Mass Spectrometry & Proteomics:
Fundamentals and Applications

Sheng Zhang

Director, Proteomics & MS Core Facility,
Cornell University

Presented at MS Workshop
October 20, 2009

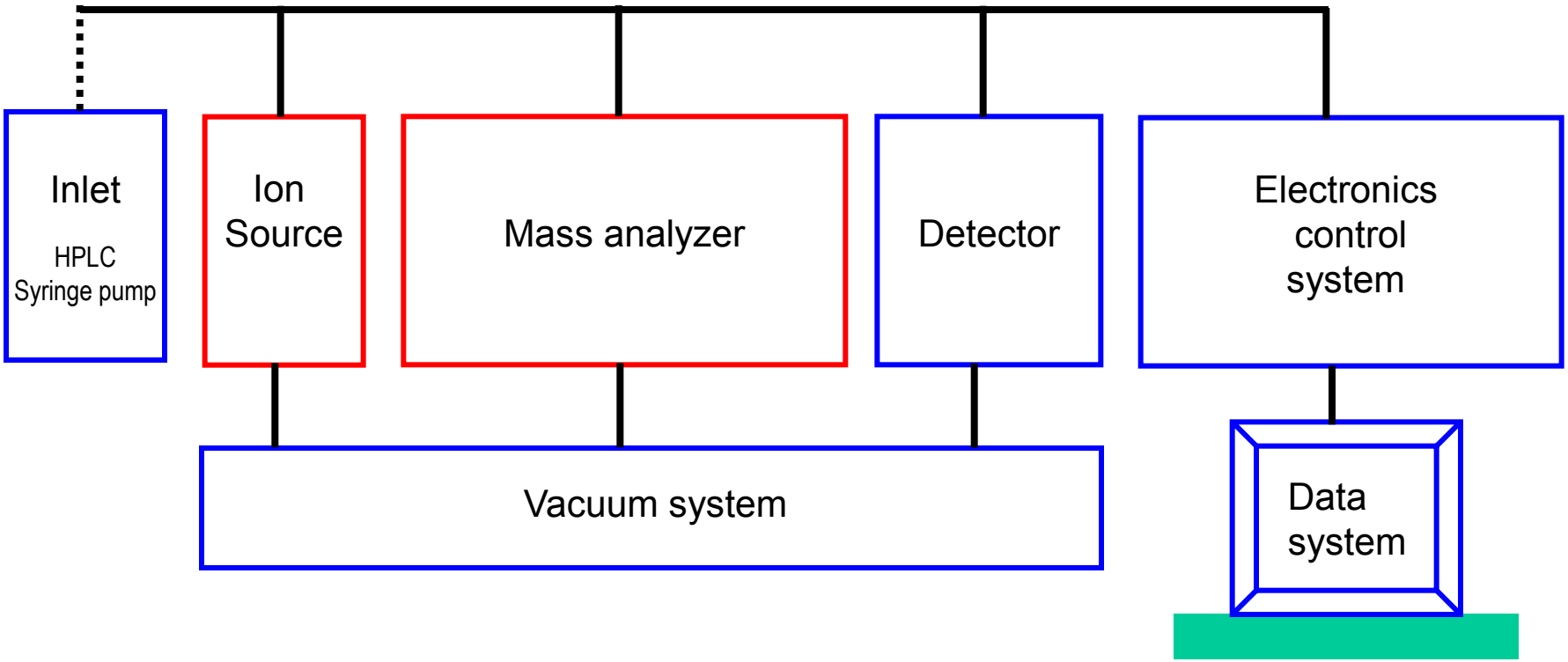
Contact: sz14@cornell.edu
http://cores.lifesciences.cornell.edu/brcinfo/?f=2

Overview

- Fundamentals of Mass Spectrometry Technologies
- Applications of Proteomics (case studies)
  - Protein Identification by GeLC-MS/MS
    - ✓ Non-targeted DDA vs targeted-based MRM-DDA
  - Quantitative proteomics
    - ✓ iTRAQ-based shotgun and MRM analysis
  - Protein PTMs
    - ✓ Phosphorylation and N/C-terminal determination
  - Protein-protein noncovalent interaction
    - ✓ A ternary mixture involved in bacterial chemotaxis

Fundamentals of Mass Spectrometry

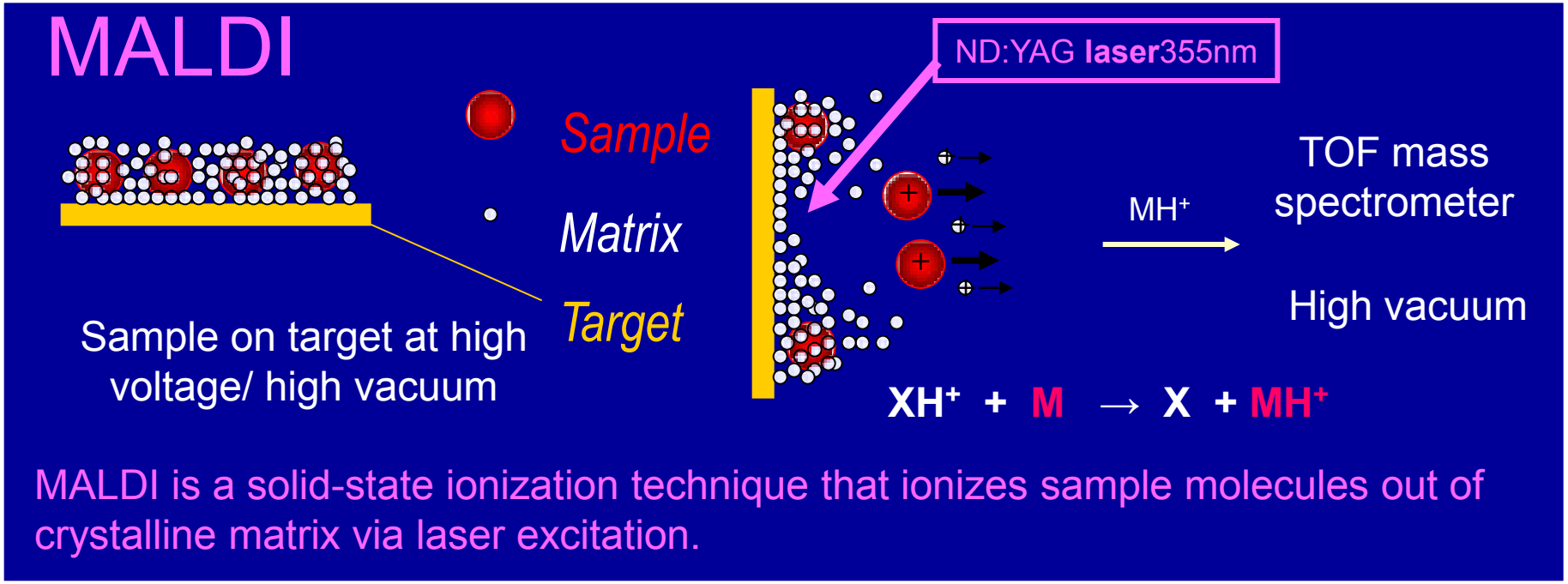
Basic Components of Mass Spectrometers



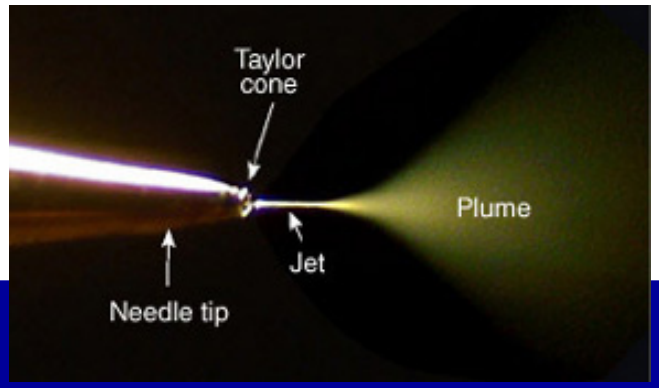
Ion Source:

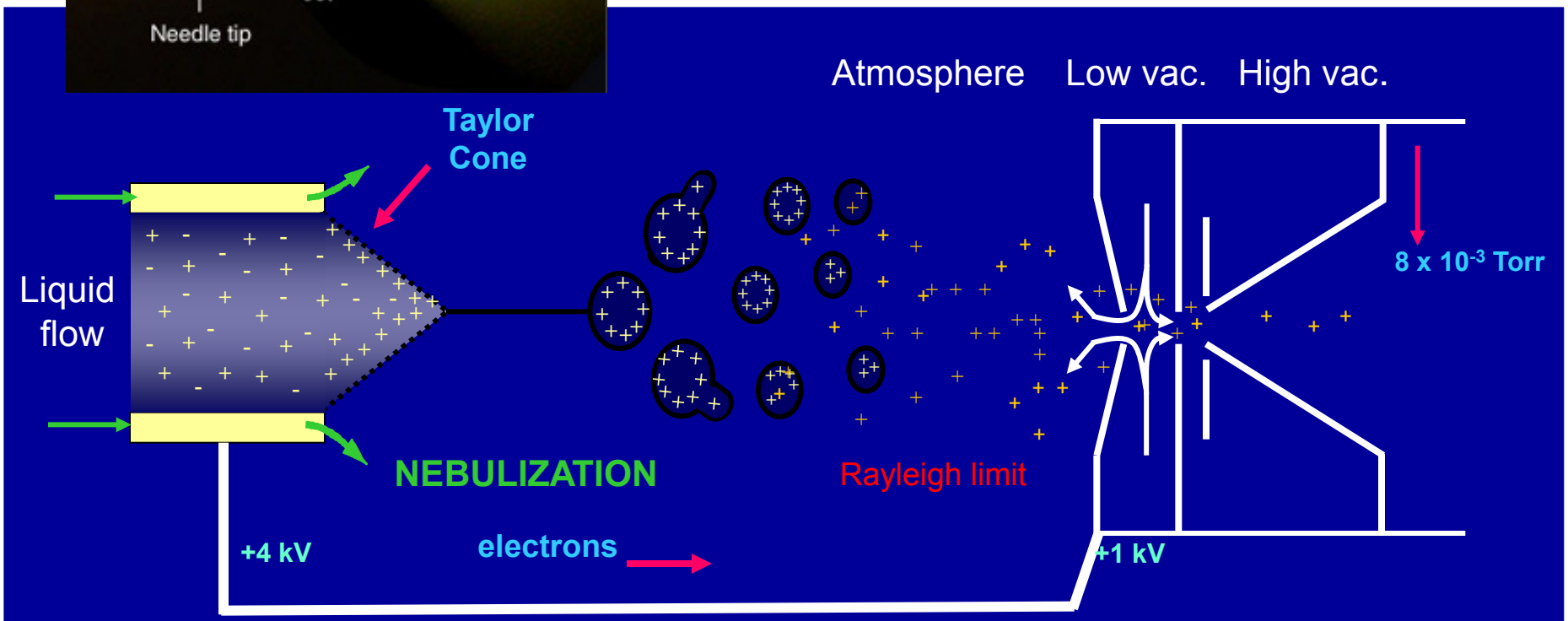
Ionization Methods:

