FUNDAMENTALS OF MASS SPECTROMETRY BASED PROTEOMICS

Ronald Holewinski, Ph.D.

Scientist II

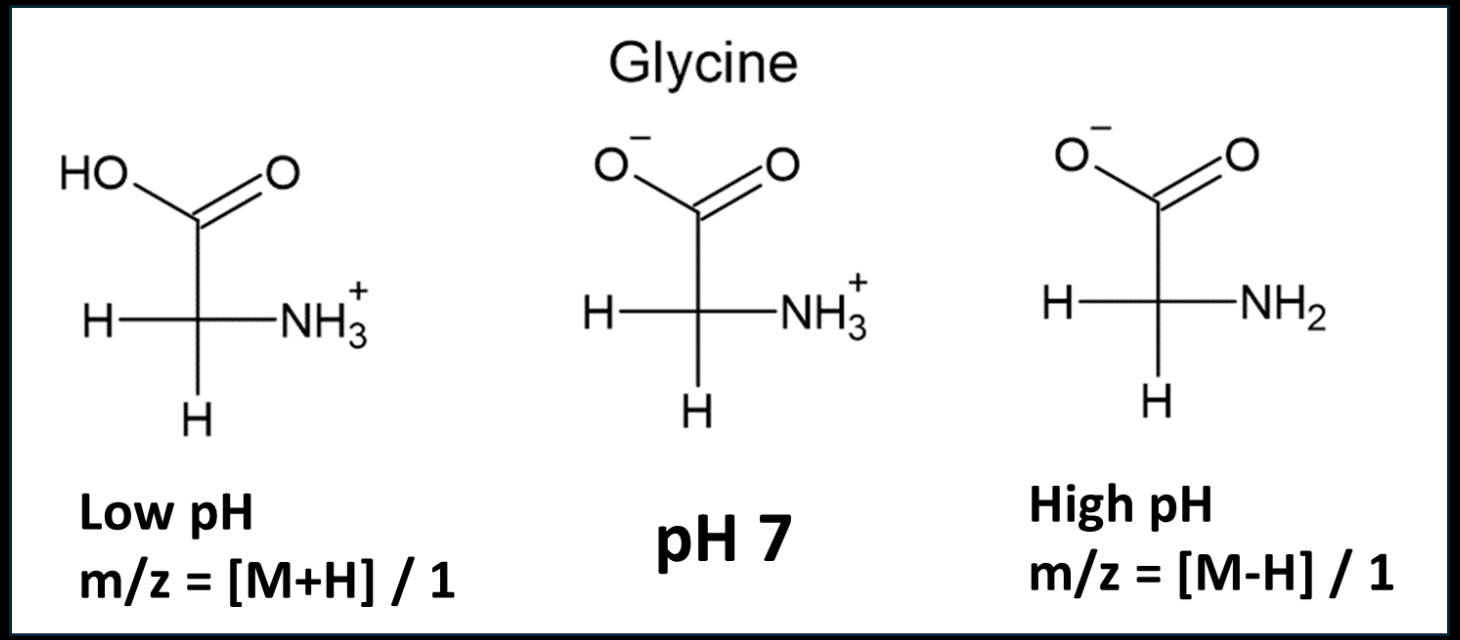
Protein Characterization Lab (NCI/CCR)

OVERVIEW

- What is Mass Spectrometry?
- How do we analyze proteins by MS?
- What are the different instruments we can use?
- What are the different types of Mass Spectrometry acquisition for protein analysis?
  - What type of analysis is right for me?
- How do we identify and quantify proteins?
- How do we identify post-translational modifications?

