NIST
National Institute of Standards and Technology
U.S. Department of Commerce

New Features of the 2017 NIST *Tandem* Mass Spectral Library

Xiaoyu Yang, Ph.D.

Mass Spectrometry Data Center

Biomolecular Measurement Division Seminar, June 27, 2017

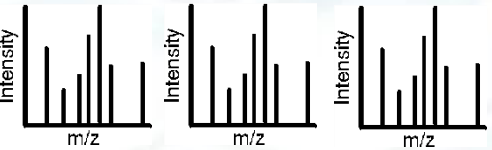
NEWS! New Release of the NIST Mass Spectral Libraries and Software

65th ASMS Conference on Mass Spectrometry and Allied Topics (American Society of Mass Spectrometry)

June 4 - 8, 2017
Indianapolis, IN

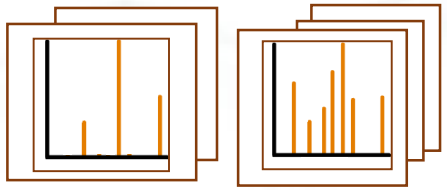
- **On June 6, 2017, the 2017 NIST Tandem Mass Spectral Library was released.**

Electrospray Ionization Tandem Mass Spectral Library



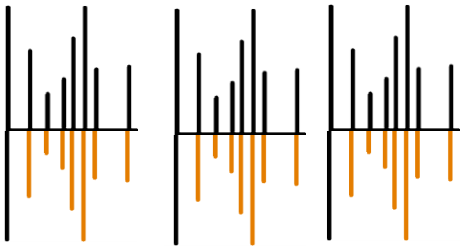
Experimental tandem mass spectrum

Searching



Reference spectra in the library

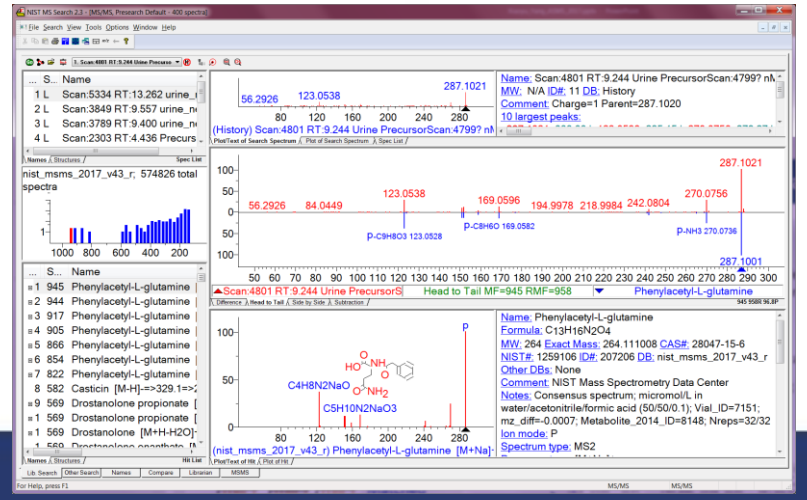
Matching



Experimental
Matched
Reference

MS Search 2.3

Batch mode searching



- ***What?***

  A collection of tandem mass spectra

- ***Why?***

  Mass spectral library searching is a fast, accurate and reliable compound identification technique for LC/MS/MS data.

- ***Why NIST Library?***

❖ High quality
❖ Comprehensive
❖ Reference

I	J	K
Metabolite	Formula	Prec.Type
Diethyl succinate	C8H14O4	[M+H-C4H10O]+
L-.beta.-Homoserine	C4H9NO3	[M+H-H2O]+
7-Hydroxychromanone	C9H8O3	[M+H]+
3,4-Dihydroxy-L-phenyl	C9H11NO4	[M+H-NH3]+
5-Aminolevulinic acid	C5H9NO3	[M+H]+
Bestatin	C16H24N2O4	[M+H-C8H15O4N]+
Benzeneethanamine	C8H11N	[M+H]+
(.+/-.)-SKF 81297	C16H16ClNO2	[M+H-C8H7ClO2]+
25-Hydroxycholesterol	C27H46O2	[M+H-H4O2]+
25-Hydroxycholesterol	C27H46O2	[M+H-H4O2]+

Procedure of Extending the Tandem Mass Spectral Library

MS^1 and MS^2 resolution: 30,000

Authentic compounds

Ion Trap (IT, IT-FT)
Collision Cell (HCD)

↓ Analyzed on Orbitrap Elite

Spectra of MS^2, MS^3, MS^4, isotopic precursors, in-source fragments

Density-based clustering

↓ Identify precursors of MS^2 spectra

MS^2 spectra with the same precursor ion

Top-down hierarchical clustering

↓ Cluster spectra of MS^2, MS^3, MS^4

Consensus spectra

Precursor accuracy
Purity
Peak intensity
Fragmentation

Quality control

↓ Manual inspection (expert evaluation)

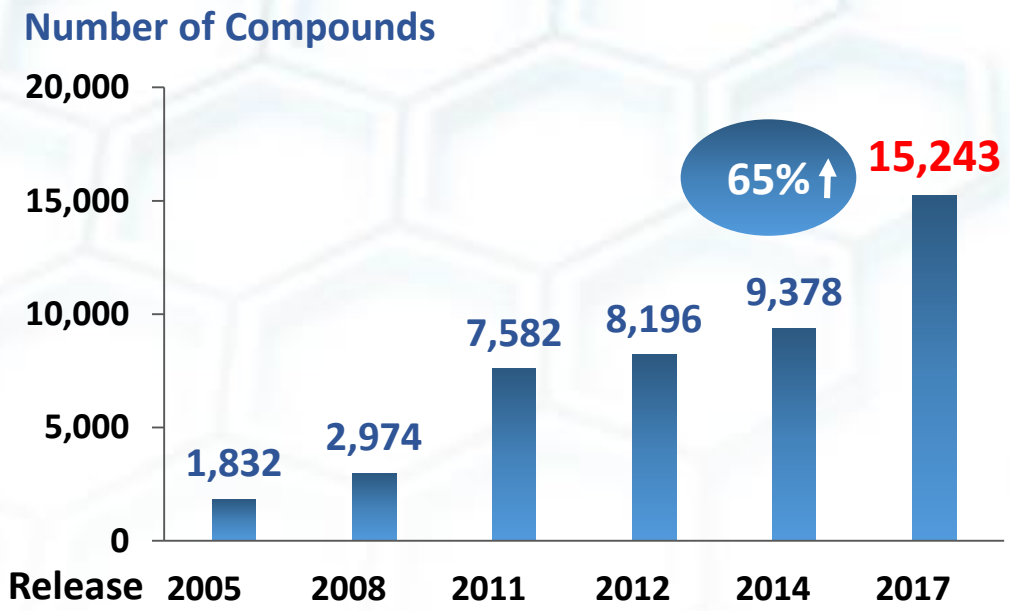
Tandem mass spectral library

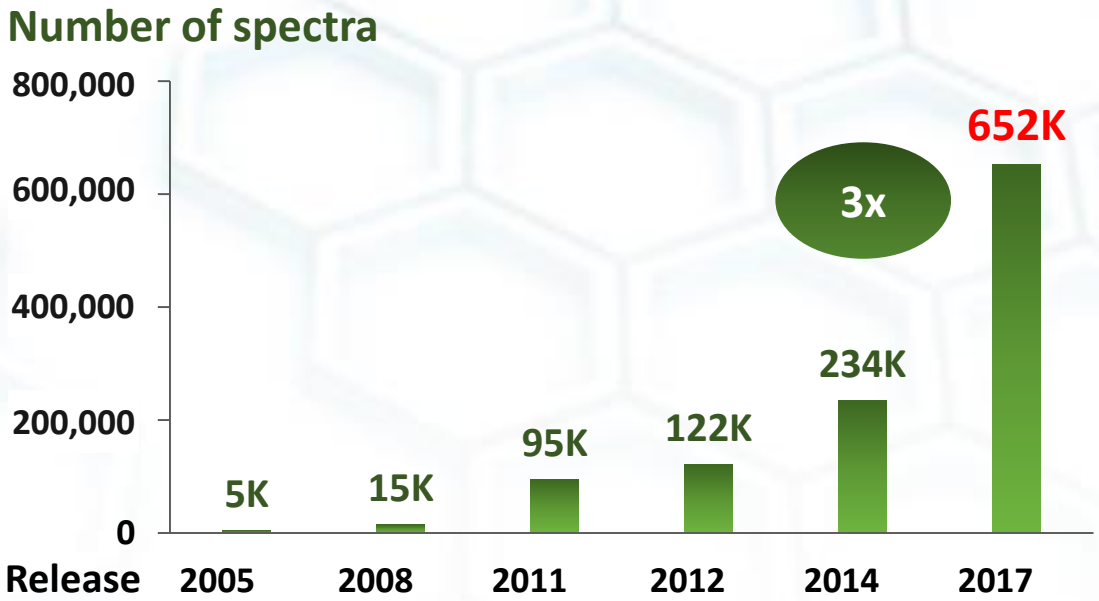
Dr. Pedatsur Neta

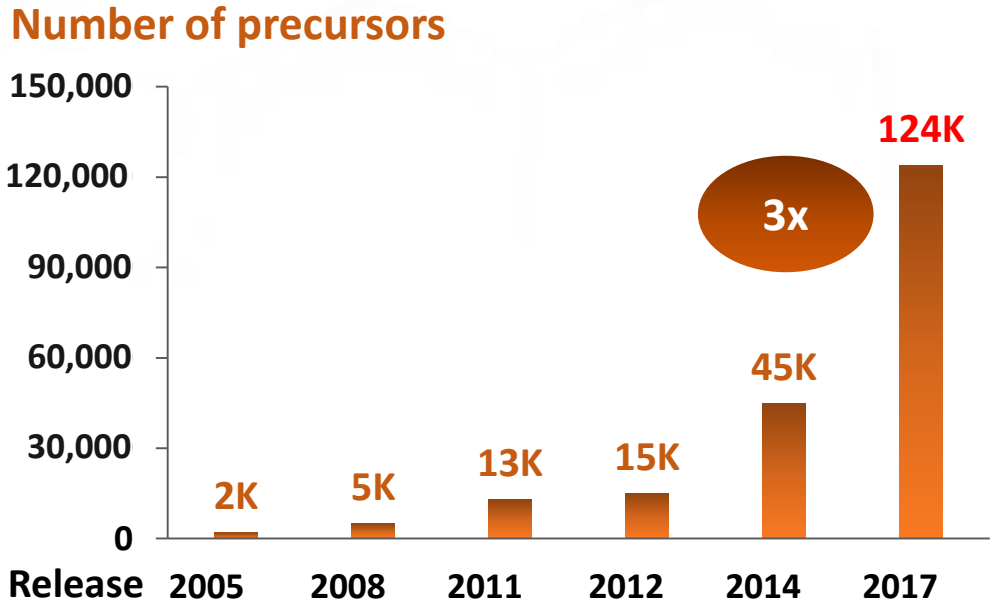
X. Yang, P. Neta, S. Stein. Extending A Tandem Mass Spectral Library to Include MS^2 Spectra of Fragment Ions Produced In-Source and MS^n Spectra. *J. Am. Soc. Mass Spectrom. (Just Accepted)*