1. **Electron ionization:** $M + e^- \longrightarrow M^{+\bullet} + 2e^-$
2. **Chemical ionization:** $M + [B + H]^+ \longrightarrow [M + H]^+ + B$
3. **Matrix-Assisted Laser Desorption/Ionization (MALDI)**
4. **Electrospray Ionization (ESI)**



Electrospray Ionization----ESI Theory

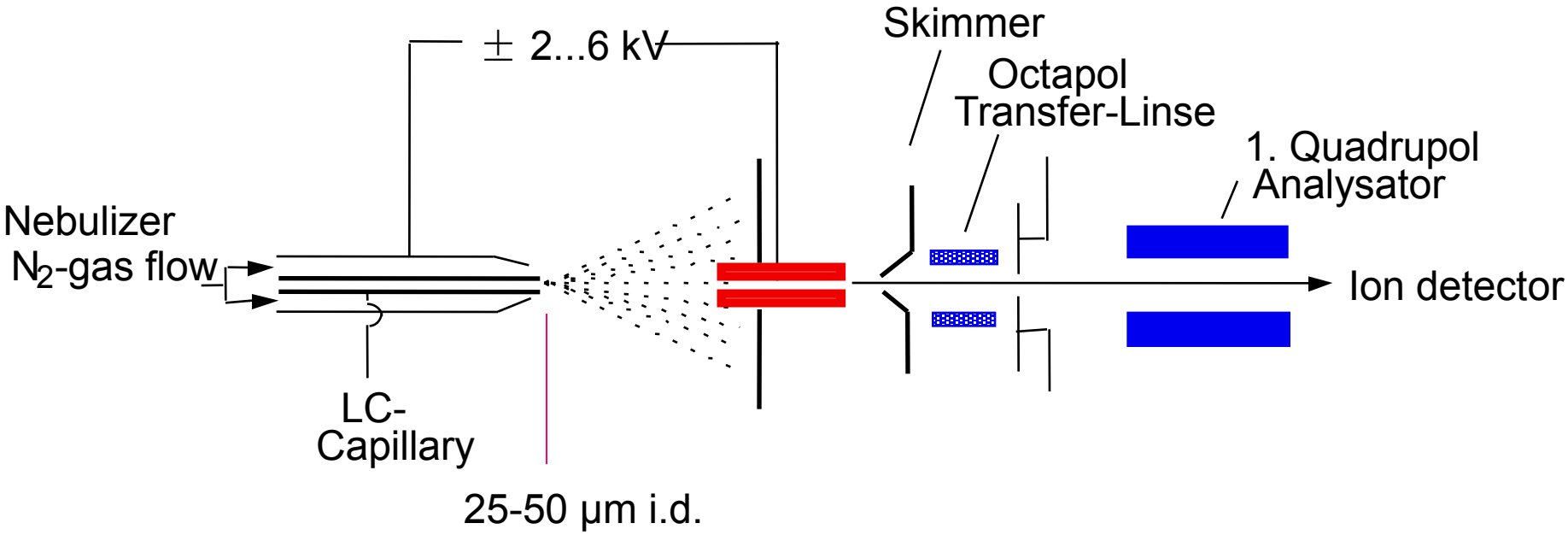




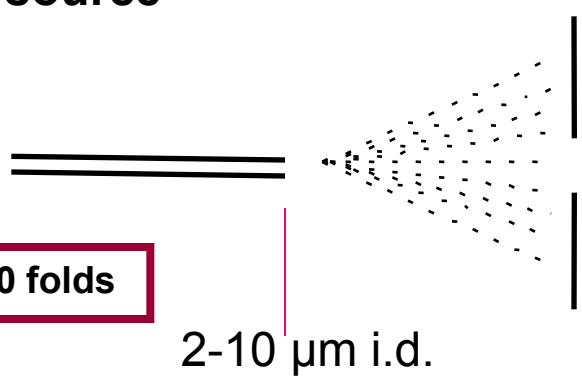
ESI is a solution technique that generates a continuous stream of ions by a 3-step process: droplet formation, droplet shrinkage and ion desorption

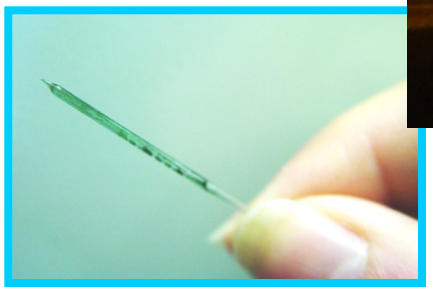
Comparison of Two LC/ESI Ion Sources

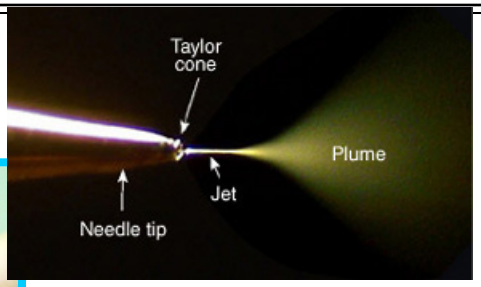
LC-Interface (turbo ion source)
4-20 µl/min



NanoESI ion source
0.2-0.5 µl/min

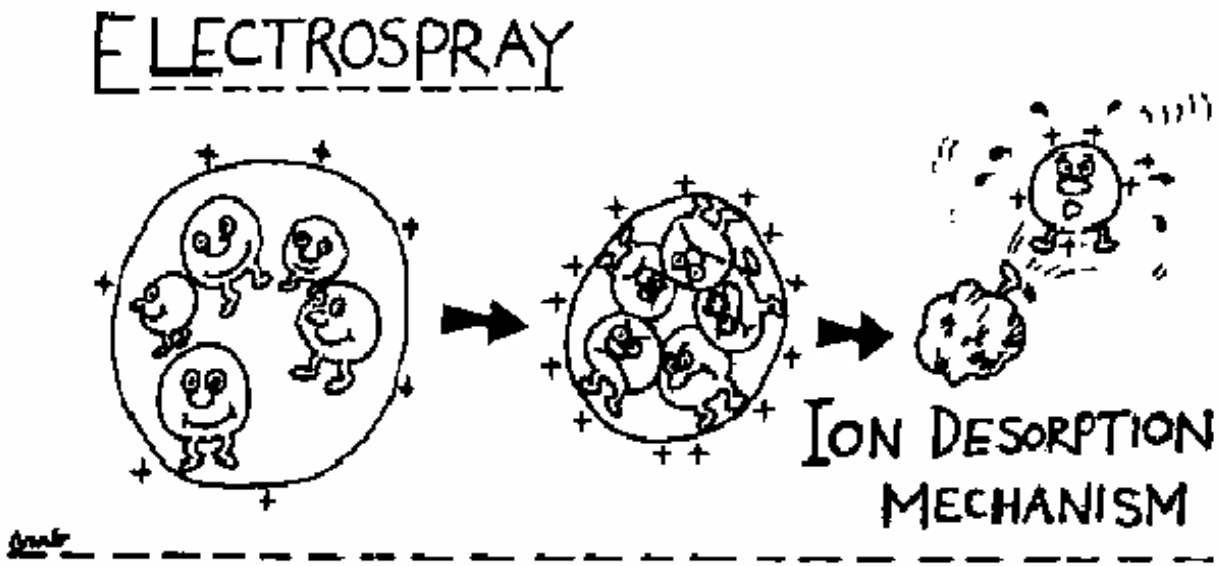




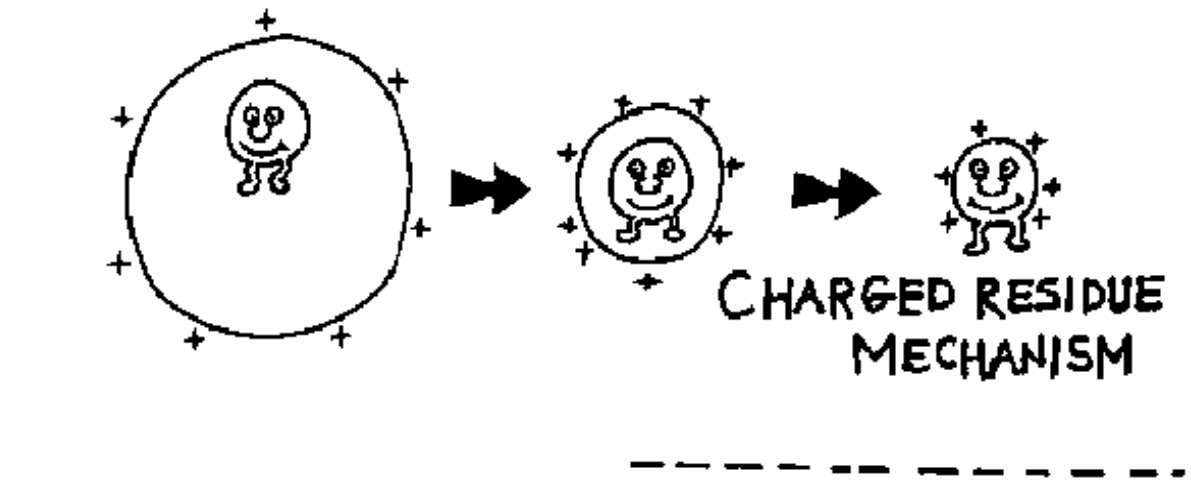


Flow rate ↓ : 20 folds

Two ESI Mechanism Models



Applies to ions with significant surface Activity (hydrophobic Molecules: peptide, Fatty acids etc.).