WHAT IS MASS SPECTROMETRY

- Mass spectrometry is not actually measuring mass of a molecule but rather the mass-to-charge ratio (m/z) of a molecule.
  - m/z = [mass of ion / charge of ion].
  - The mass of the ion is the molecular mass in the ionic state and can differ for the same molecule depending if it carries a positive or negative charge
  - If a molecule cannot be ionized to carry a positive or negative charge, then that molecule cannot be analyzed by MS



TYPES OF MS INSTRUMENTS

- **Orbitraps** – ThermoFisher
  - High Resolution, High Mass Accuracy, Speed
- **Quadrupole Time of Flight (QToF)** – Sciex, Agilent, Bruker, Waters
  - Speed, Sensitivity
- **Triple Quadrupole** – Variety of Vendors
  - Selectivity
- **MALDI-TOF** – Variety of Vendors
  - MS Imaging, Intact mass

WHAT CAN WE ANALYZE BY MS?

- Metabolites
- Lipids
- Drug metabolism (PK)
- DNA/RNA
- Proteins/Peptides

HOW DO WE ANALYZE PROTEINS BY MS?

Intact Protein

- Measure m/z of intact protein
- Protein-Drug conjugates

Top Down

- Measure m/z of intact protein
- Fragment to get protein identification/PTMs
- Provides "proteoform" information
- Proteins < 150kDa