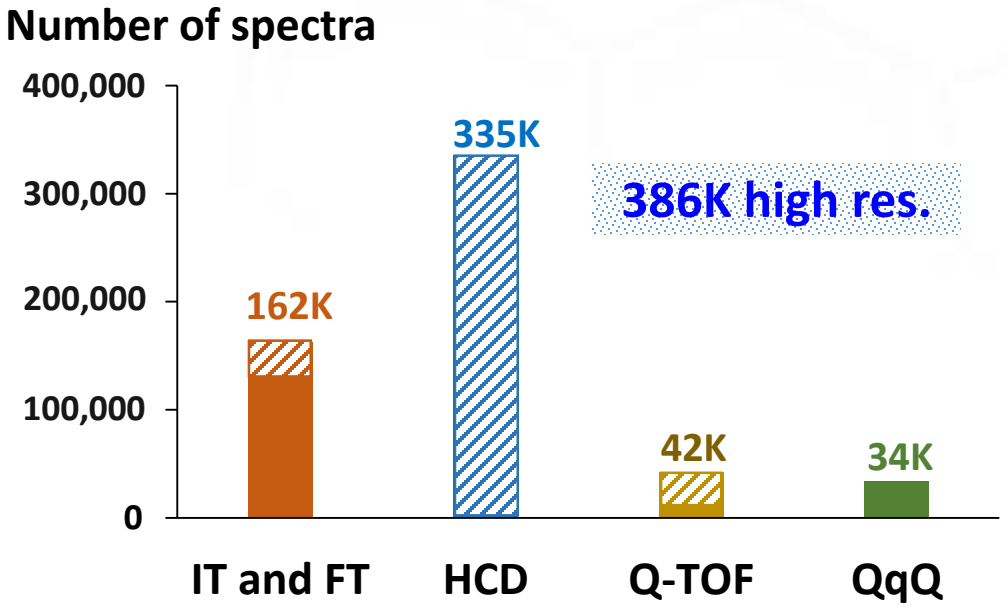
X. Yang, P. Neta, S. Stein. Quality Control for Building Libraries from Electrospray Ionization Tandem Mass Spectra. *Anal. Chem.*, 2014, 86: 6393-6400.

Dramatic Increase: Number of Compounds and Tandem Mass Spectra









New Precursor Ions and Mass Spectra

13,045

Positive Precursors:

$[M+H]^+$, $[M+2H]^{2+}$,
$[M+H-H_2O]^+$, $[M+H-NH_3]^+$, $[M+H-OH]^+$,
$[M+H+H_2O]^+$, $[M+NH_4]^+$

$[2M+H]^+$, $[3M+H]^+$

$[M+Na]^+$, $[M-H+2Na]^+$, $[M-2H+3Na]^+$,
$[M+K]^+$, $[M-H+2K]^+$, $[M-2H+3K]^+$,
$[M+Li]^+$, $[M-H+2Li]^+$, $[M-2H+3Li]^+$

6,001

Negative Precursors:

$[M-H]^-$, $[M-2H]^{2-}$,
$[M-H-H_2O]^-$, $[M-H-NH_3]^-$,
$[M-H+H_2O]^-$, $[M-H+NH_3]^-$

$[2M-H]^-$, $[3M-H]^-$

plus

2017 New Library

555

Isotopic Precursors with Cl, Br or Sn

9,590

MS^n

MS^3 and MS^4 spectra of the most intense peaks in the MS^2 and MS^3 spectra, respectively.

7,675

In-source Fragments: $[M+H\text{-neutral}]^+$

$[M\text{-}H\text{-neutral}]^-$

Fragments from the original target compound within the ESI source

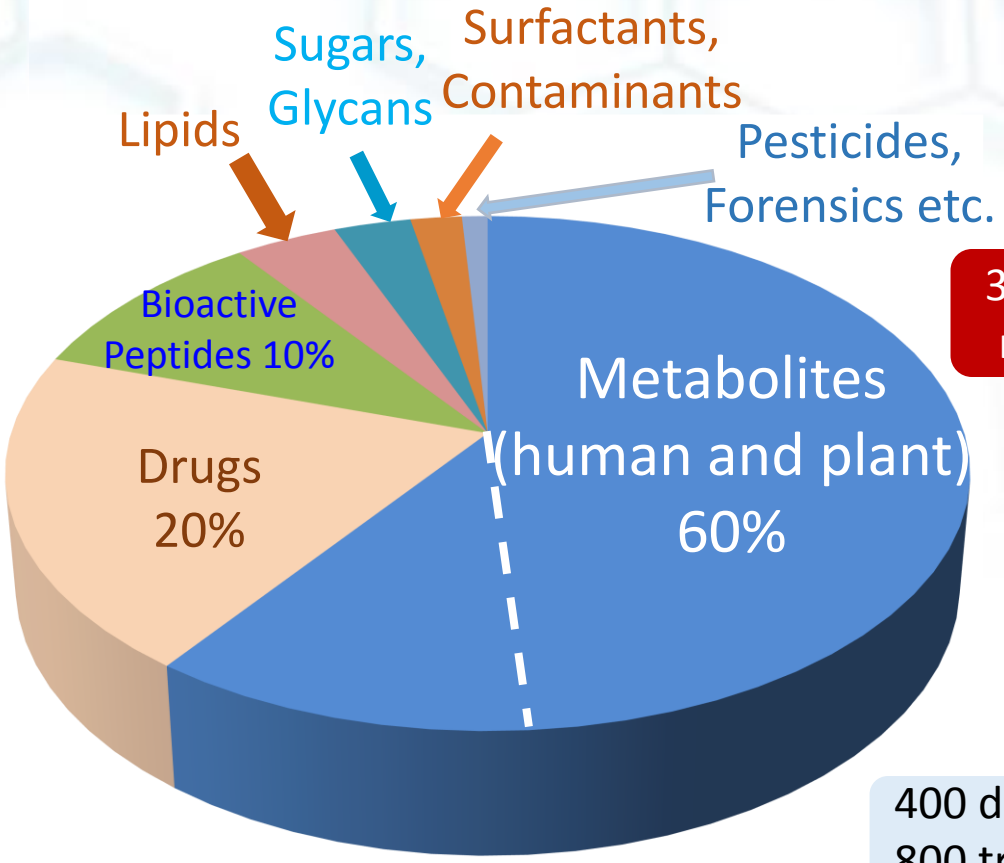
2017 NIST Tandem Mass Spectral Library

60%

3x

15,243 compounds
123,881 precursor ions
574,826 spectra

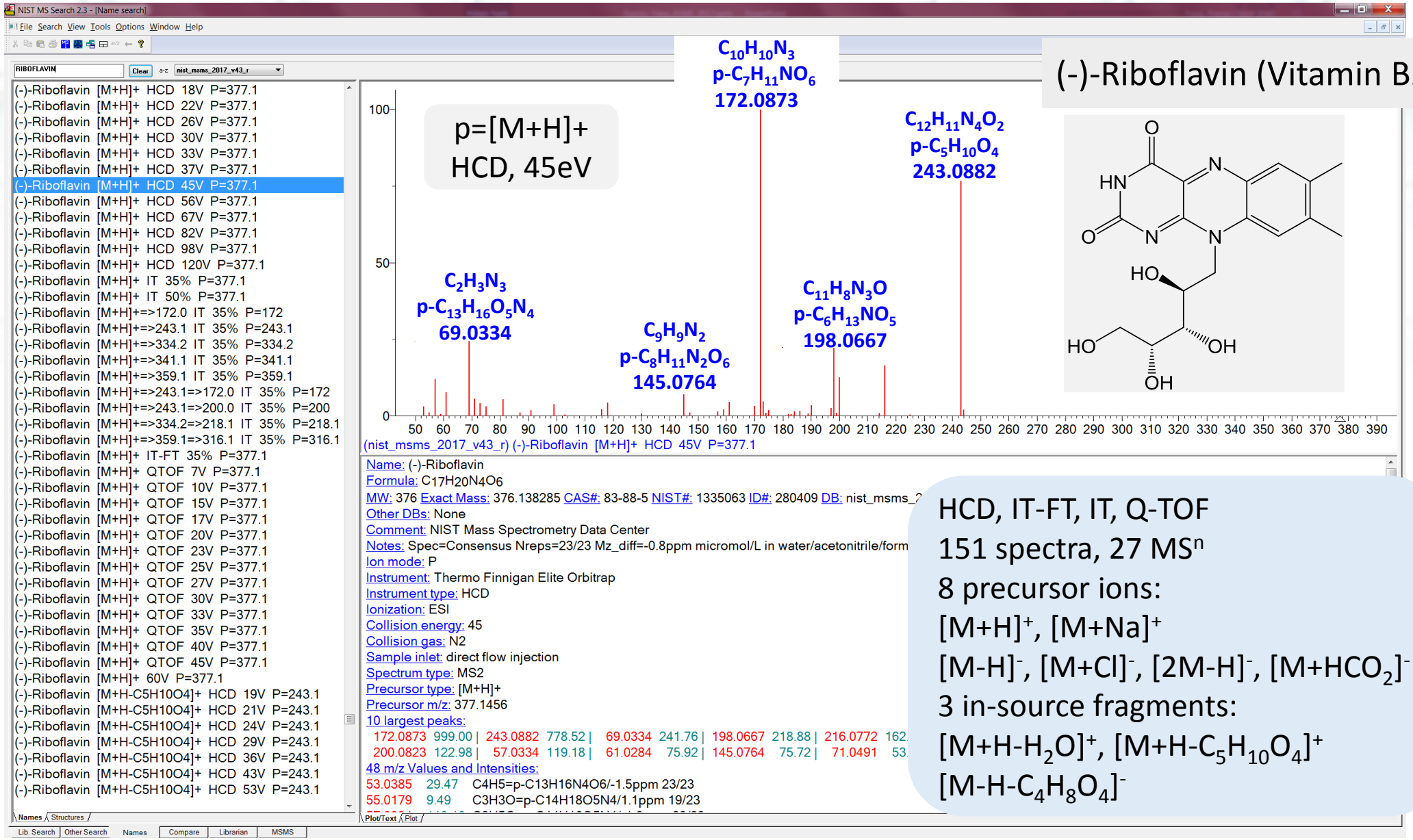
80% (+), 20% (-)
27% MS^2 of in-source
17% MS^n