Applies largely to hydrophilic species Molecules: protein, metal cations.

Ionization Methods for Proteomics
ESI vs MALDI (complementary)

- **ESI Advantages**
  - Direct coupling of LC to MS
  - High quality MS/MS data
  - Gentle ion source for protein interaction studies
  - Excellent for quantitation
  - Good for small molecules
- **ESI Disadvantages**
  - Limited time for MS/MS
  - No repeat analysis
  - Low tolerance to salt
- **MALDI Advantages**
  - Offline coupling of LC to MS
  - Not time limited for MS/MS
  - Low and high mass MW peptides
  - Allow sample re-analysis
  - High tolerance to salt
- **MALDI Disadvantages**
  - Usually low quality MS/MS
  - Not for NCI studies
  - Not first choice for quantitation

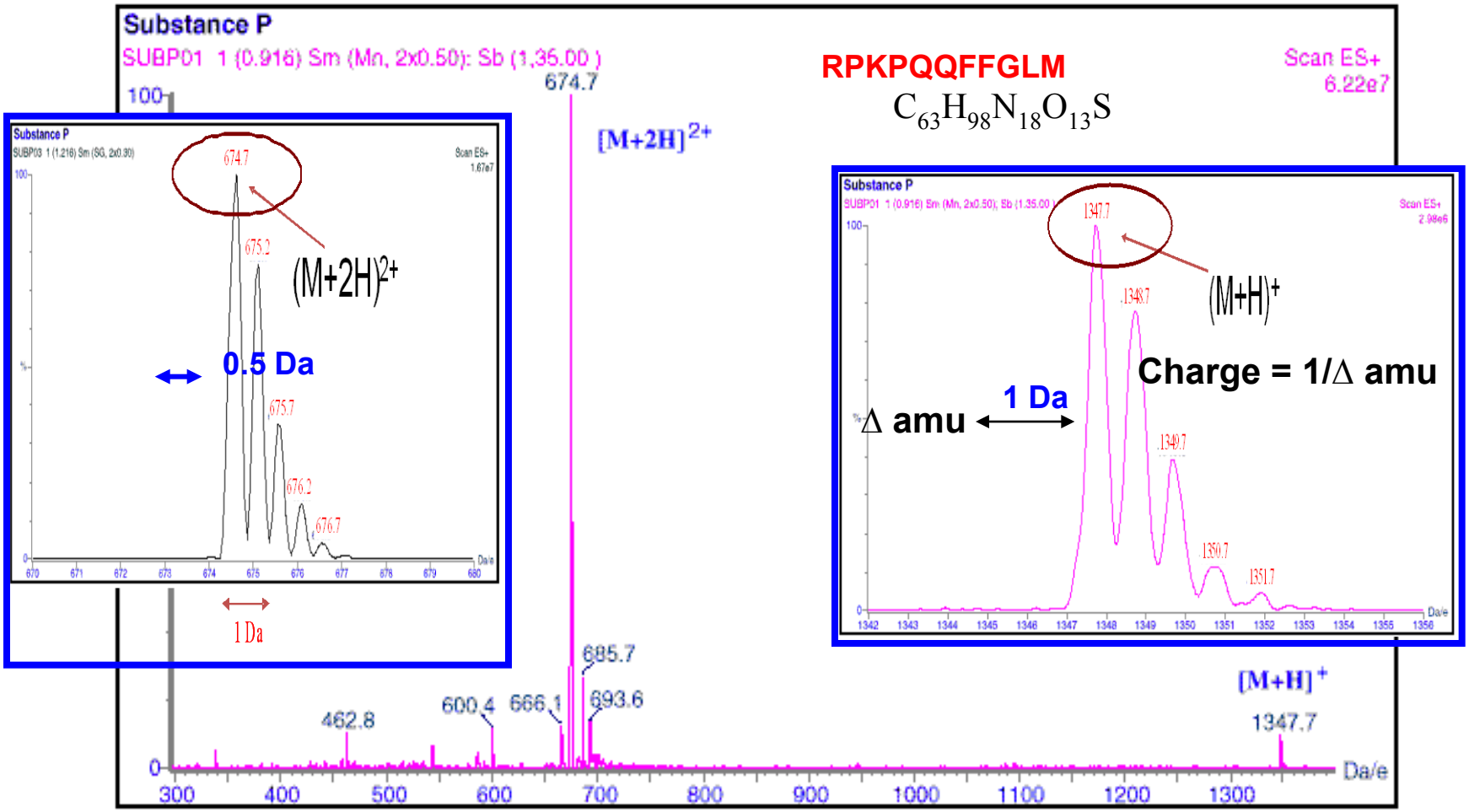
Mass analyzers

- **Mass analyzer performs two vital functions:**
  - ✓ Sort all ions according to their mass-to-charge (m/z) ratio
  - ✓ Focus all mass-resolved ions at a single focal point.

- **The nature of mass analyzer determines key parameters of mass analysis:**
  - ✓ mass resolution
  - ✓ mass accuracy
  - ✓ mass range
  - ✓ scan speed
  - ✓ linear dynamic range, sensitivity, ion transmission efficiency, etc.

- **Most common mass analyzers:**
  - ✓ Quadrupole mass filter (based on trajectory stability)
  - ✓ Ion trap mass analyzer (based on m/z resonance frequency)
  - ✓ Time-of-Flight mass analyzer (based on velocity)

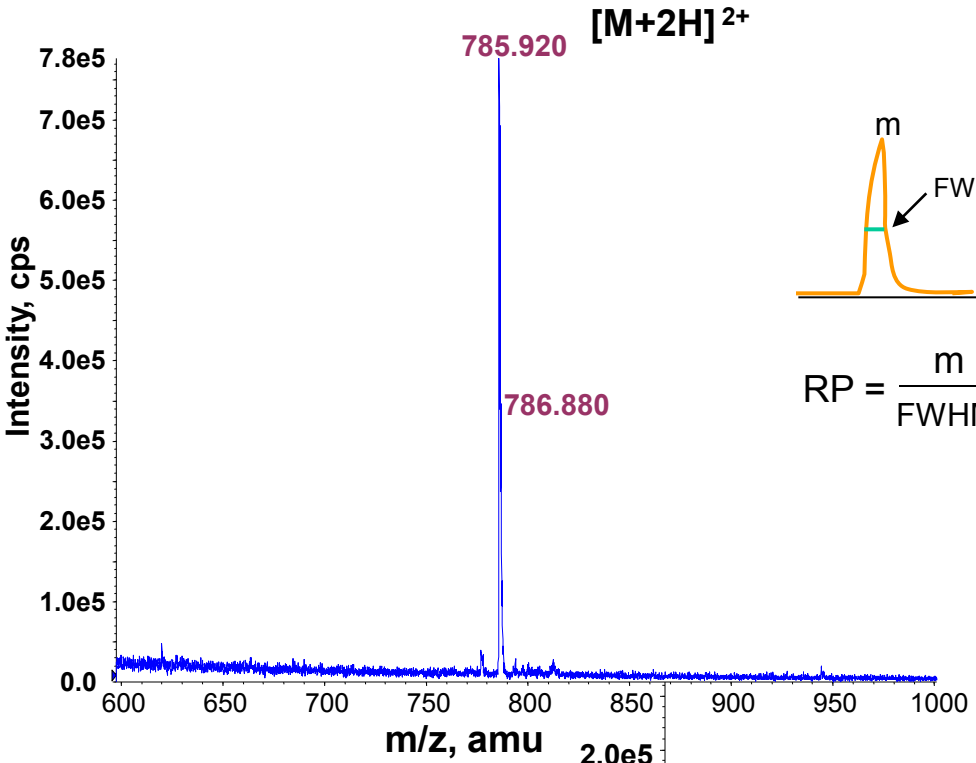
Recognizing Multiply-Charged Ions



Mass spectrometers operate on the basis of mass-to-charge ratio (*m/z*)
Carbon isotopes of doubly charged ions are separated by 0.5 Da