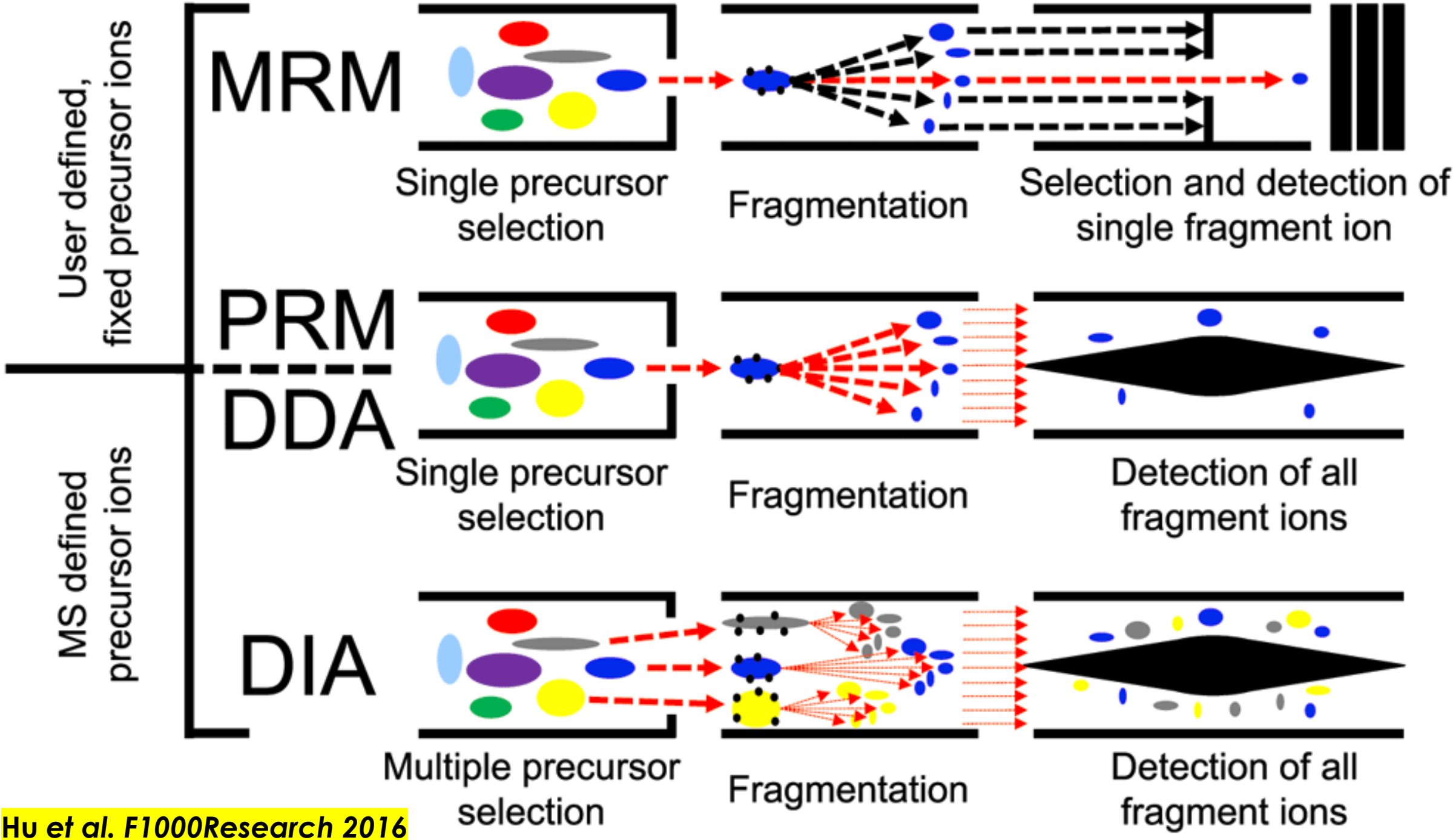
Bottom Up

- measure m/z of digested peptides
- Lose "proteoform" information
- Constrained by protease specificity
- Deepest depth/most versatile

DATA ACQUISITION

- Data-Dependent Acquisition (DDA) – No prior knowledge needed
  - MS1 survey scan to detect peptides within specified m/z range
  - MS2 scan of top abundant ions for peptide sequencing
- Data-Independent Acquisition (DIA) – Library based (*a priori*) or Library-Free
  - MS1 survey scan
  - MS2 scan of wide m/z range to theoretically fragment all ions eluting at a given RT

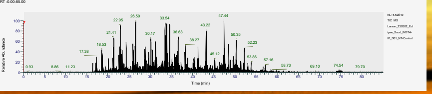
- Parallel Reaction Monitoring (PRM) – Orbitraps, TOFs
  - MS2 scan of *specified* m/z value(s) followed by MS2 scan of **ALL** fragmentation products
- Multiple Reaction Monitoring (MRM) – Quadrapoles
  - MS1 scan of *specified* m/z value(s) followed by MS2 scan of **SINGLE** fragmentation product

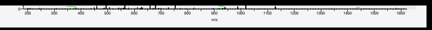


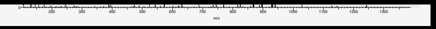
3x Quad

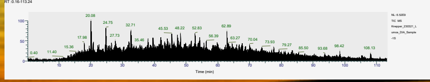
Orbitraps, QToFs

Hu *et al. F1000Research 2016*


RT :0.00-85.00
NL: 5.52E10
TIC MS
Larson_230302_Ecl
ipse_Sood_INST4-
IP_S01_NT-Control
Relative Abundance
Time (min)
0.93
8.86
11.23
17.38
18.53
21.41
22.95
26.59
30.17
33.54
36.63
38.27
43.22
45.12
47.44
50.35
52.23
53.86
57.16
58.73
69.10
74.54
79.70