3,500 human metabolites

400 dipeptides,
800 tryptic tripeptides

Mass Spectra of Different Precursor Ions and Instruments



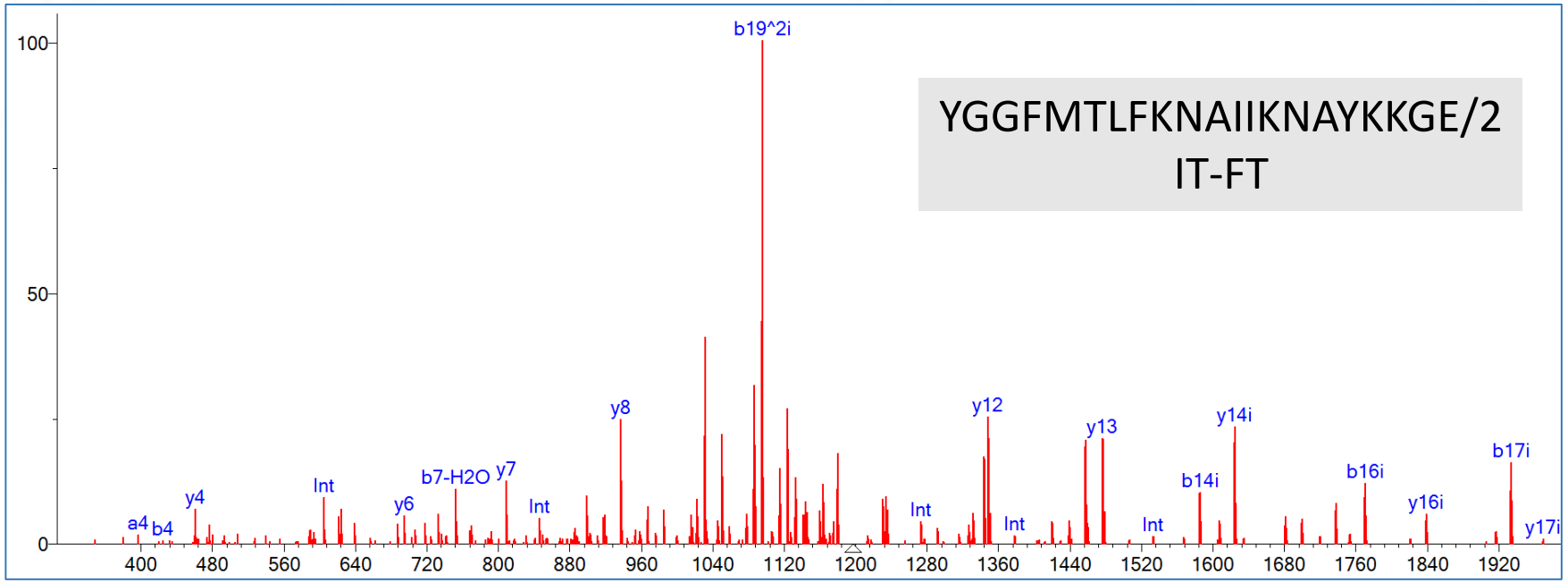
(-)-Riboflavin (Vitamin B2)

HCD, IT-FT, IT, Q-TOF
151 spectra, 27 MS^n
8 precursor ions:
$[M+H]^+$, $[M+Na]^+$
$[M-H]^-$, $[M+Cl]^-$, $[2M-H]^-$, $[M+HCO_2]^-$
3 in-source fragments:
$[M+H-H_2O]^+$, $[M+H-C_5H_{10}O_4]^+$
$[M-H-C_4H_8O_4]^-$

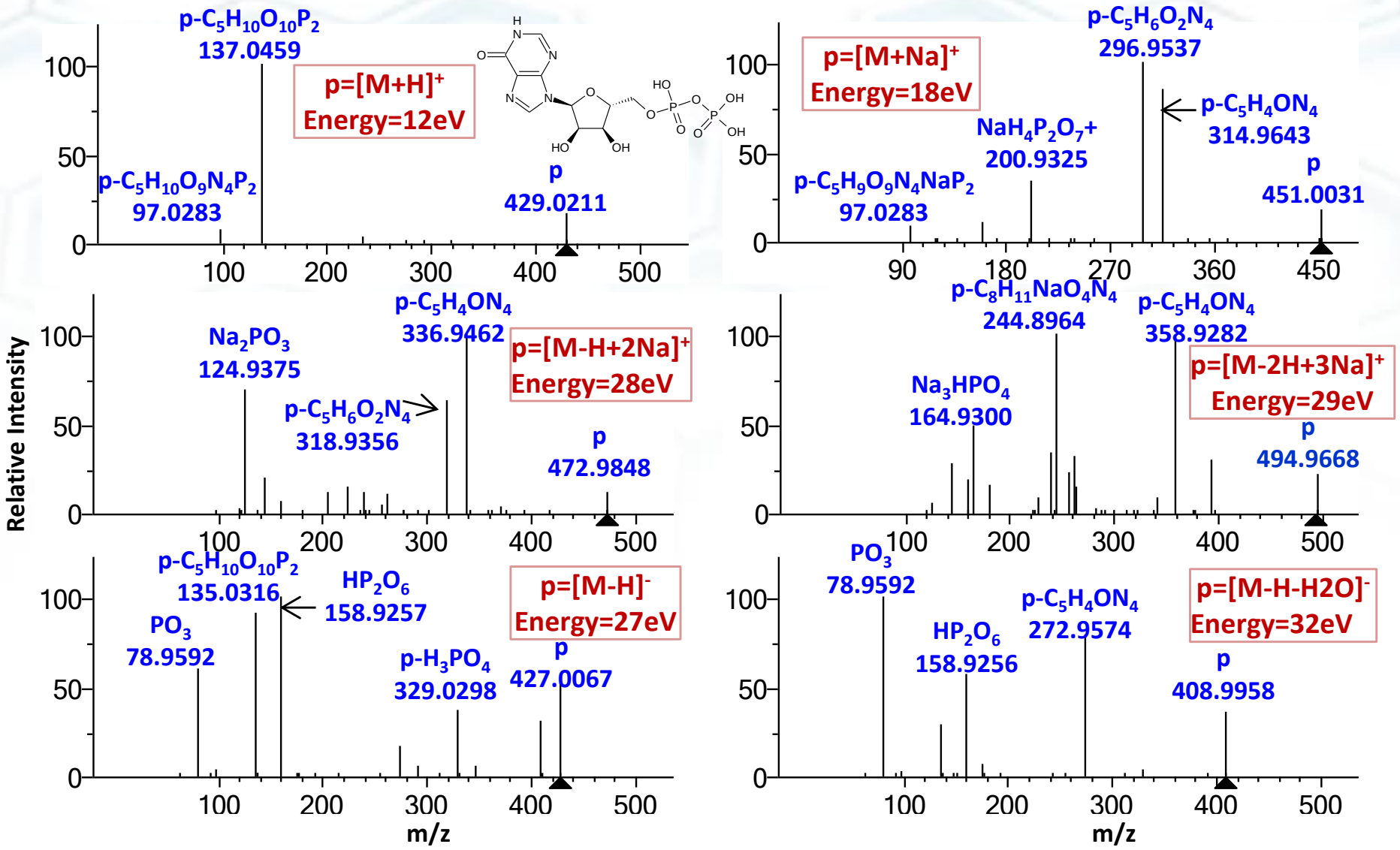
Annotating Product Ions of Peptide Mass Spectra

Software for free download (.dll and .exe):

- Essential software for annotating millions of spectra for the NIST Peptide Mass Spectral Library;
- IgG, HSA, TMT, and ITRAQ spectra;
- p, y, b, a, c, internal fragments;
- 104 neutral losses (35 from modifications); 39 fragments from modifications; 114 immonia; 955 modifications; internal fragments associated with proline.

