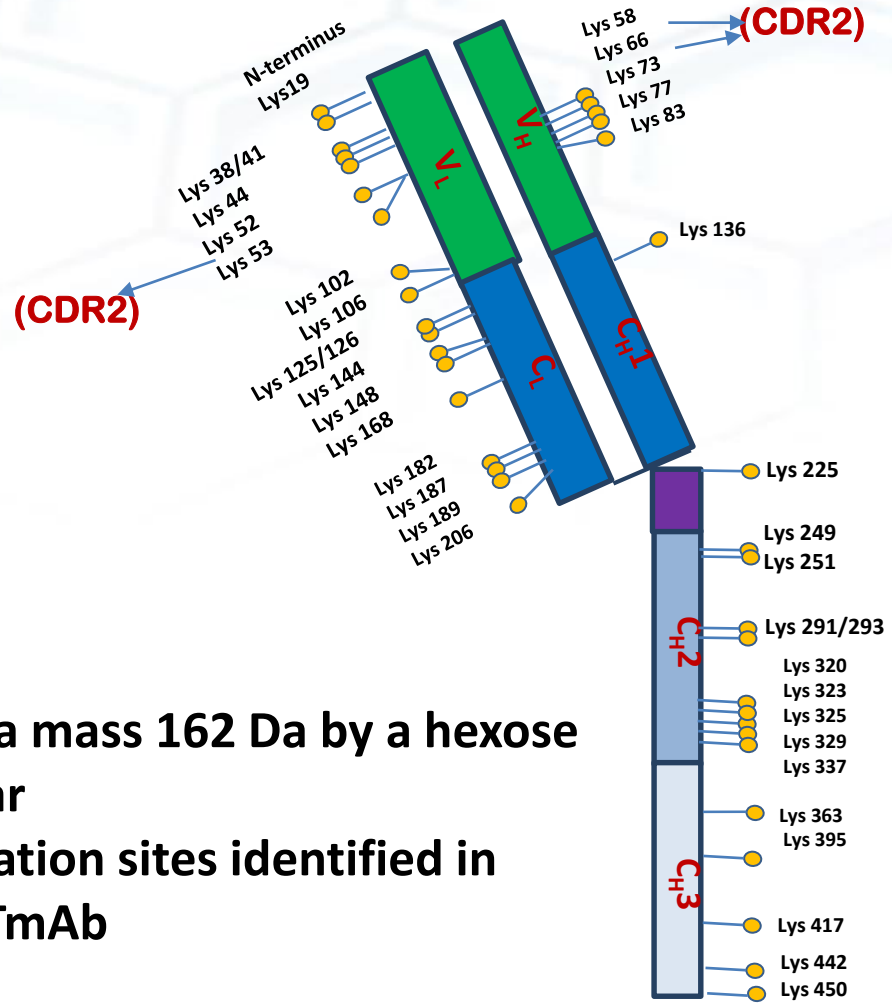
1. Dong, Q. Xinjian Yan, Yuxue Liang, and Stephen E. Stein. *Journal of Proteome Research* 2016 *15* (5), 1472-1486.
2. Prien, J. M. et al. The NISTmAb Case Study. Orthogonal Technologies for NISTmAb N-Glycan Structure Elucidation and Quantitation.

20-Fraction 2D-LC Separation of Glycopeptides from Peptides and Sialylated from Non-Sialylated Species



This greater 2D separation resulted in significant signal intensity enhancement comparing to those in the 1D study, especially for sialylated species.