Mass Resolution and Mass Accuracy

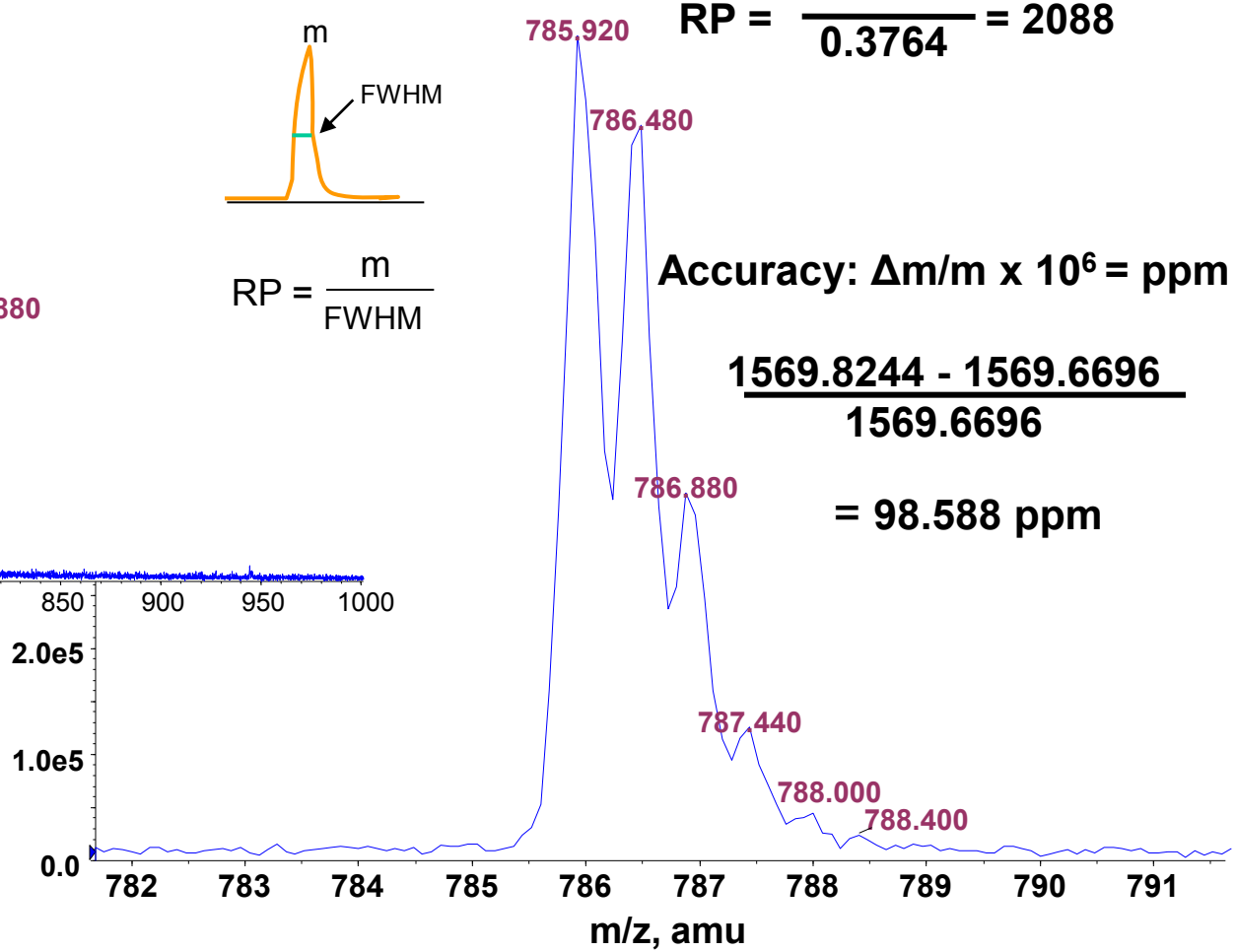
50 fmol/ µL EGVNDNEEGFFSAR



$$RP = \frac{m}{FWHM}$$

Peak width @50% = 0.3764 amu

$$RP = \frac{785.920}{0.3764} = 2088$$

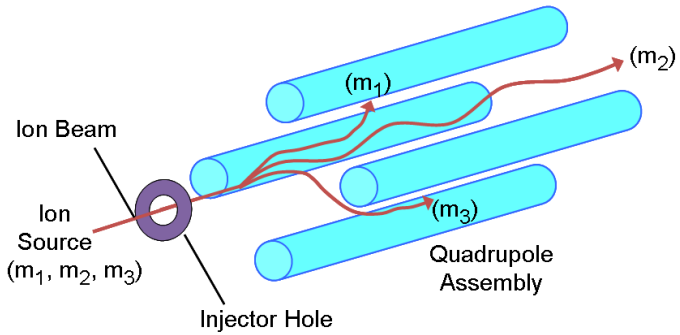


Accuracy: $\Delta m/m \times 10^6$ = ppm

$$\frac{1569.8244 - 1569.6696}{1569.6696}$$

= 98.588 ppm

Quadrupole Theory



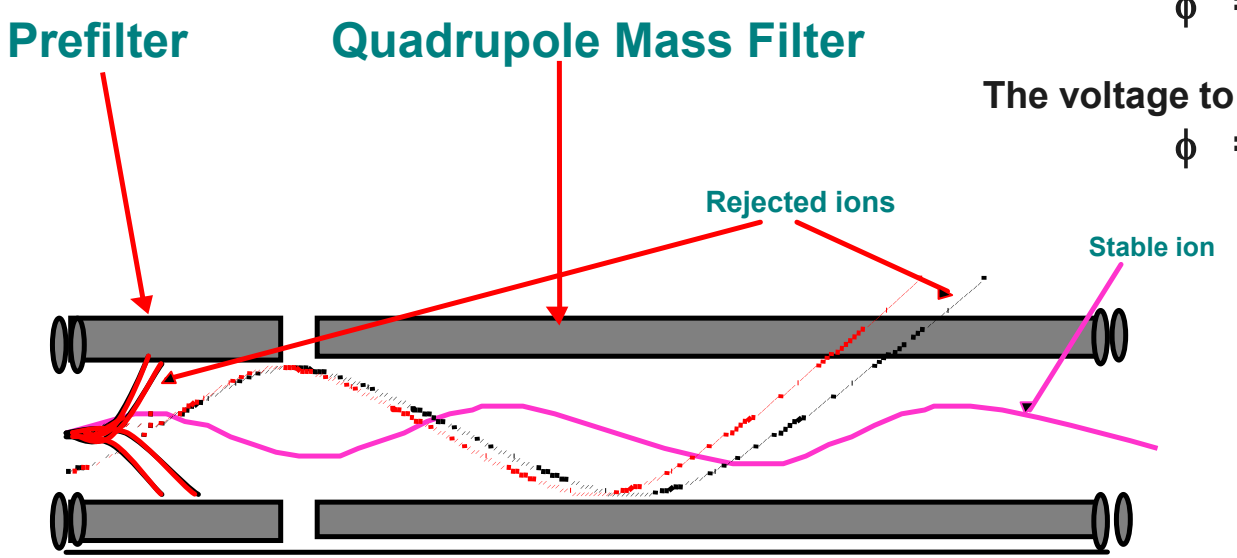
- Four metallic rods arranged symmetrically as two sets in a square array.
- A combination of RF and DC voltages are applied to each pair of rods.
- Produce a complex, oscillating movement of the charged ions as they move across the Q.
- Under a set of defined DC and RF potentials, ions of selected m/z value pass through the Q.

The voltage applied to an opposing pair of rods:

$$\phi = U - V \cos \omega T$$

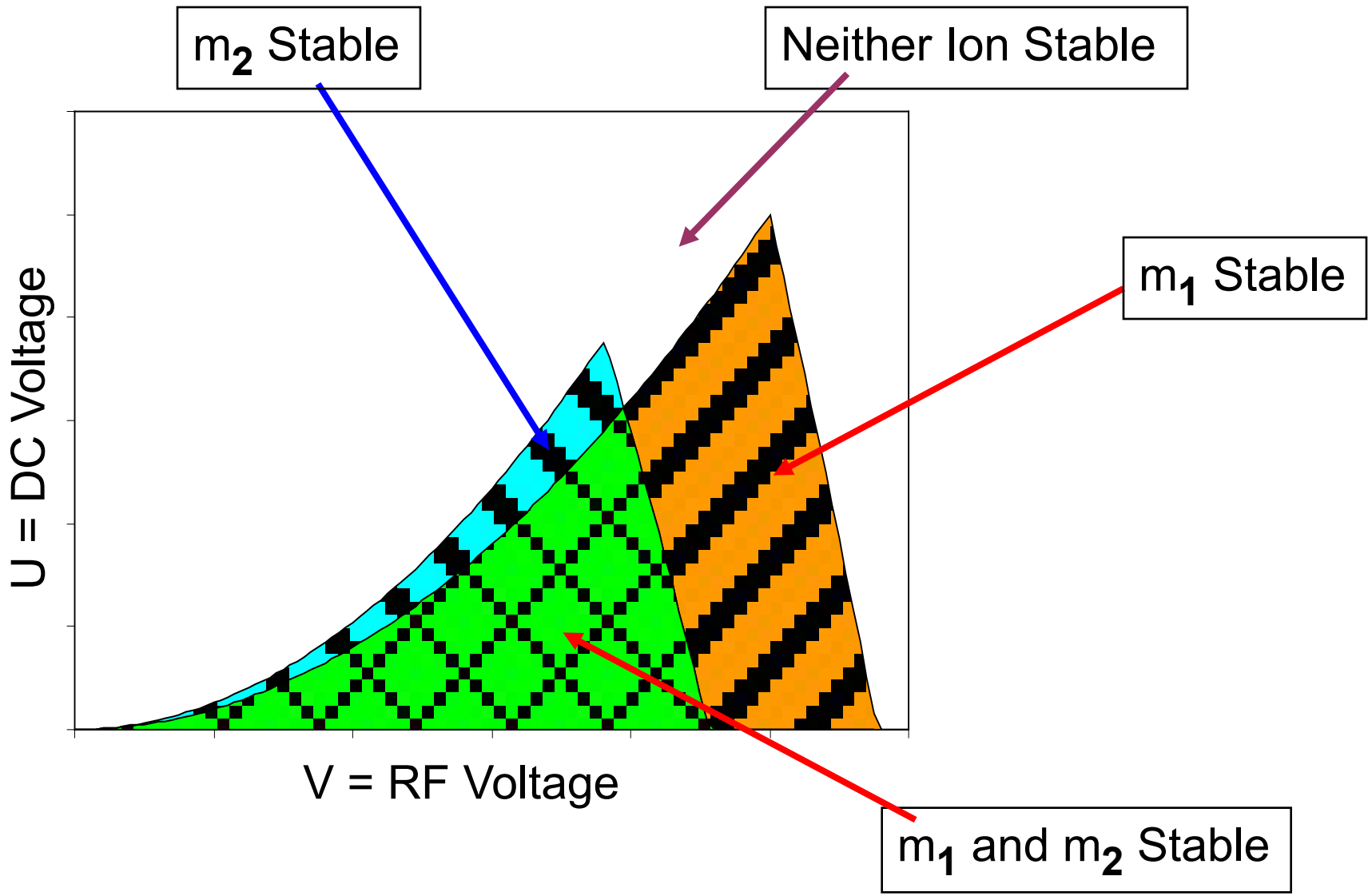
The voltage to the other pair of opposing rods:

$$\phi = -U + V \cos \omega T$$

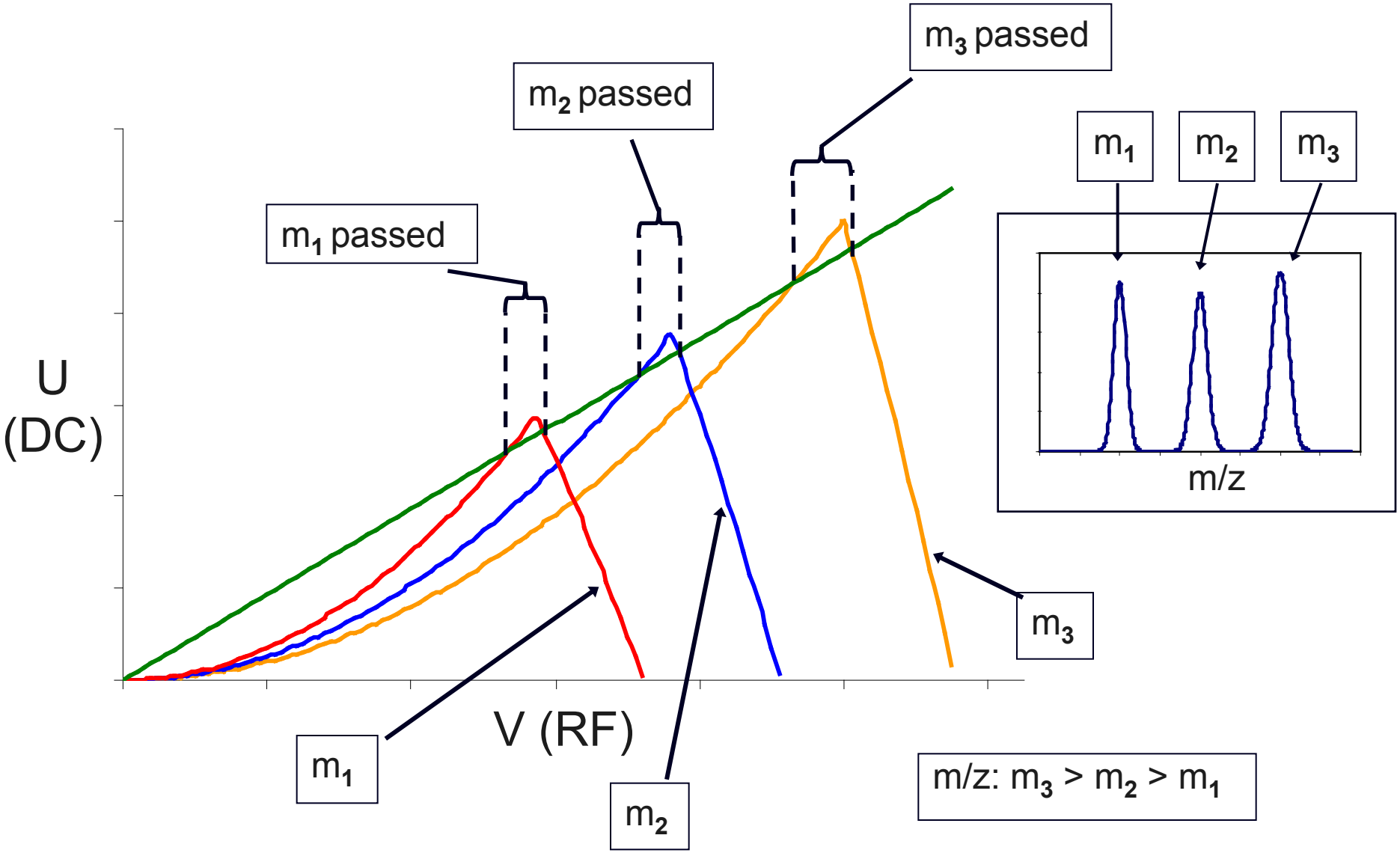


U = DC Voltage
V = RF Voltage
ω = RF frequency (1 mHz)

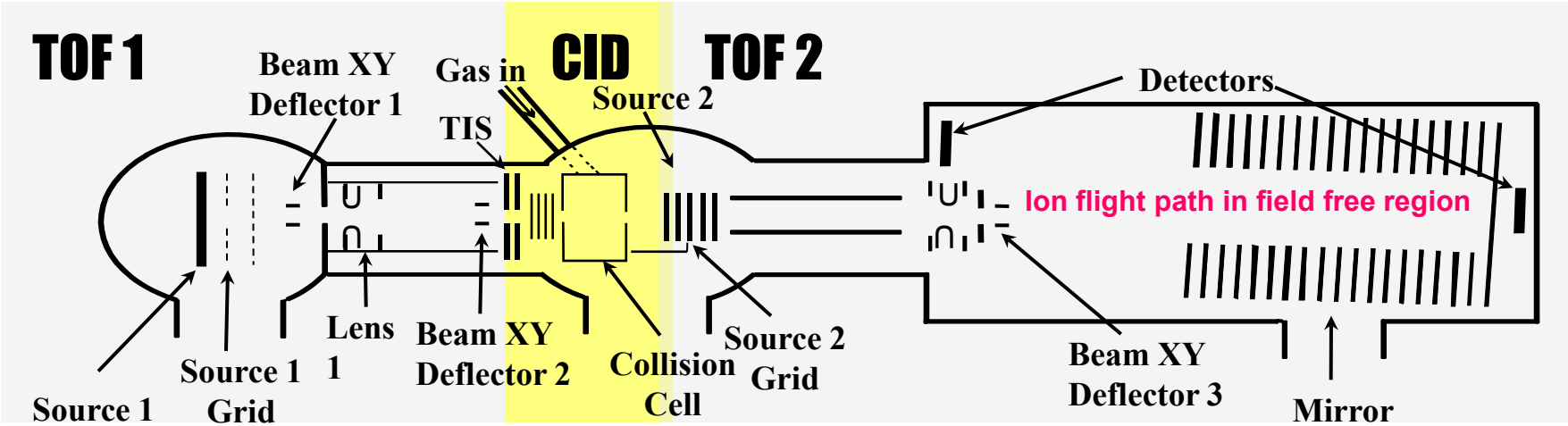
A Stability Diagram: ions are separated in a quadrupole based on the stability of their trajectories in the oscillating electric fields



Quadrupole Mass Filter with U/V = Constant: Example of Better Quadrupole Operational Line