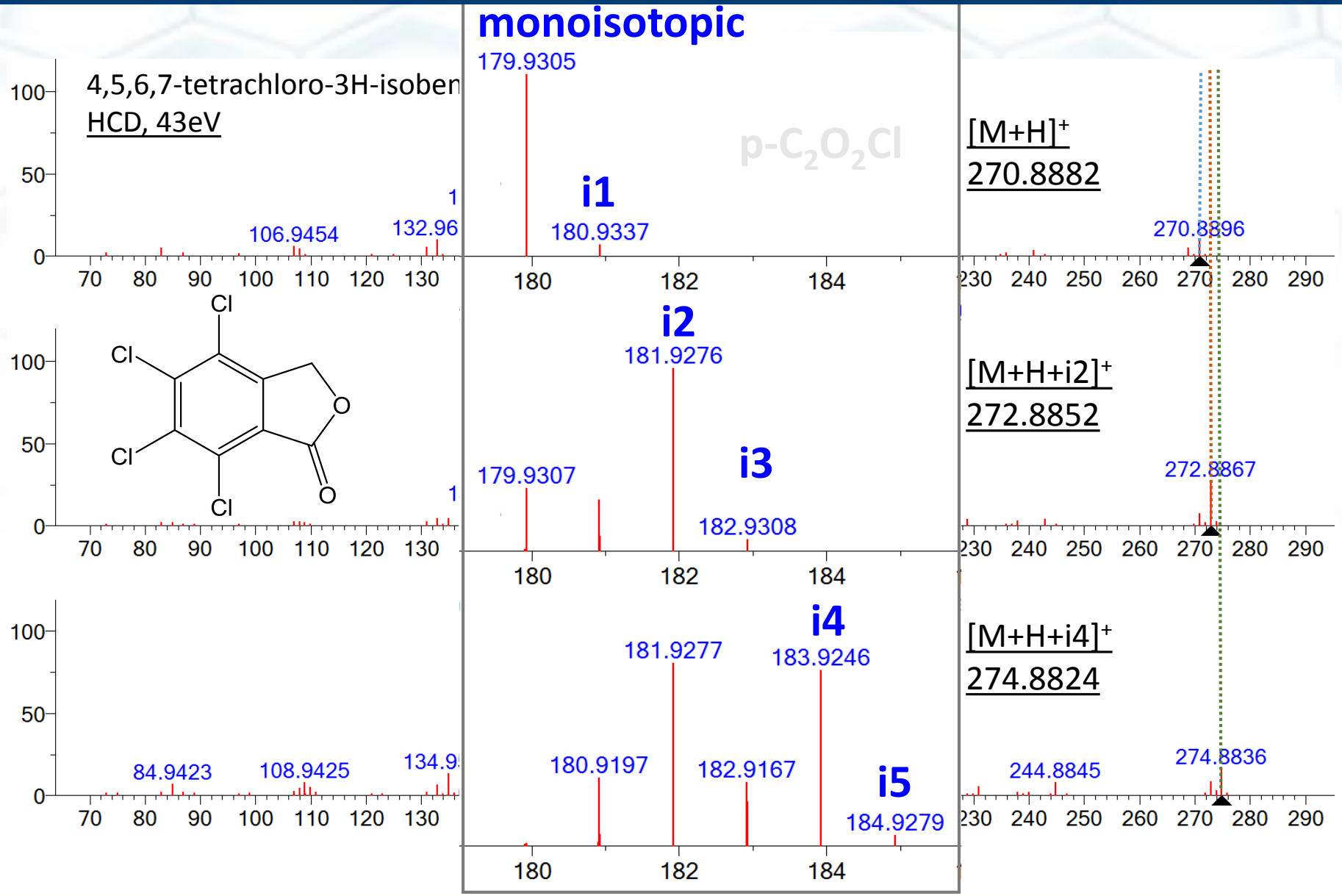


Spectra of Multiple Precursor Ions for More Robust Compound Identification



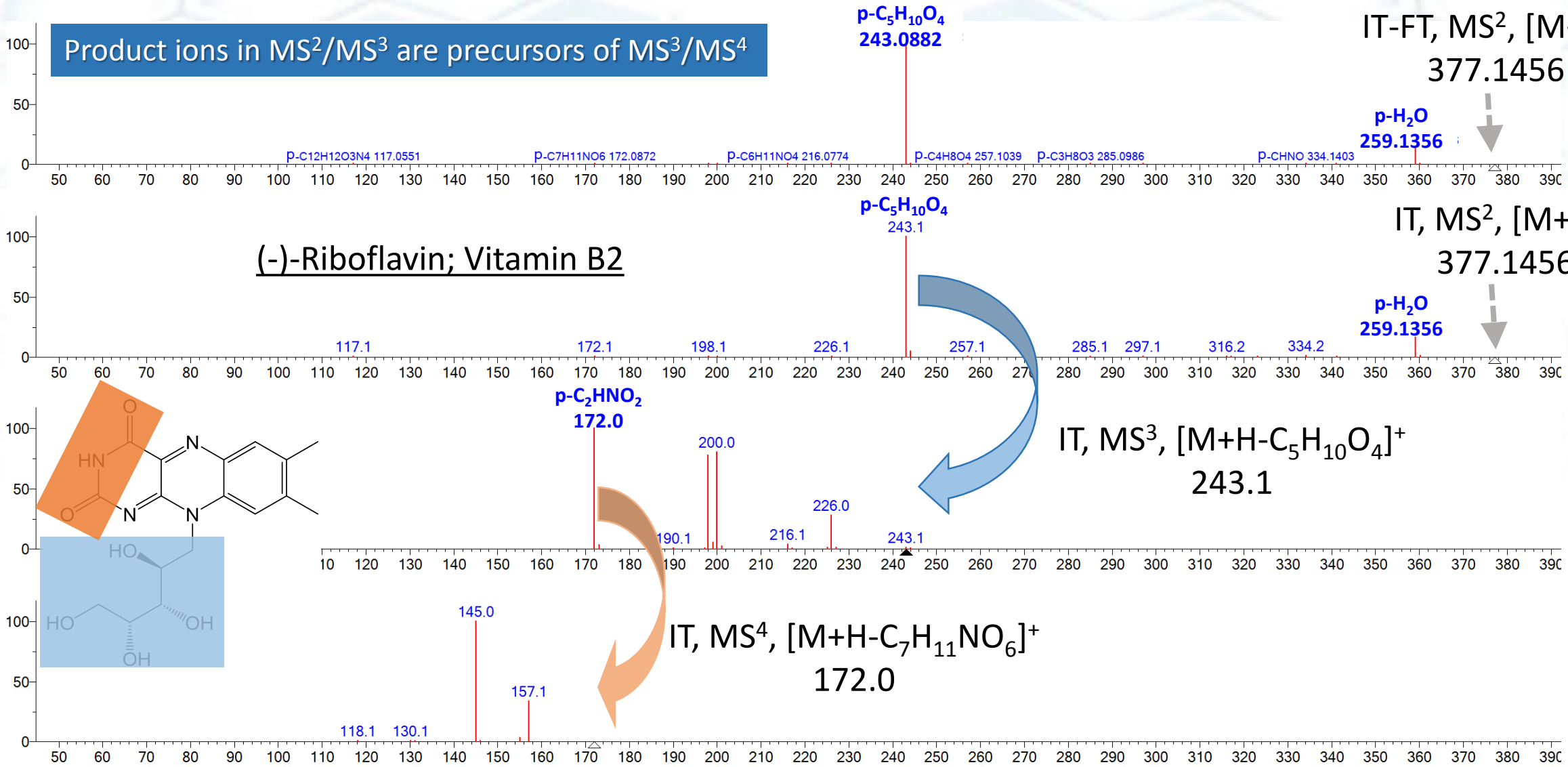
Inosine 5'-diphosphate acquired on HCD

MS² Spectra with Isotopic Precursors



MS^n for Better Identifying Compounds

Product ions in MS^2/MS^3 are precursors of MS^3/MS^4



IT-FT, MS^2, $[M+H]^+$
377.1456

IT, MS^2, $[M+H]^+$
377.1456

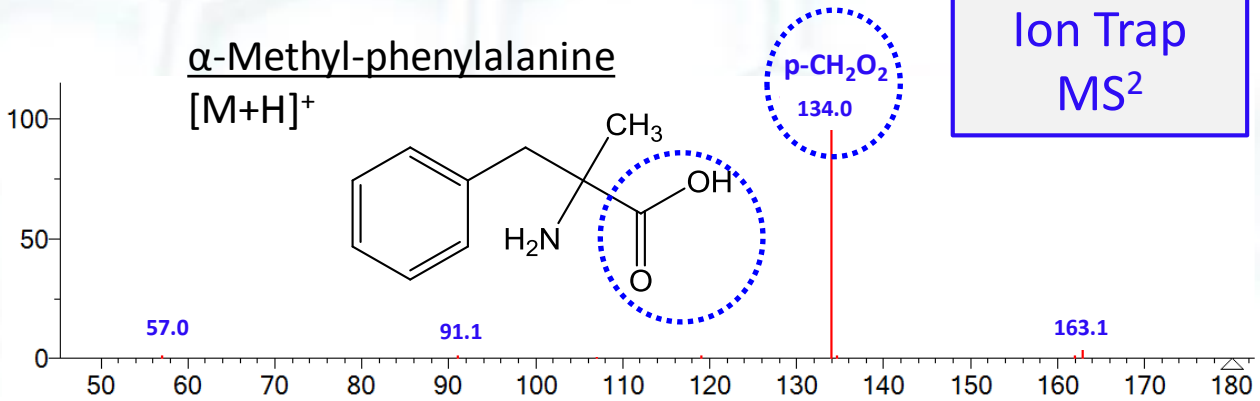
(-)-Riboflavin; Vitamin B2

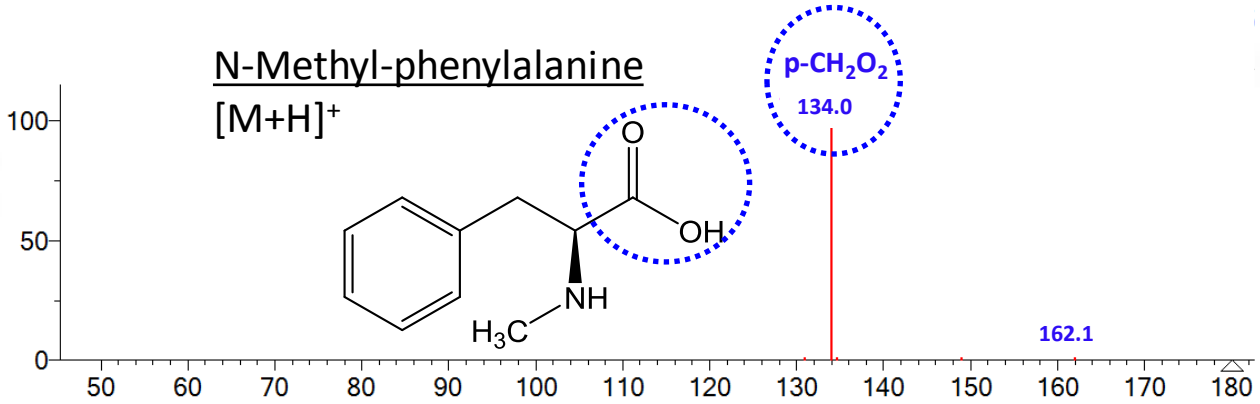
IT, MS^3, $[M+H-C_5H_{10}O_4]^+$
243.1

IT, MS^4, $[M+H-C_7H_{11}NO_6]^+$
172.0

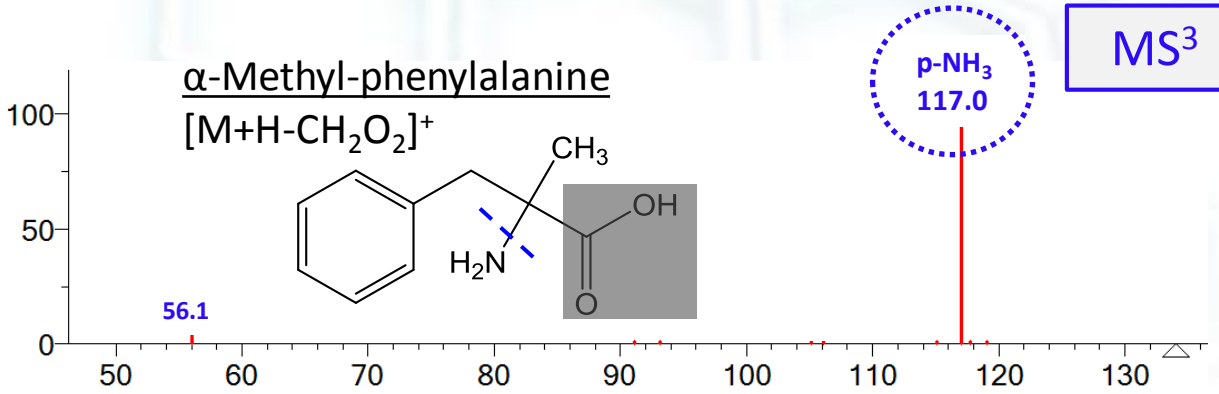
Ion Trap MS^n for Distinguishing Isomers