A Complete Glycation Profile for the NISTmAb Based on 1D and 2D Studies



- **delta mass 162 Da by a hexose sugar**
- **glycation sites identified in NISTmAb**

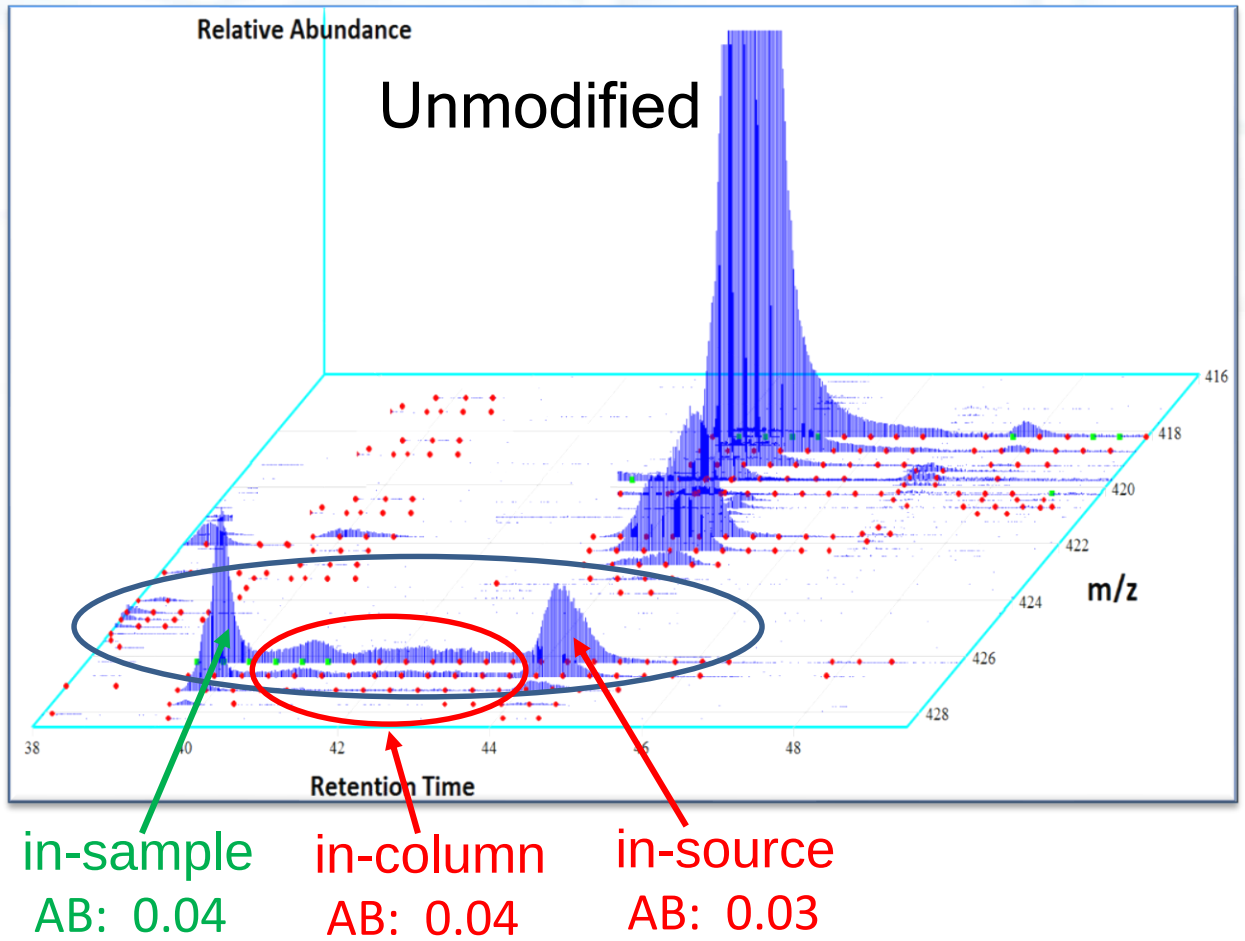
Results

- A total of 590 spectra of 92 glycated peptides identified by 1D and 2D LC-MS/MS analyses
- The glycation profile mapped 21/36 (58%) and 15/15 (100%) of glycation sites in the heavy and light chains of the NISTmAb

Observation of In-Sample, In-Column and In-Source Oxidation In Peptide Mapping Experiments