Time-of-Flight Mass Analyzer



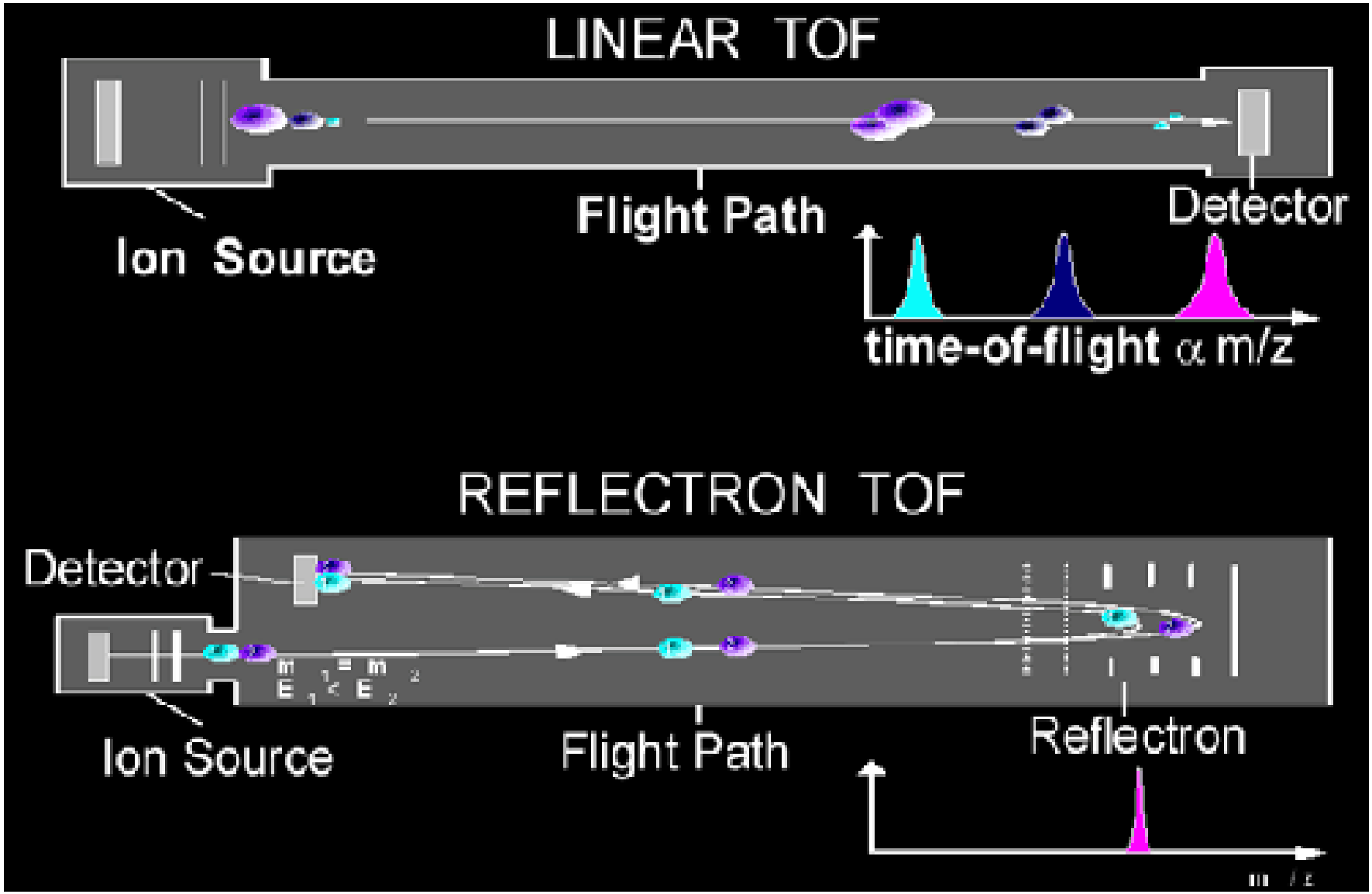
$$KE = V * z = \frac{m * v^2}{2}$$

$$v = \left(\frac{2*V*z}{m}\right)^{1/2}$$

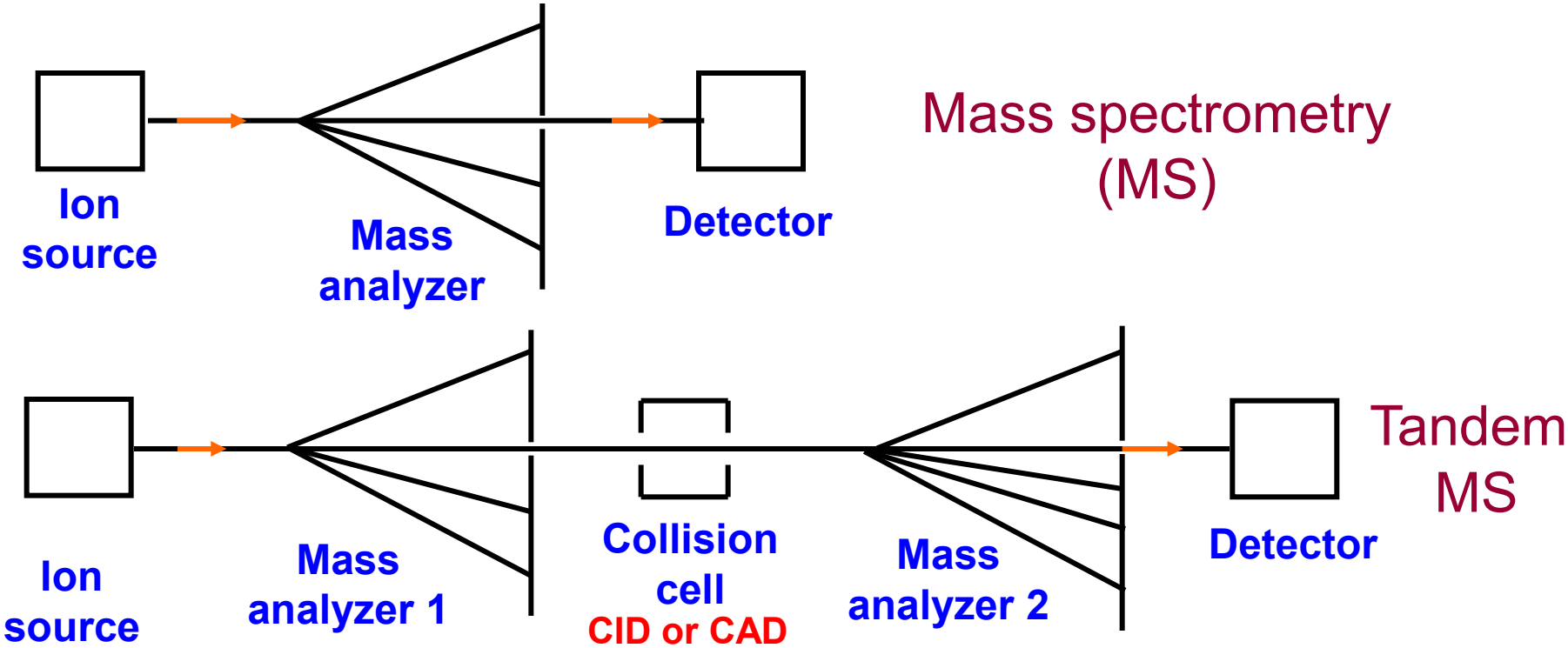
$$t = \frac{L}{v} = L * \left(\frac{m}{2*V*z}\right)^{1/2}$$

- ✓ An ion is given a fixed amount of KE by acceleration in an electric field generated by high voltage (20-30kV).
- ✓ The ion enters a field-free region traveling at a velocity that is inversely proportional to its m/z.
- ✓ The time required for the ion to travel the length of the field-free region is measured.
- ✓ The travel time is to used to calculate the velocity and ultimately the m/z of the ion.

Time-of-Flight Mass Analyzer



Schematic Diagram of MS vs MS/MS

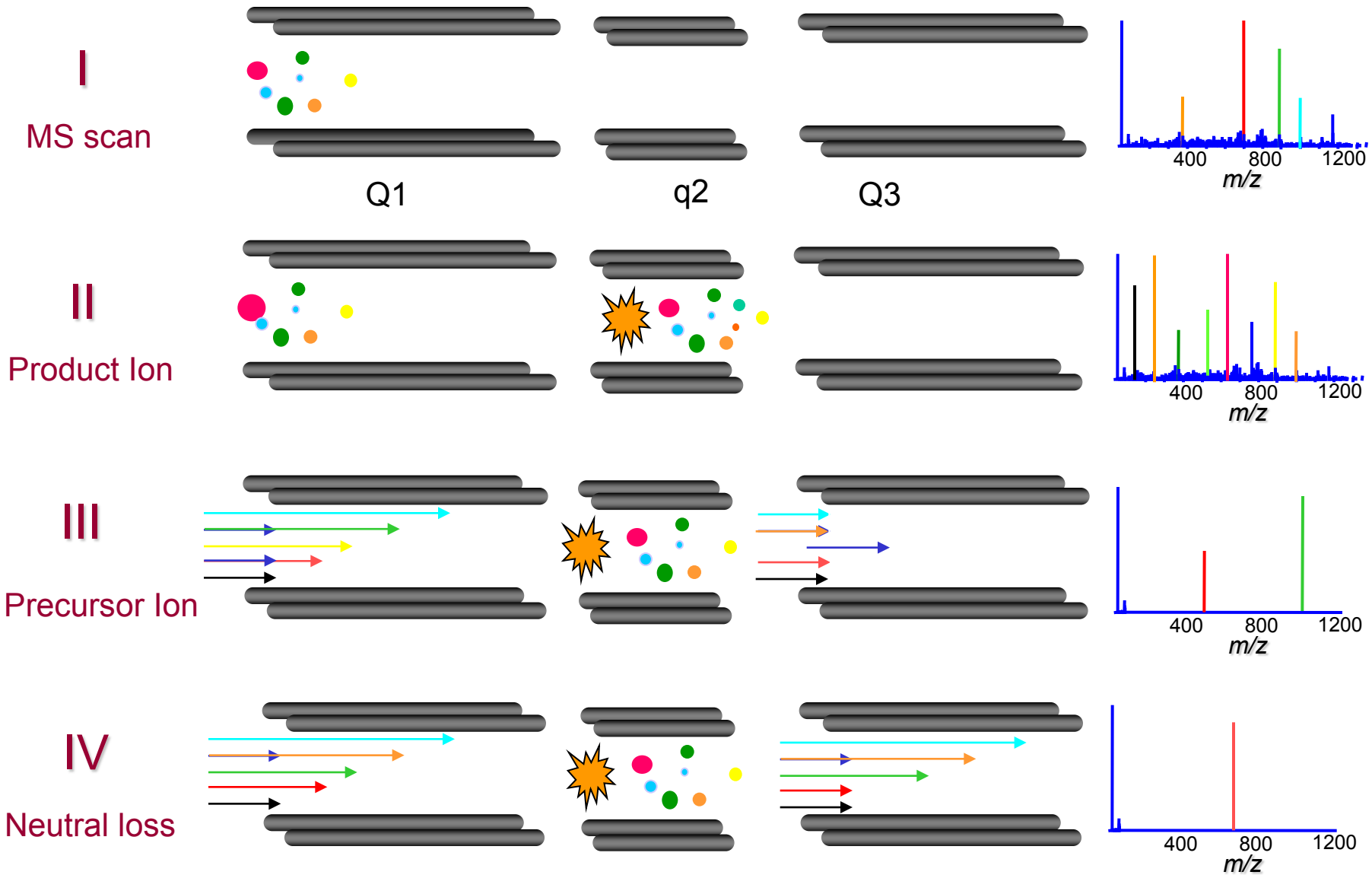


Tandem in space	Tandem in time
Triple quadrupole	3D Ion trap
Quadrupole TOF	Linear Ion trap
Hybrid quadrupole linear ion trap	Fourier transform (FTMS)

Comparison of Tandem Mass Spectrometer's Features

Features	Triple Quadrupole	Quadrupole TOF	Quadrupole Linear Ion Trap
Sensitivity	++	+++	++++
Resolution	+	++++	++
Accuracy	+	++++	++
Quantitation	++++	++	++
Scan speed	++	+++	++++
Mass range	+	++++	++
Dynamic range	++++	++	+++