m/z


m/z

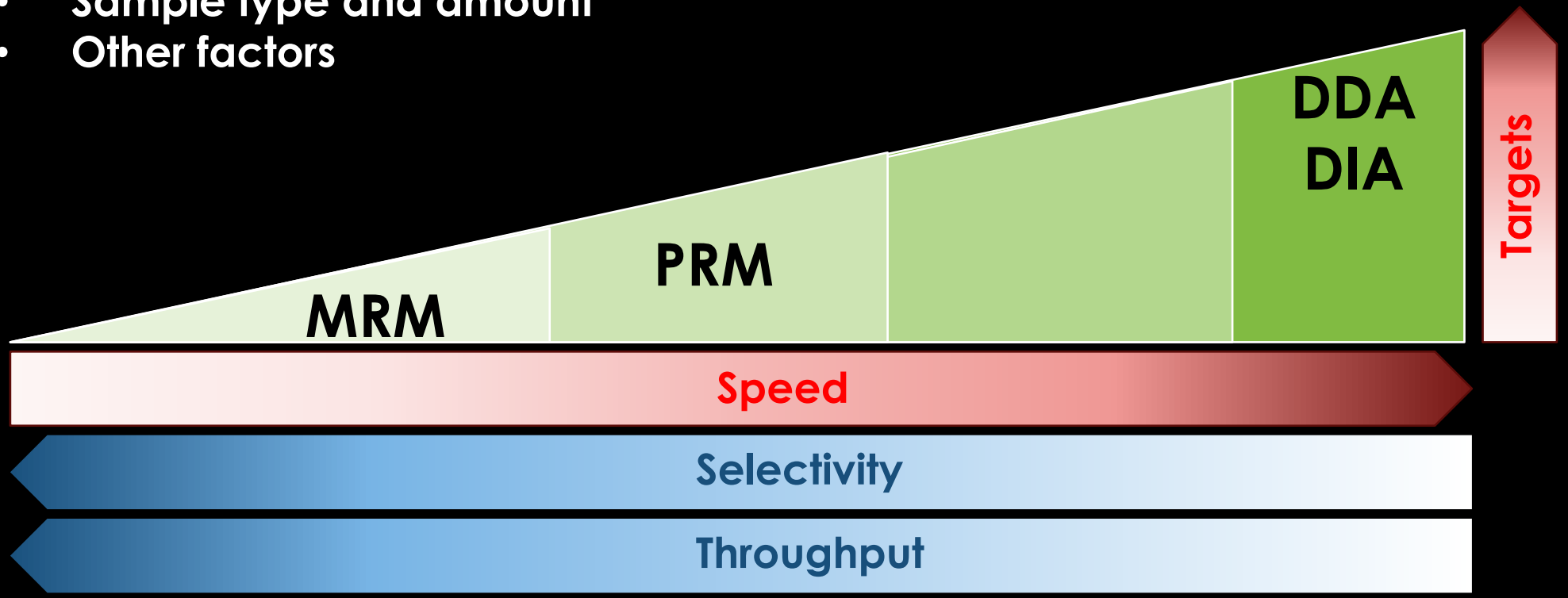

RT :0.16-113.24
NL: 6.52E9
TIC MS
Knepper_230321_L
umos_DIA_Sample
-1S
Time (min)
20.08
24.75
32.71
62.89
45.53
48.22
52.83
17.98
27.73
35.46
56.39
63.27
70.04
73.93
79.27
85.50
93.68
98.42
108.13
15.36
11.40
0.40