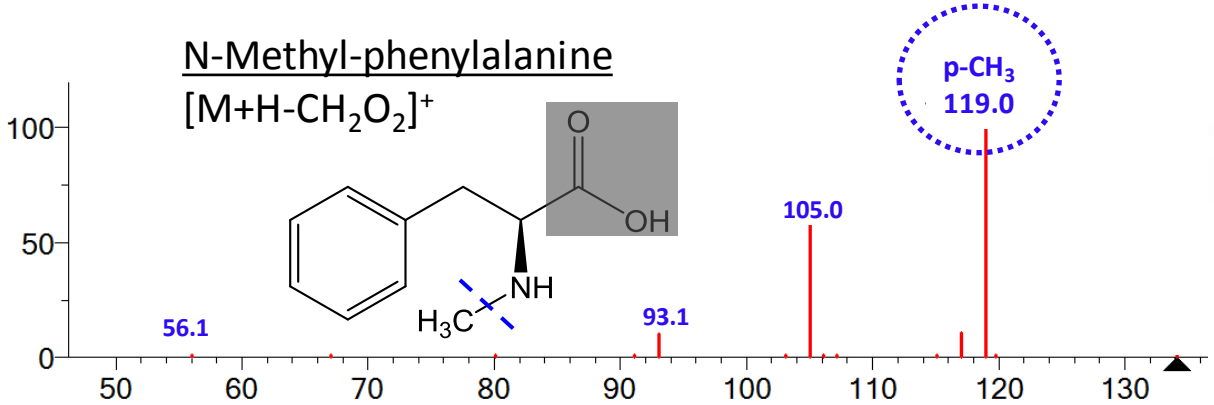
Ion Trap MS^2





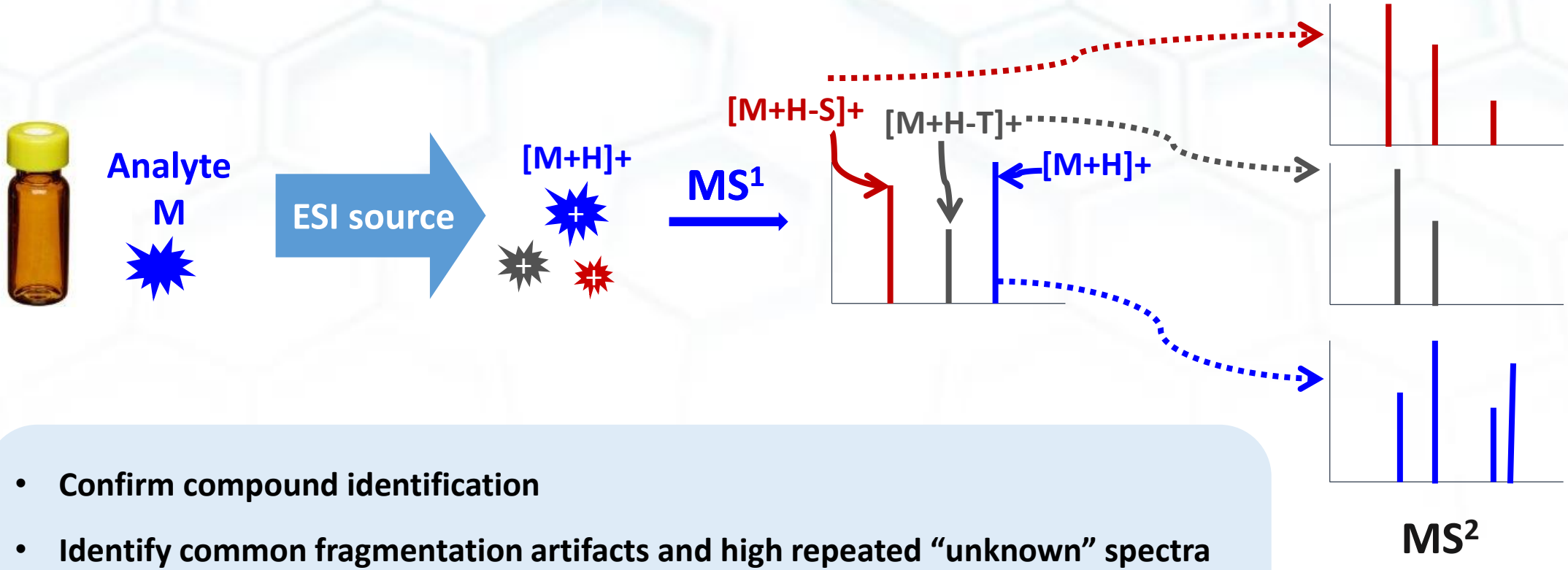
MS^3





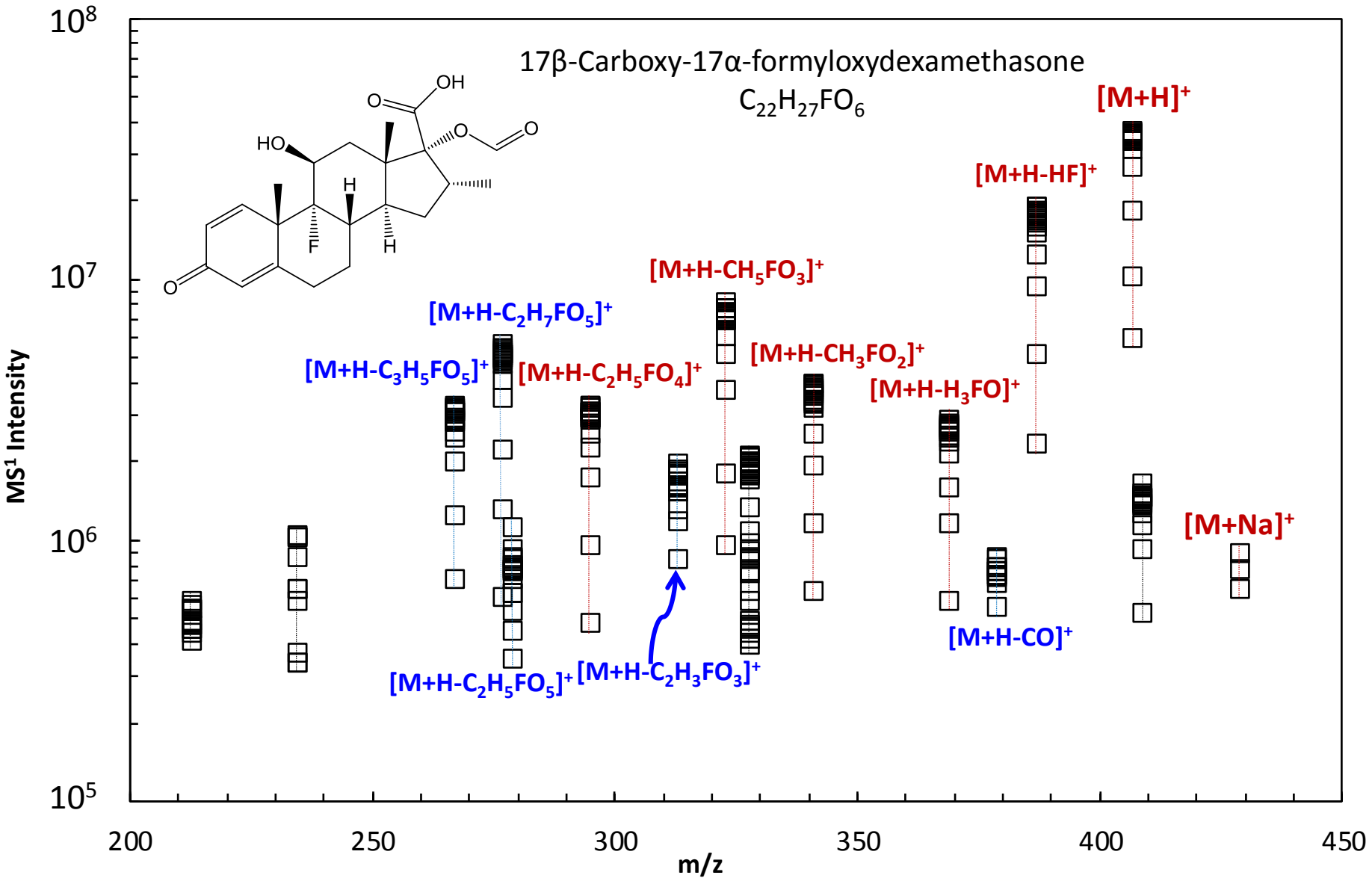
In-source Fragmentation

- In-source Fragmentation: Fragments from the original target compound within the ESI source at the same elution time

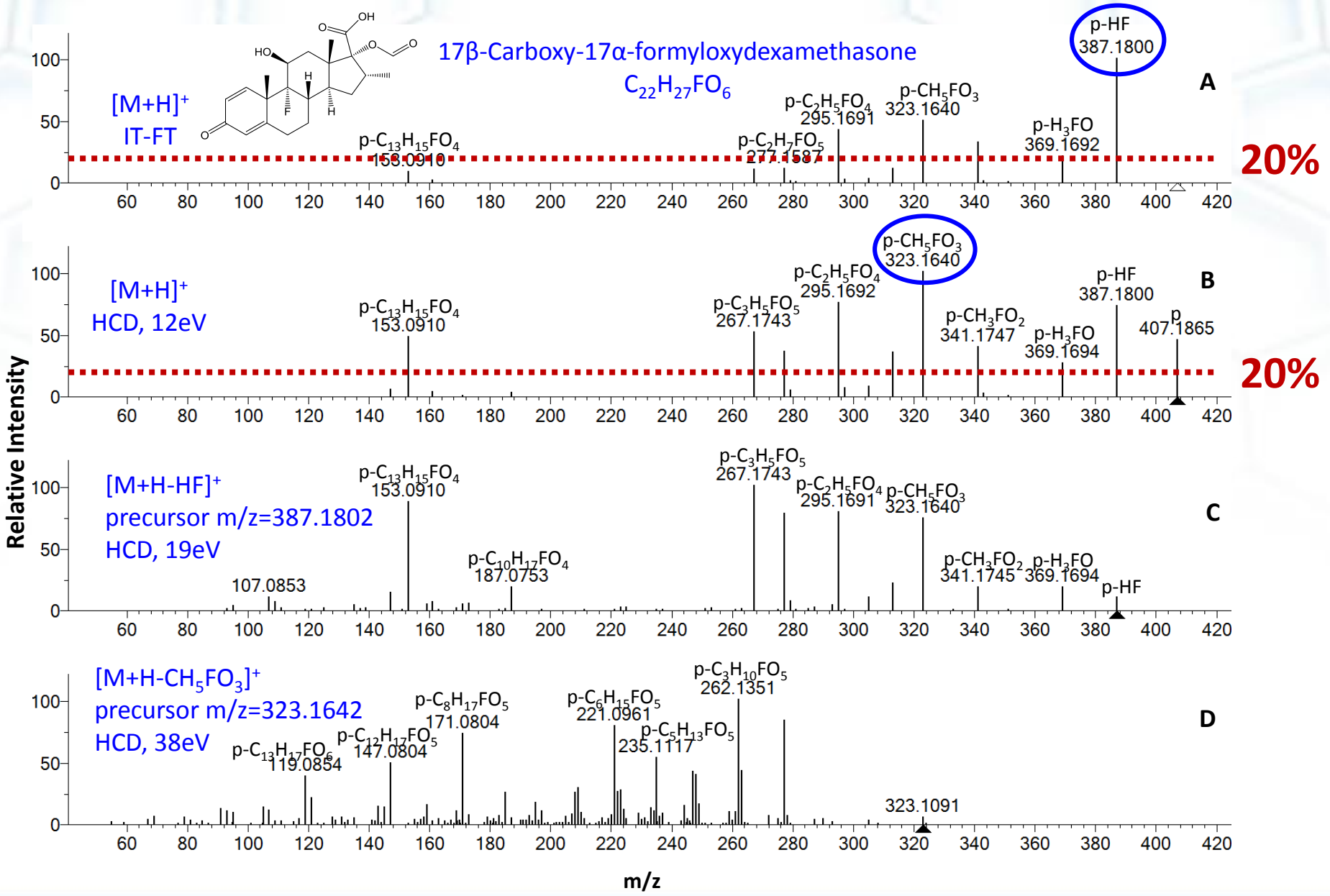


- **Confirm compound identification**
- **Identify common fragmentation artifacts and high repeated "unknown" spectra**
- **Let users be aware that the identified compounds might not be in their sample**
- **Help identify compounds that are not in the library**

In-source Fragmentation Ions



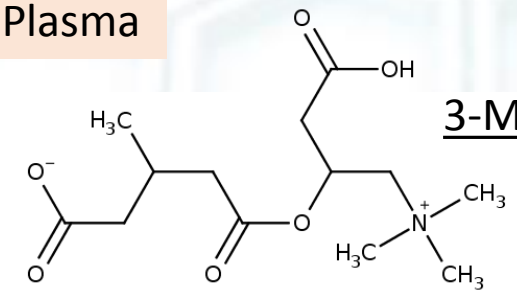
MS² Spectra of In-source Fragmentats in the Library