Selected ion chromatograms of oxidized/unmodified DTLMISR

NISTmAb
2+ DTLMISR
Met 255



Mono-oxidized

m/z 418.221

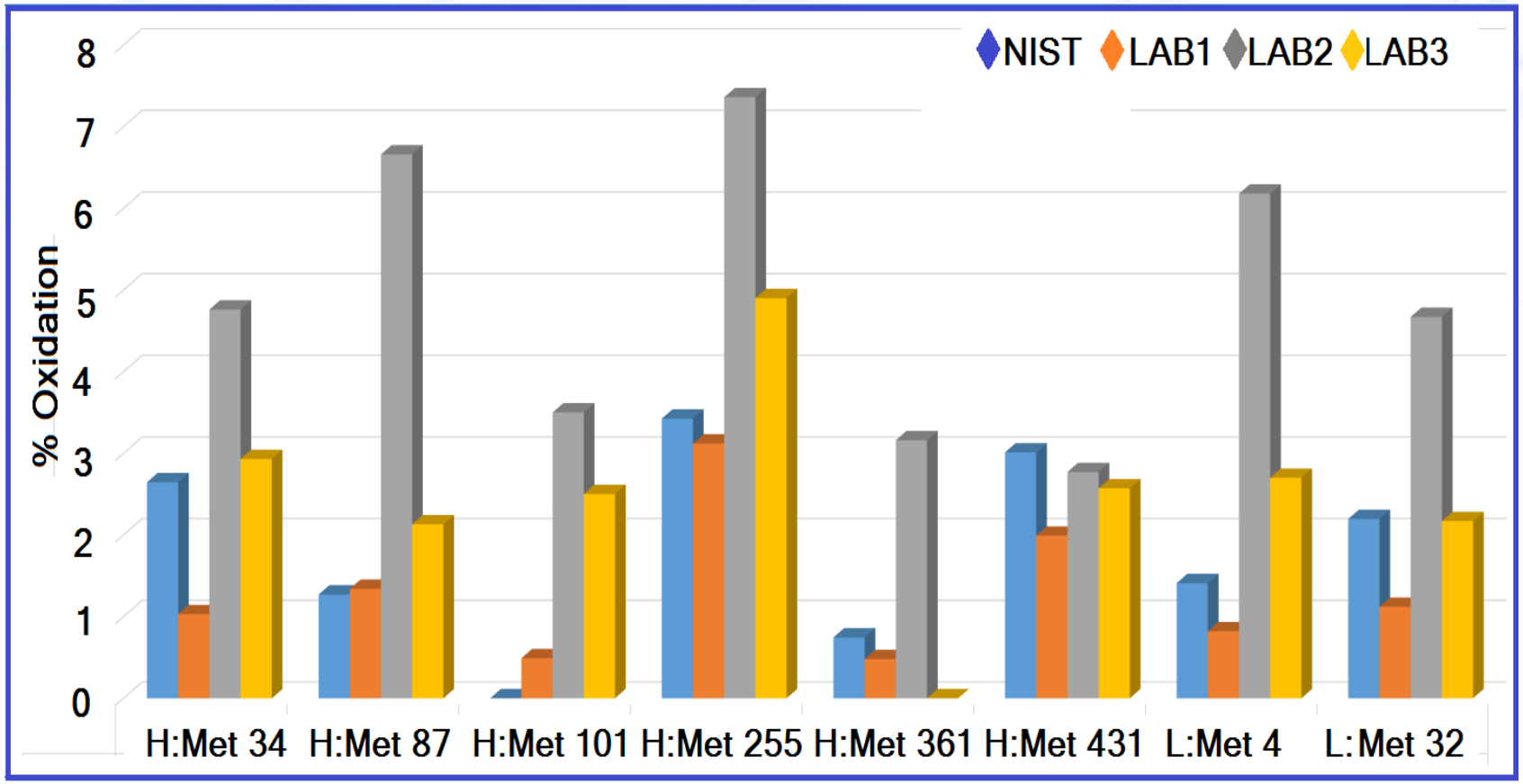
m/z 426.218

in-sample
AB: 0.04

in-column
AB: 0.04

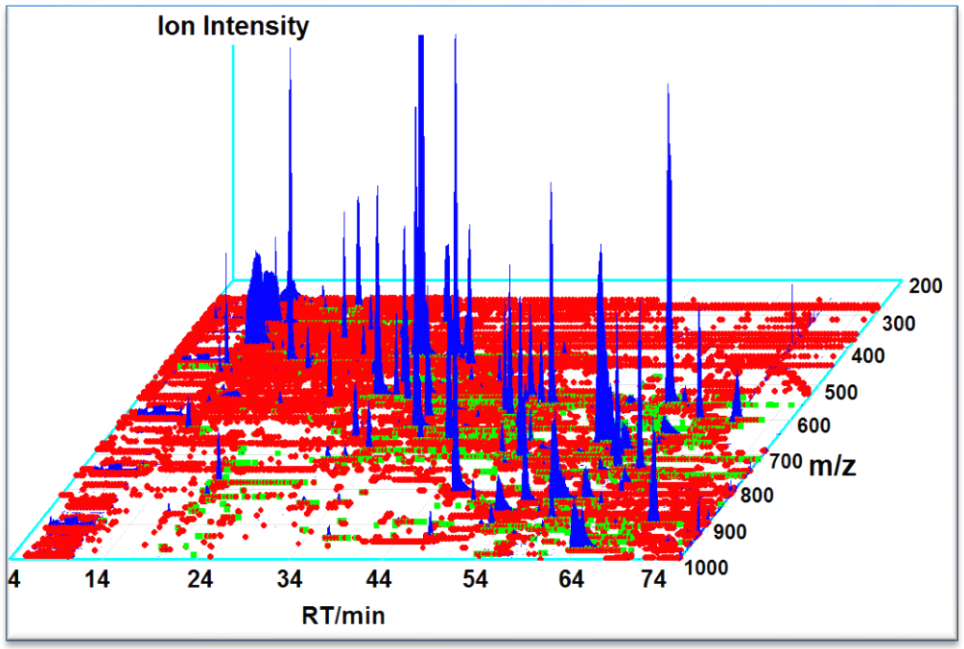
in-source