DATA ACQUISITION

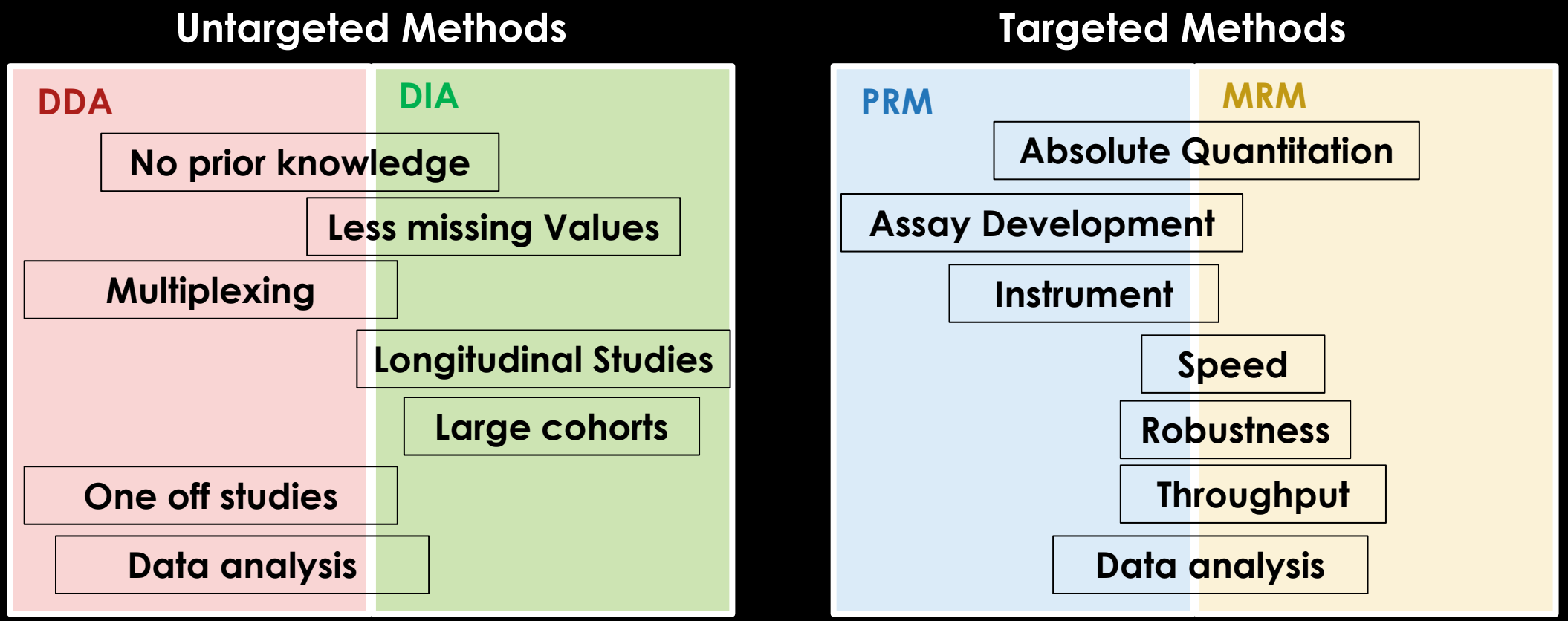
Q: What type of analysis do I need?

A: It depends…..