MS Scan Modes

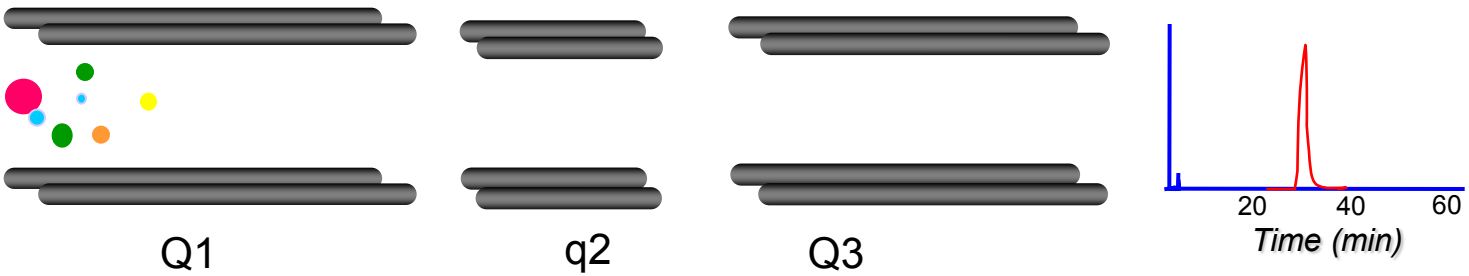


MS Scan Modes

Selected Ion Monitoring

V

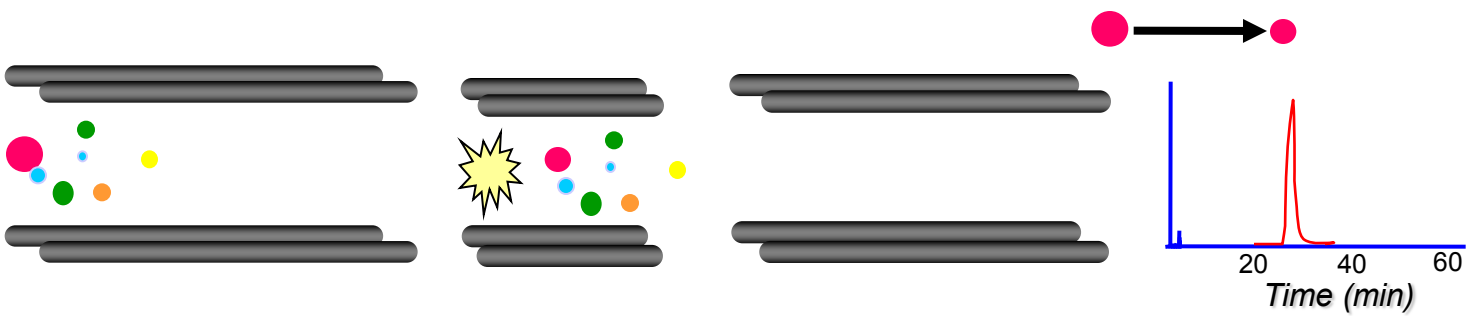
SIM



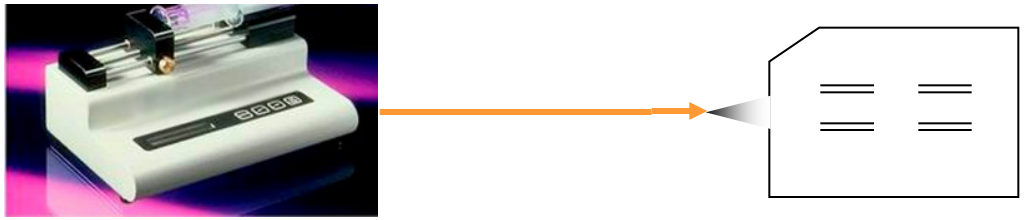
Selected Reaction Monitoring

IV

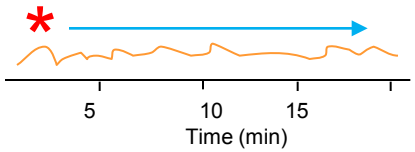
SRM/MRM



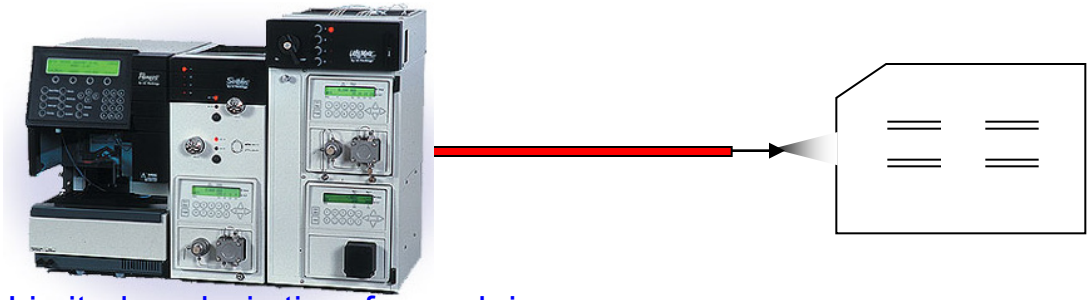
Sample Delivery: Infusion vs On-line LC



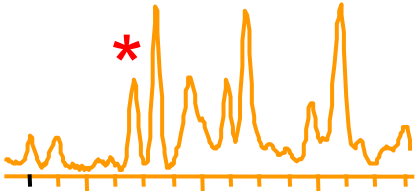
TIC of Infusion



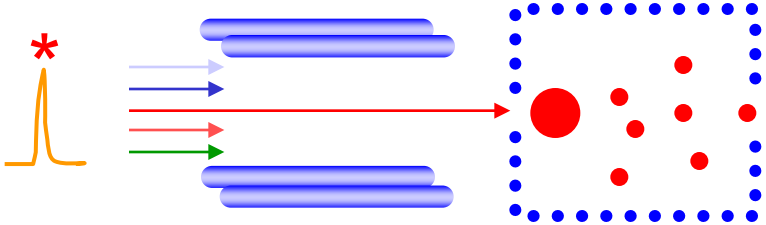
- Extended analysis time
- Clean sample required
- More samples needed
- Relatively low sensitivity

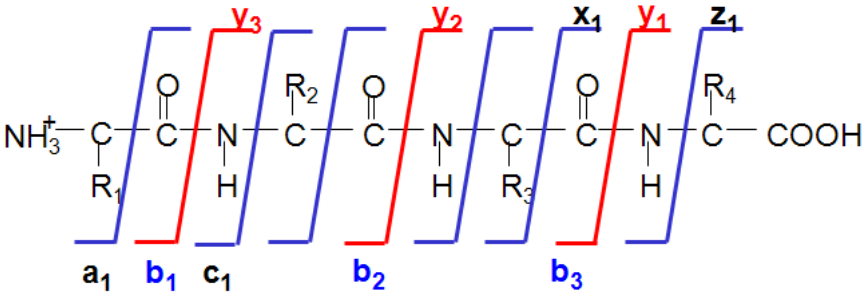


TIC of LC/MS/MS

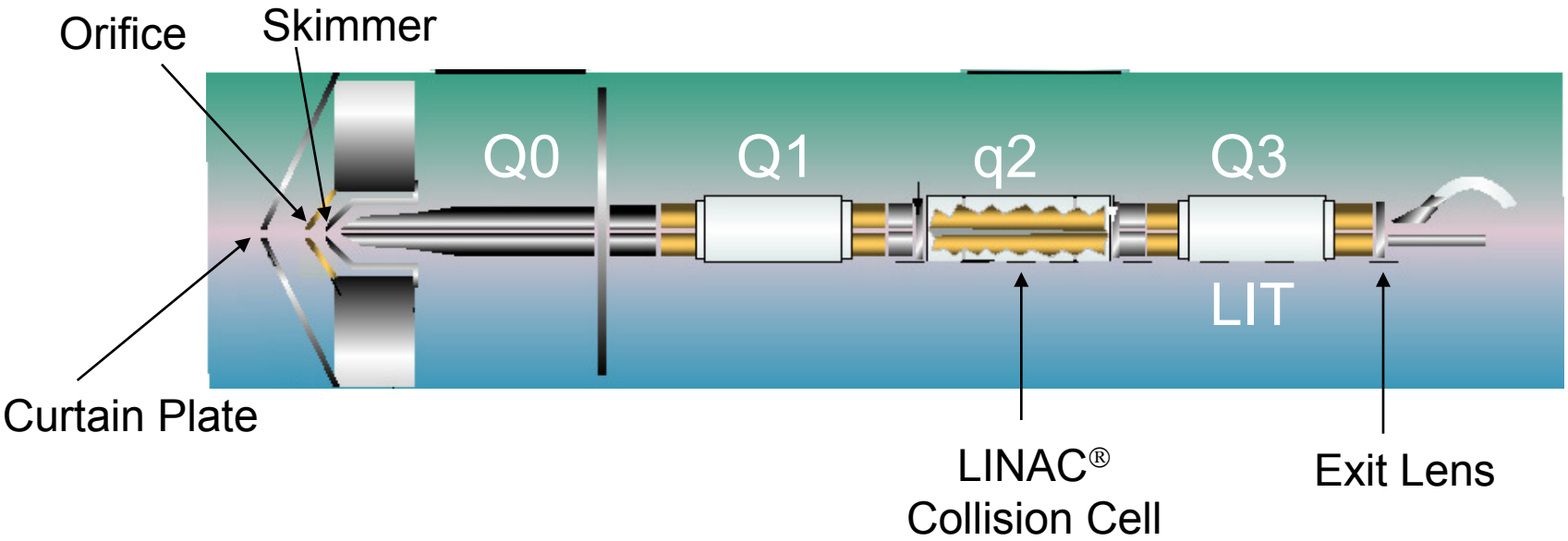


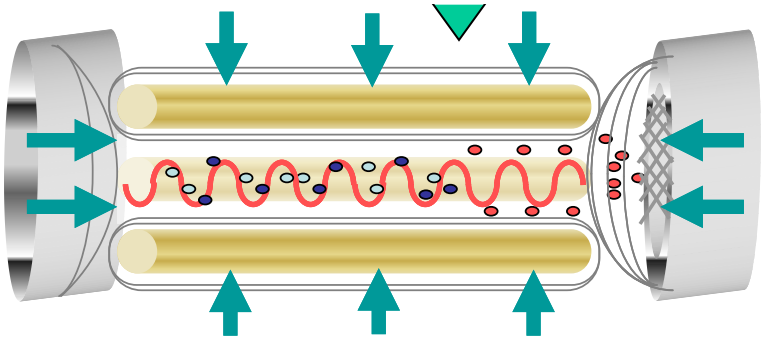
- Limited analysis time for each ion
- Less sample required
- No pre-analysis cleanup
- High sensitivity