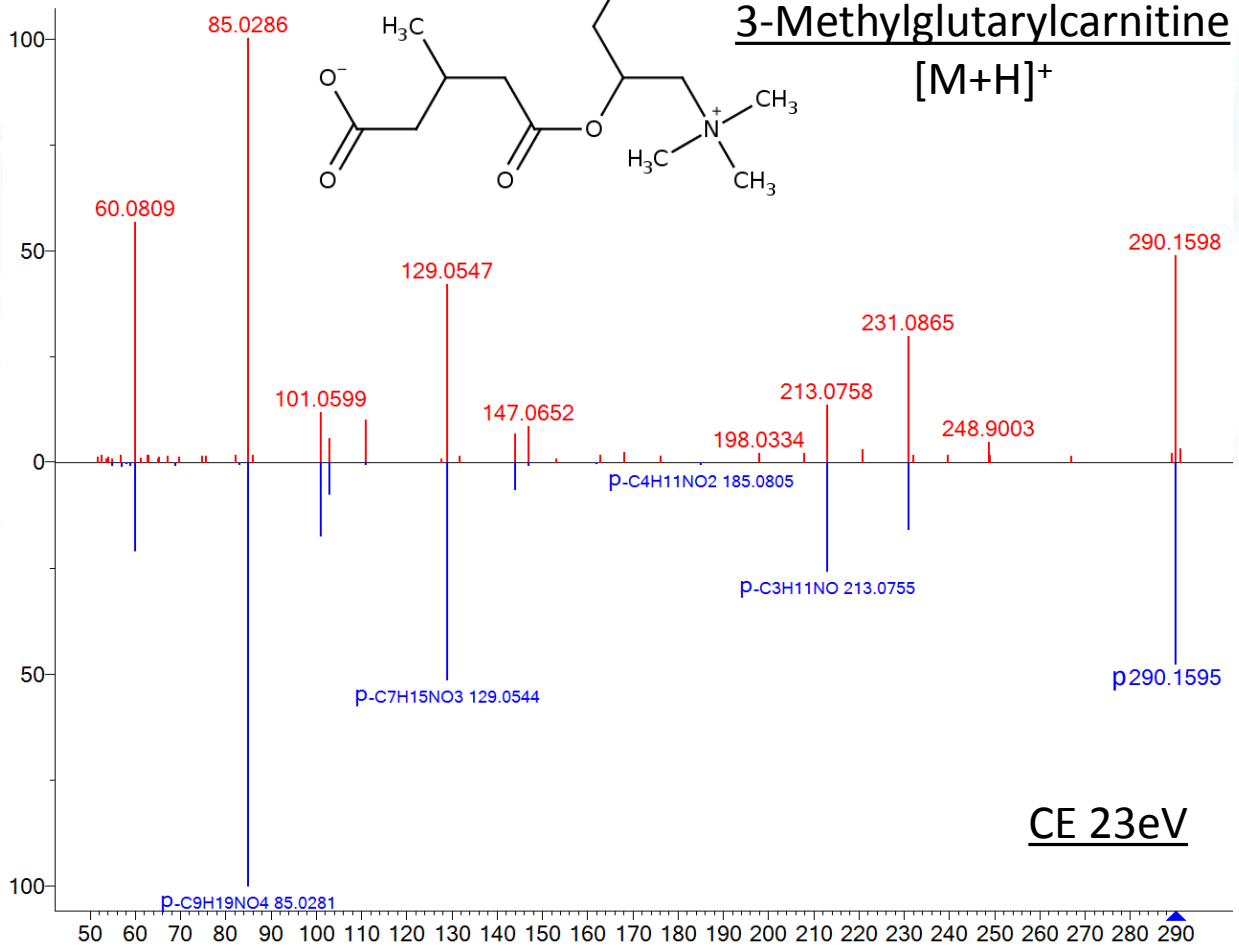
Identifying Metabolites Using the Tandem Mass Spectral Library