- on what you want to accomplish
- Sample type and amount
- Other factors



CHOOSING BETWEEN METHODS



Sample → *Tissue, Cells, IP, Media, Biofluids*

Digestion ← *Trypsin, LysC, Chymotrypsin, GluC*

Digestion → TMTpro

TMTpro → PTM enrichment

PTM enrichment → LC/MS/MS Acquisition

TMTpro → LC/MS/MS Acquisition

LC/MS/MS Acquisition → Database Search/Quan

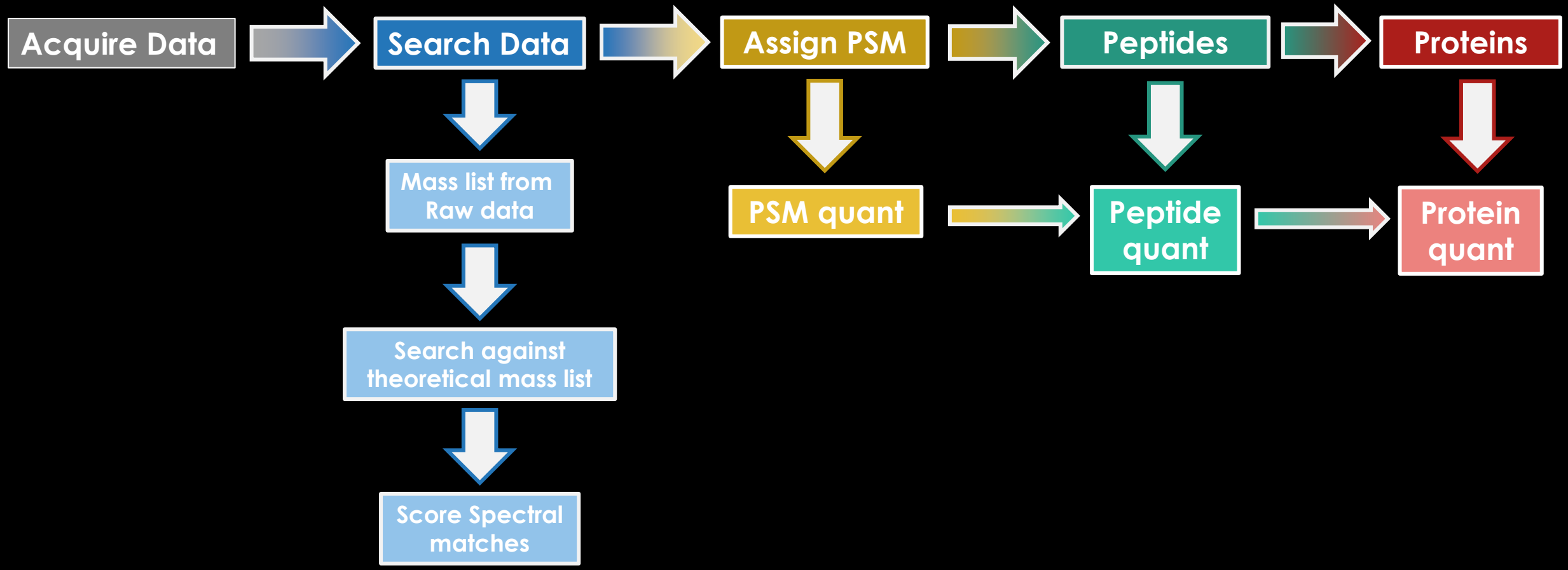
Database Search/Quan → *Sequest, Maxquant, Mascot* (DDA)

Database Search/Quan → *Skyline, Vendor Specific* (PRM/MRM)

Database Search/Quan → Data Analysis

Data Analysis → *Quantitation/Stats, Volcano plots, Pathway analysis*

IDENTIFYING PROTEINS BY MS



IDENTIFYING PTMS BY MS

What PTMs have the most publications in PubMed?

modification	modification motif	amino acid modification site	number of publications (2017–2022)
phosphorylation	HPO_3	Y, S, T, H, D	68208
methylation	$(-CH_3)$	K	21750
glycosylation	(N-Glc-NAc)	N-X-S/T, X ≠ P	10660
ubiquitination	ubiquitin	K	9505
acetylation	CH_3CO	S, K	8590
glycation	$(\text{-Hex})_n$	K, R	5331
disulfide bond	S–S	C	4580
SUMOylation	small ubiquitin-like modifier	protein chain	1919
hydroxylation	(−OH)	side chain	1458
S-nitrosylation	(−NO)	cysteine thiol group	1395
lipidation	lipid/fatty acid	protein chain	631
succinylation	$CO–CH_2–CH_2–CO_2H$	K	584
glutathionylation	GSH	cysteine thiol group	561

Bobalova et. al. *J. Agric. Food Chem.* 2023

Challenges

- Low stoichiometry to unmodified peptides
  - Enrichment often necessary but not available for most PTMs
- Reduced Ionization
- Fragmentation issues
- Constrained by sequence
- Changes in hydrophobicity
- Search ambiguity

SUMMARY/MAIN POINTS

1. Proteomic analysis by MS can be done by a variety of techniques

2. MS has a lot of different applications in proteomic analysis

3. The type of analysis is dependent upon the experiment

QUIZ

A TRIPLE QUADRUPOLE IS USED TO PERFORM WHICH ANALYSIS?

A. DDA

B. DIA

C. MRM

D. PRM

WHAT IS THE MOST COMMON PROTEASE USED FOR BOTTOM-UP PROTEOMICS?

A. GluC

B. Trypsin

C. LysC

D. Chymotrypsin

WHICH OF THE FOLLOWING DOES PRM HAVE AN ADVANTAGE OVER MRM?

A. Throughput
B. Speed
C. Selectivity
D. Assay Development

THANK YOU