AB: 0.03

Significant Variations in Quantifying All 8 Methionine Sites in the NISTmAb Obtained From Four Separate Laboratories



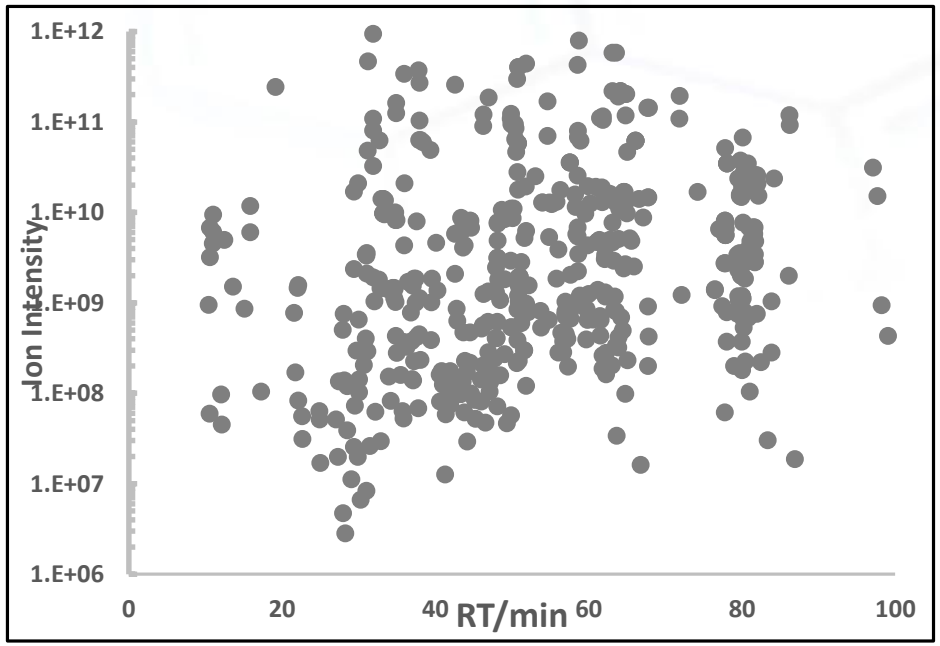
Use of Conventional Database Search Programs to Identify Common Peptides and Modifications