UPLC, HCD CE 25%

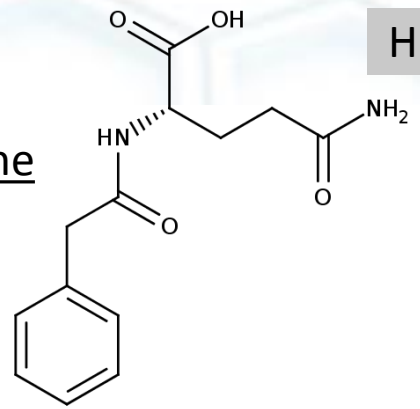
Human Plasma



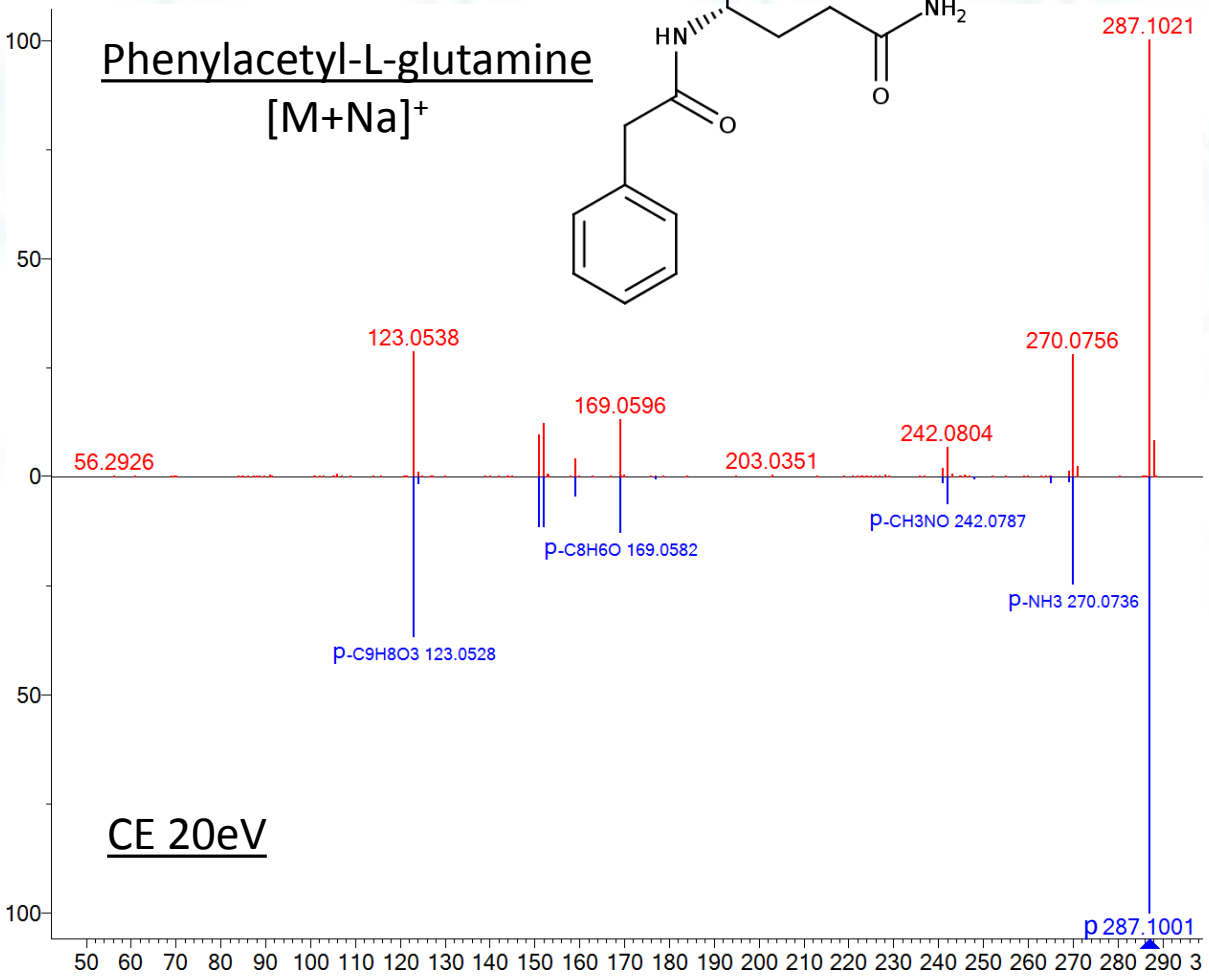
3-Methylglutarylcarnitine
$[M+H]^+$



Human Urine



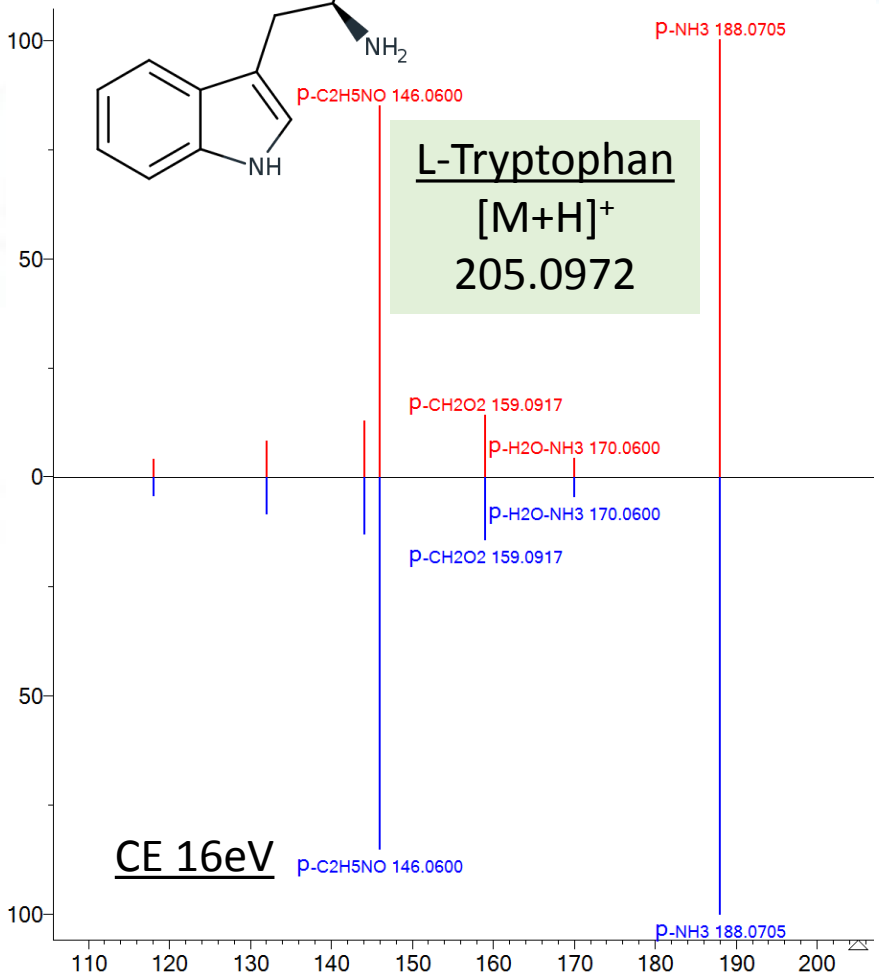
Phenylacetyl-L-glutamine
$[M+Na]^+$

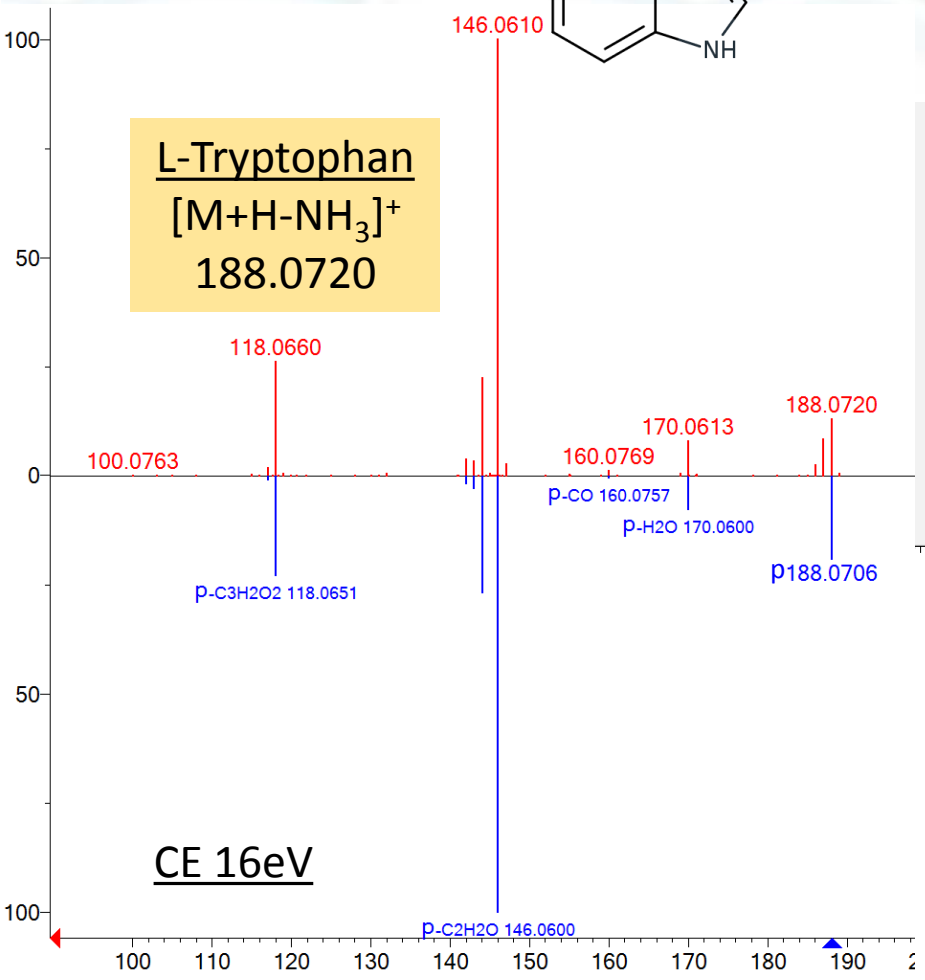


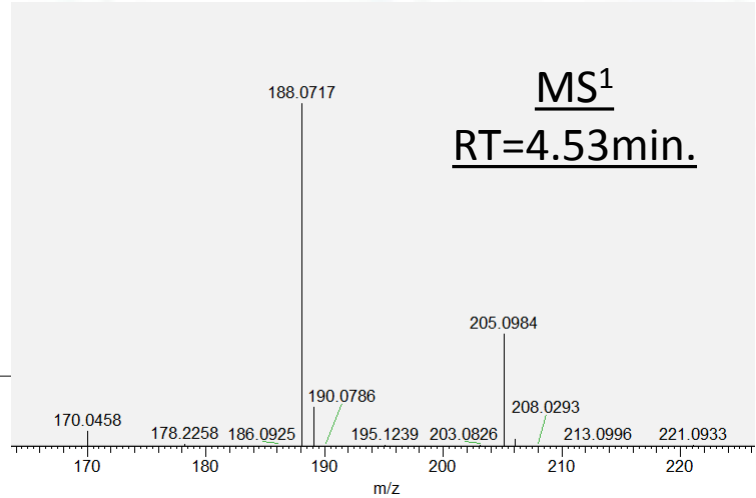
Library

In-source Fragmentation: the Identified Compound Might Not Be in the Sample

Human Urine, UPLC, HCD CE 25%





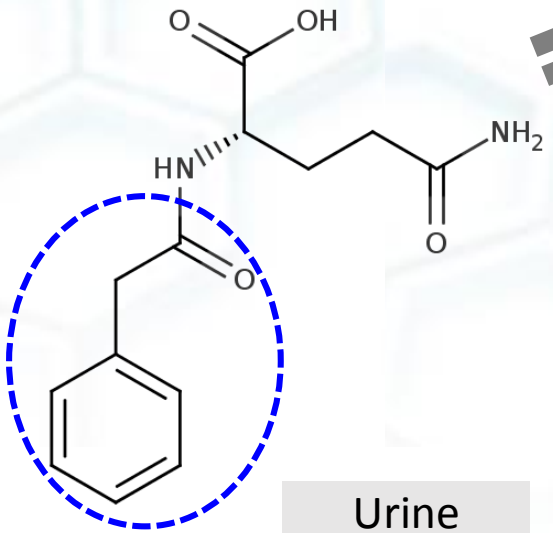


3-Indoleacrylic acid

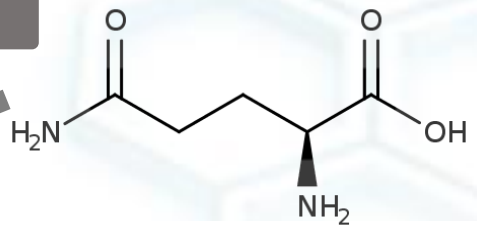
Library

Identifying Metabolites by Searching the Negative Mode Spectra in the Library

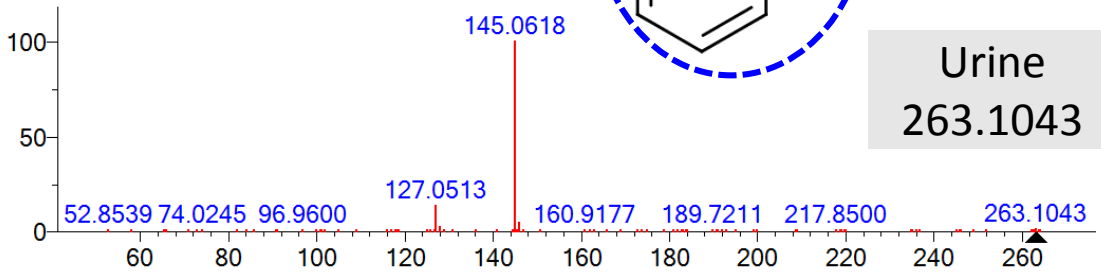
Human Urine, UPLC, HCD CE 25%

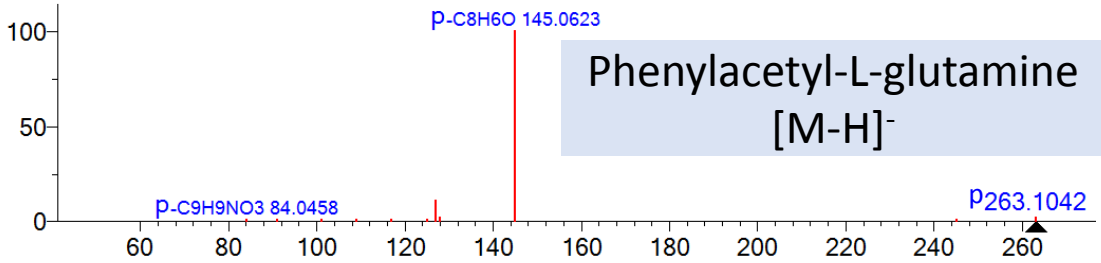


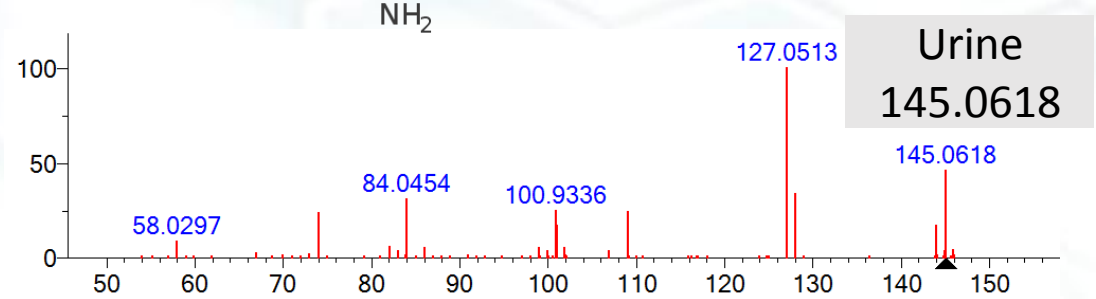
Phenylacetyl-L-glutamine

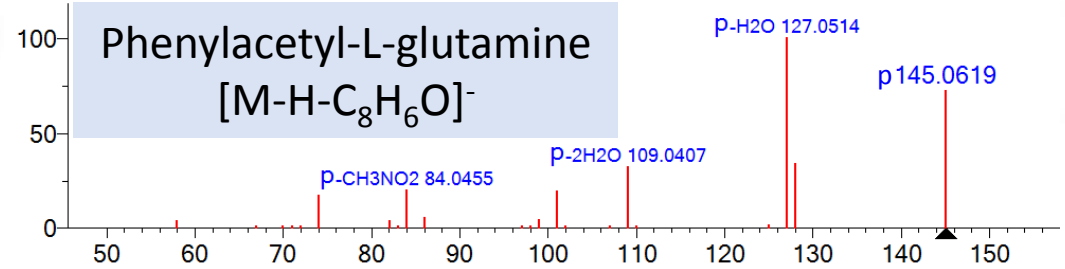


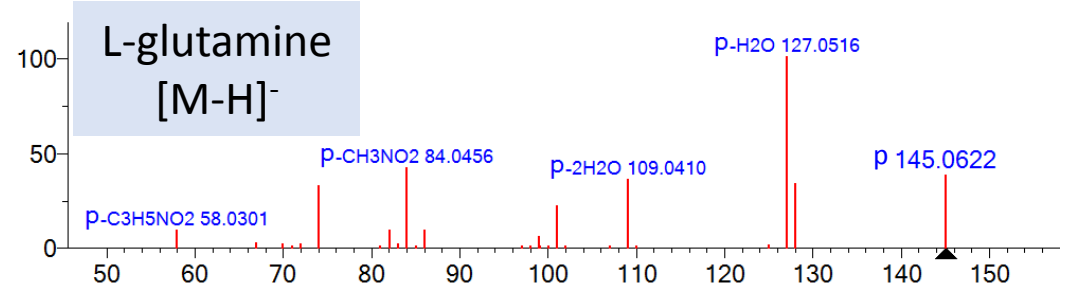
L-glutamine





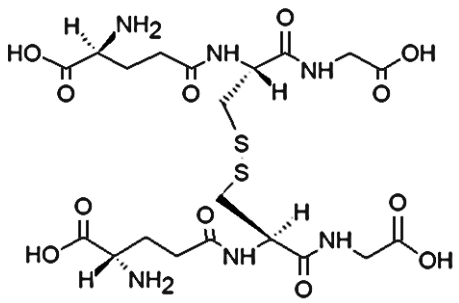




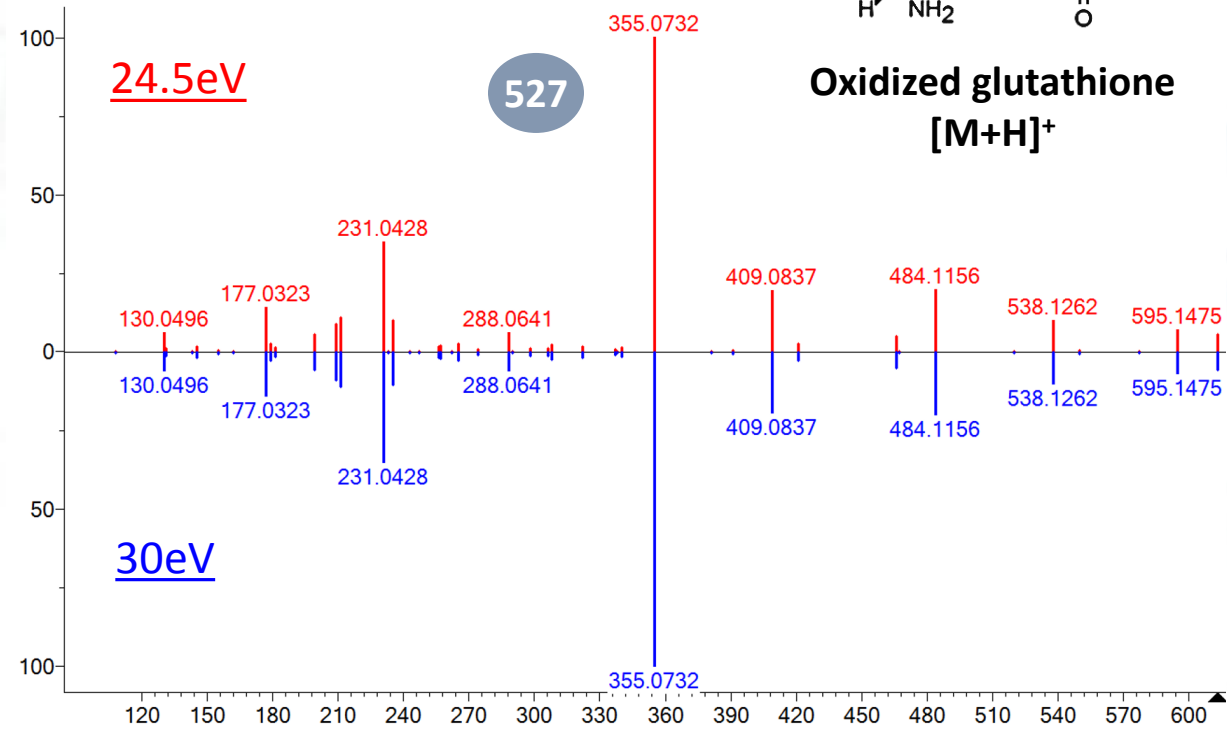


MS² of In-source Fragmentation for Identifying High Repeated "Unknown" Spectra

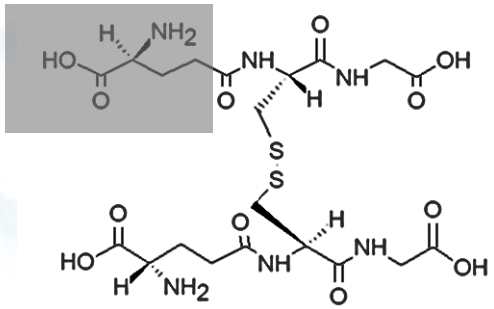
E. Coli
HCD



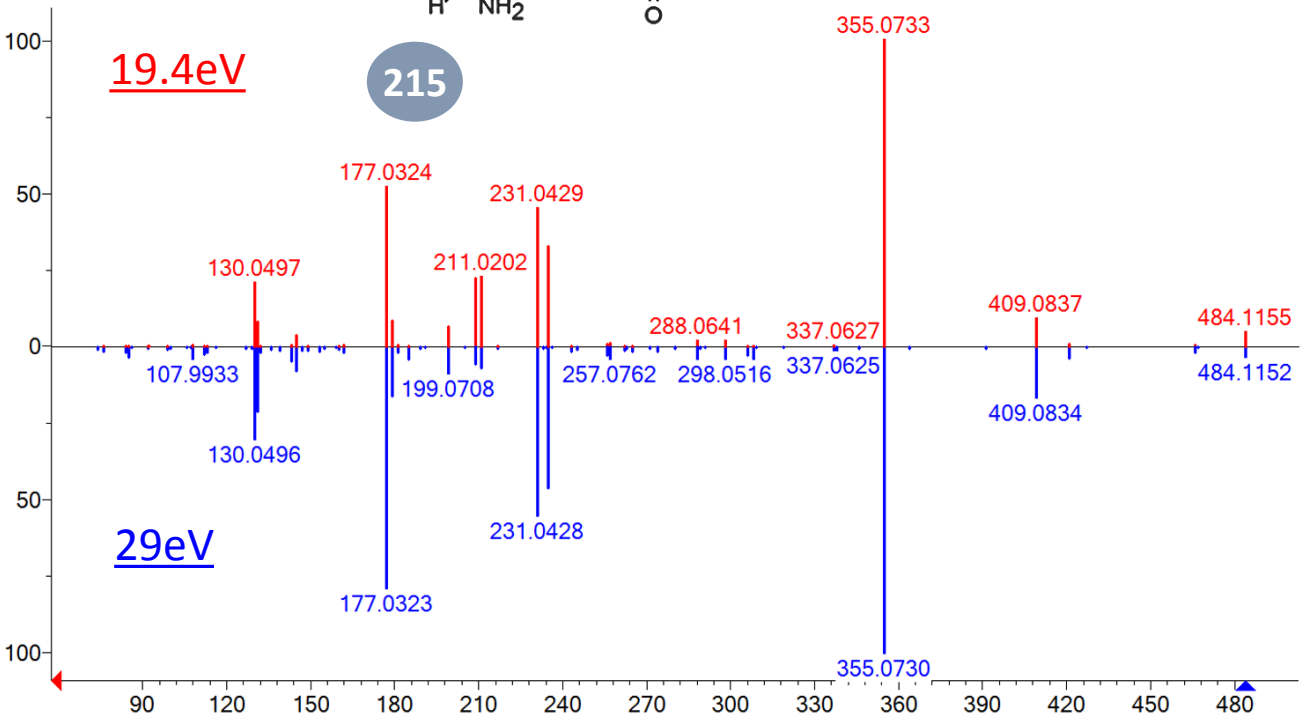
Oxidized glutathione
$[M+H]^+$



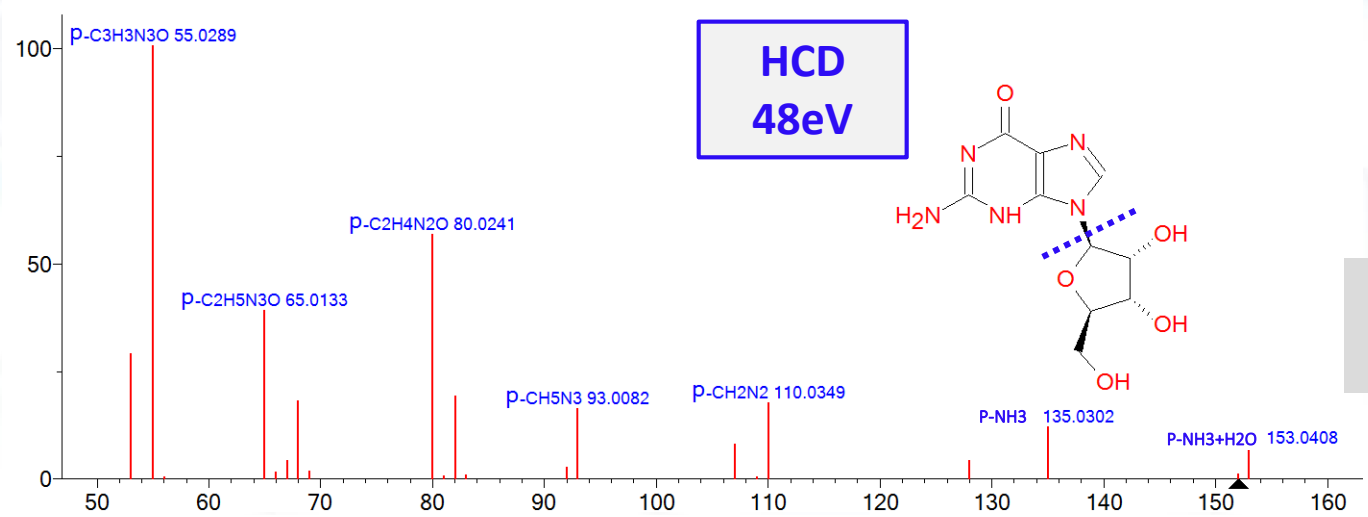
Library



Oxidized glutathione
Fragment
$[M+H-C_5H_7NO_3]^+$

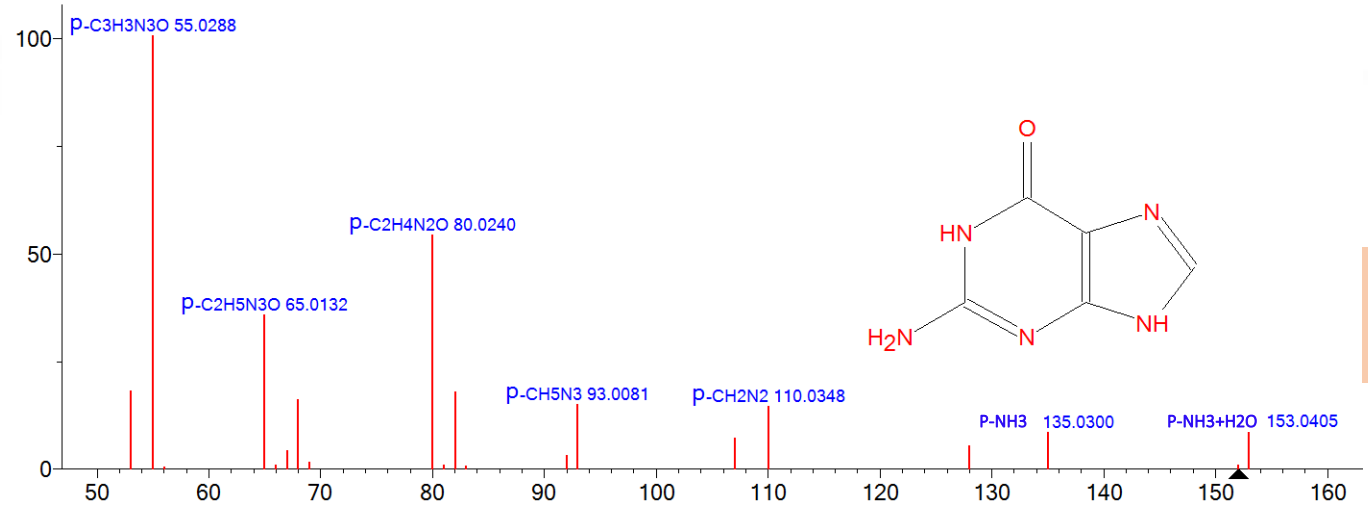


Help Identify Compounds that Are Not in the Library



Guanosine
$[M+H-C_5H_8O_4]^+$

In-source fragment



Guanine
$[M+H]^+$

Authentic

Summary

- We extended the tandem mass spectral library with a wider range of compounds (including 3,500 human metabolites).
- The tandem mass spectral library was extended with spectra of
  - in-source fragments
  - MS^n
  - Multiple precursor ions including isotopic precursors
- This comprehensive high quality reference library (15K compounds with 124K precursor ions and 574K spectra) was tested and can be used for accurate, fast and reliable compound identification.

Acknowledgements

Pedatsur Neta
Stephen E. Stein
Yuri Mirokhin
Dmitrii Tchekhovskoi
Oleg Toropov
Sanford Markey
Yuxue Liang
Yamil Simon
Kelly Telu
Tallat H. Bukhari
William E. Wallace