The LC/MS analysis of a **HUMIRA (CHO)** tryptic digest. Humira was digested for 2 h after denaturing in 6 M guanidine at room temperature.

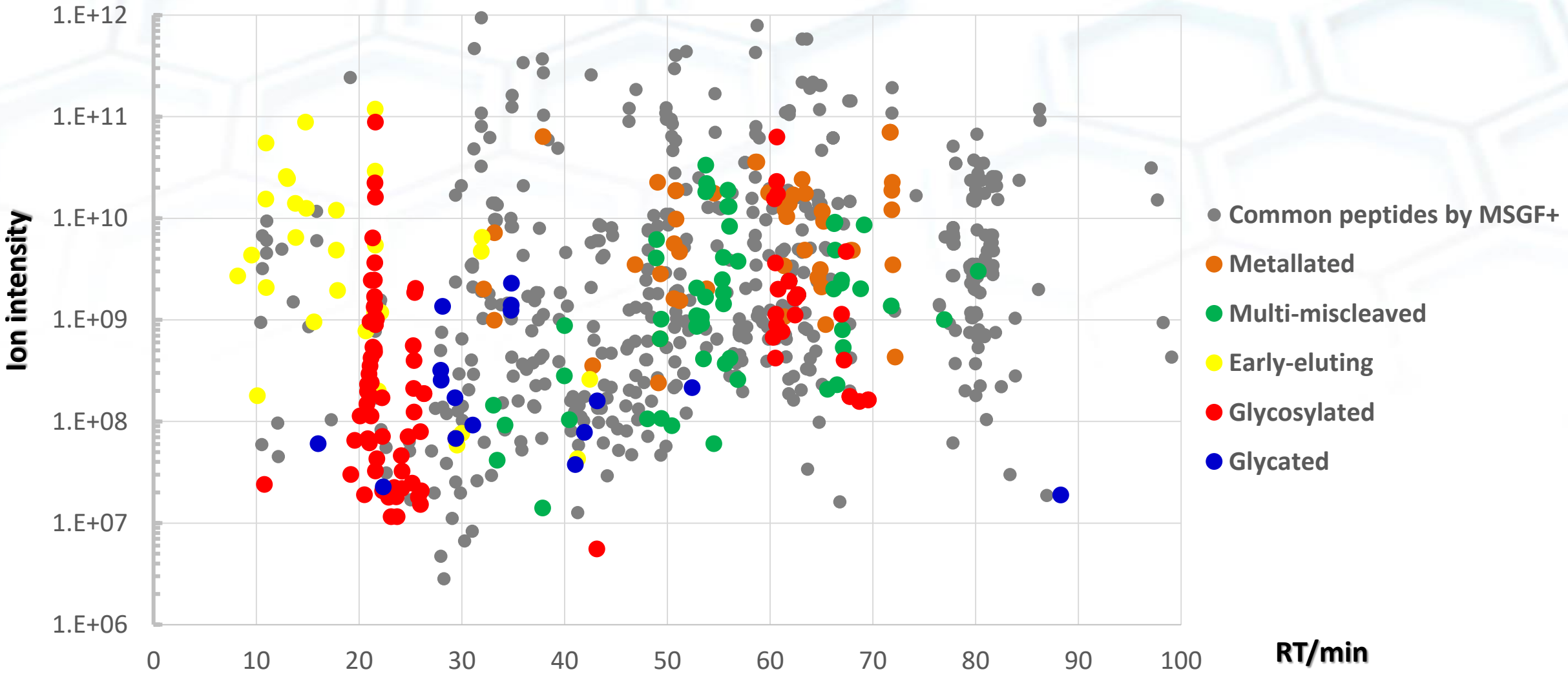


Identify

Peptides and modifications identified in a **HUMIRA** tryptic digest by a search program **MS-GF+**



Use of the NISTmAb Spectral Library to Add a Variety of New and Difficult Peptide Identifications



How to Use NIST Spectral Libraries

- The libraries are provided in two formats:
  - NIST MS library binary format
  - text format (MSP file)

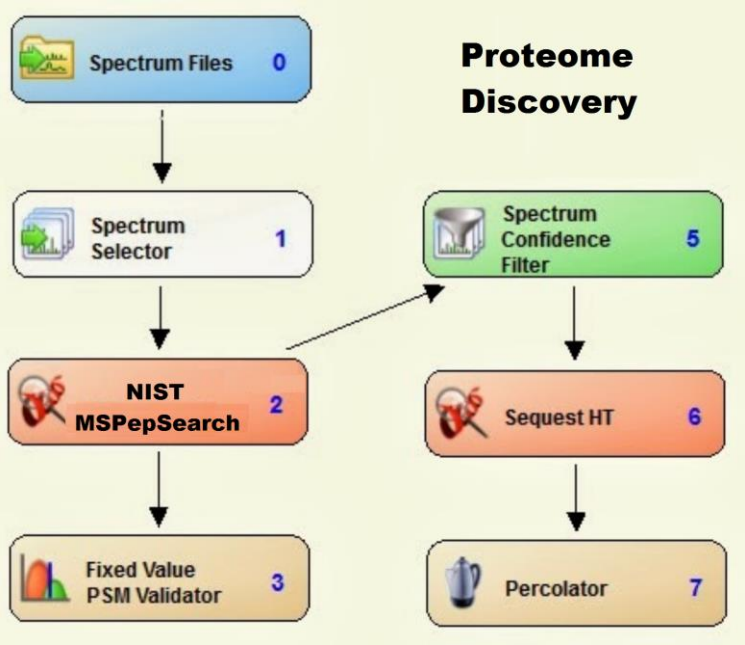
- The NISTmAb libraries of annotated mass spectra are freely available at: https://chemdata.nist.gov/dokuwiki/doku.php?id=peptidew:cdownload

- You can use MSP files with many software programs

- You can use two free software tools, the NIST MS Search and MSPepSearch, to search, browse, and compare the library in binary format

Integrating Library Search into Existing Data-Analysis Workflows

(1) Thermo Proteome Discovery

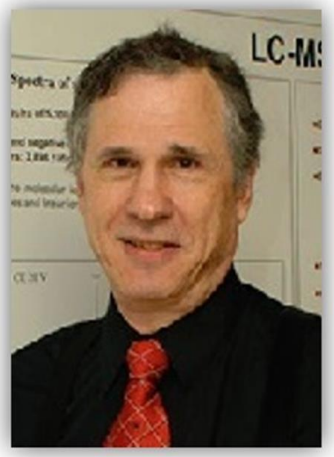


(2) Mascot Database Search

Summary

1. We welcome end-users of pharmaceutical companies and others to take advantage of this NIST reference mAb and its free comprehensive spectral libraries
2. The NISTmAb spectral library is invaluable in biotherapeutic analysis because of its comprehensive and unbiased matches of all classes of peptides and modifications
3. Library search typically exhibits faster, more sensitive and more reliable identification of low abundance peptides, including many hard-to identify modifications
4. Large numbers of modifications identified by the library can be used for monitoring critical quality attributes in MS-based daily routine analysis

NIST National Institute of Standards and Technology U.S. Department of Commerce | Mass Spectrometry Data Center



Stephen Stein

Yuxue Liang

Eric Yan