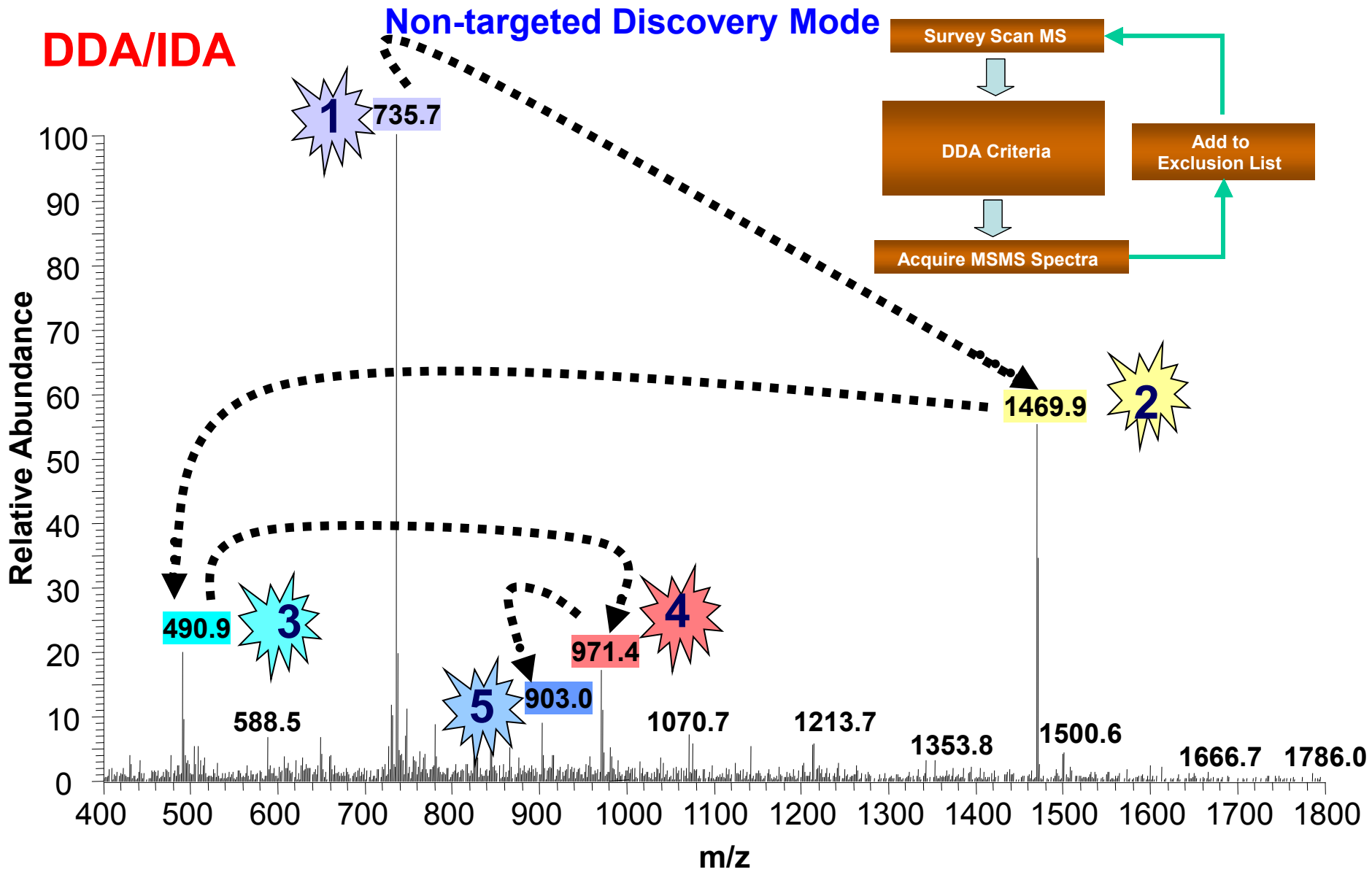
Hybrid Triple Quadrupole/Linear Ion Trap Technology on the 4000 Q TRAP® System





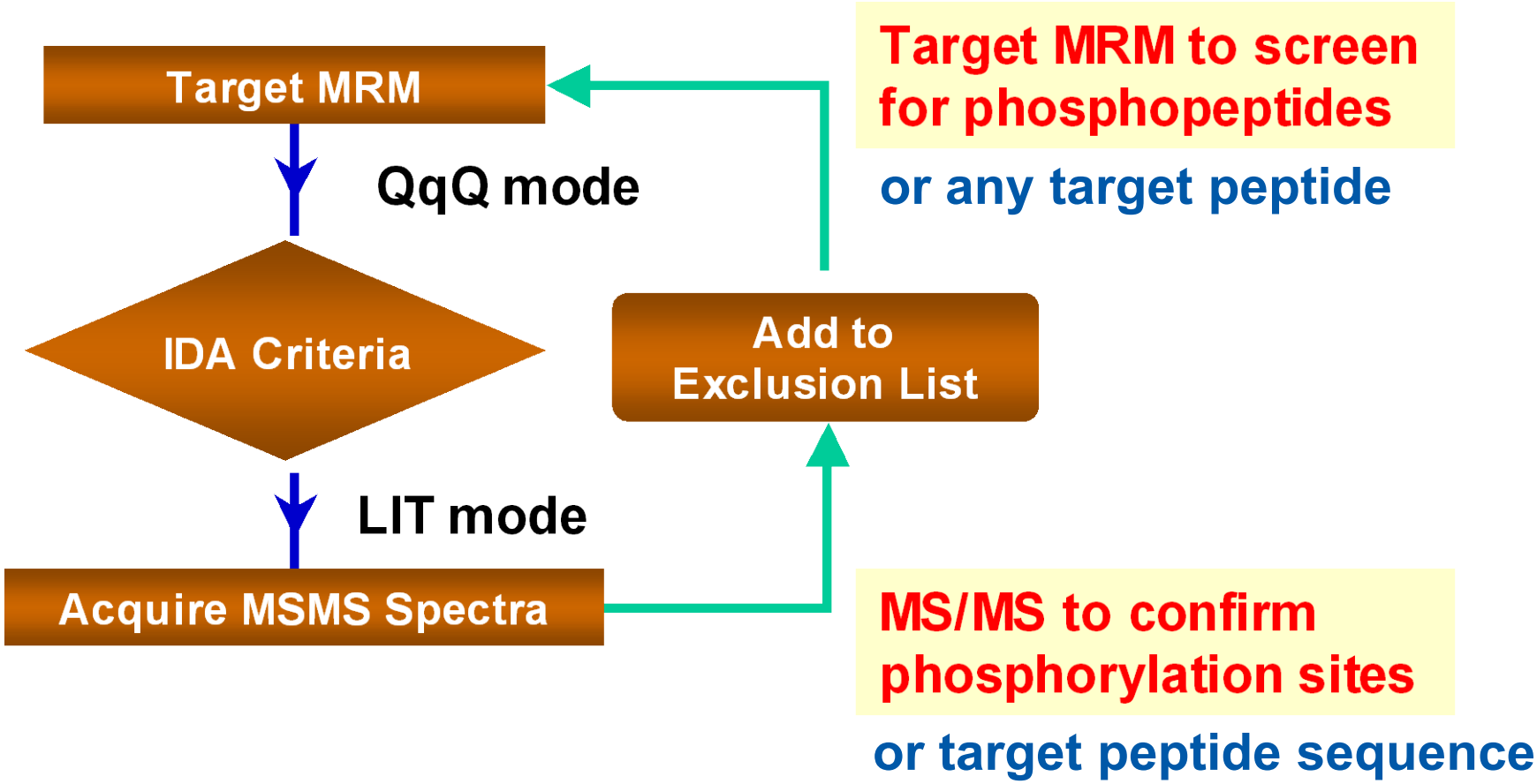
Video Shows a Hybrid Triple Quadrupole/Linear Ion Trap Analyzer on the 4000 Q TRAP® System

Data-Dependent Acquisition & Dynamic Exclusion



Targeted Approaches by MRM-Triggered IDA

- For confirmation of Quantitation Results
- For targeted "fishing" or ID in a sample
  - Quantitative information is a bonus

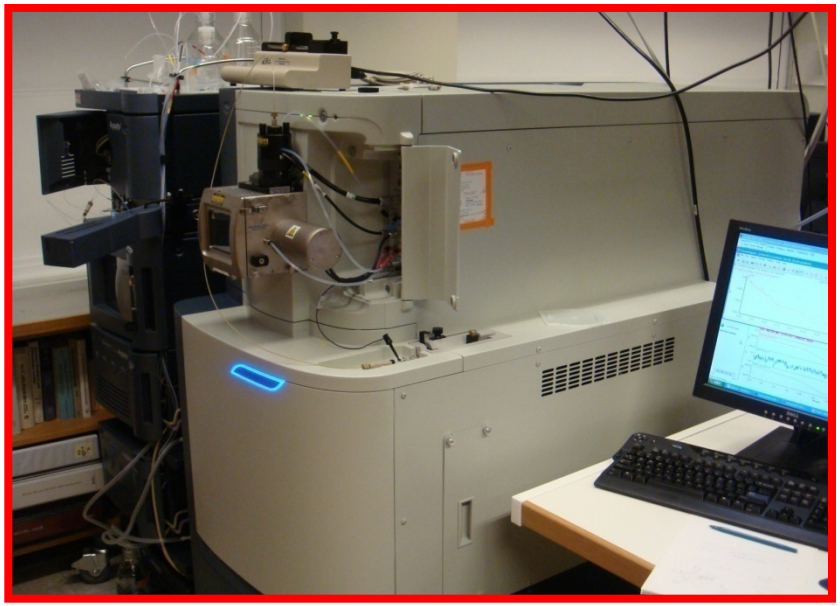


Cornell MS Facility

Facility Instrumentation

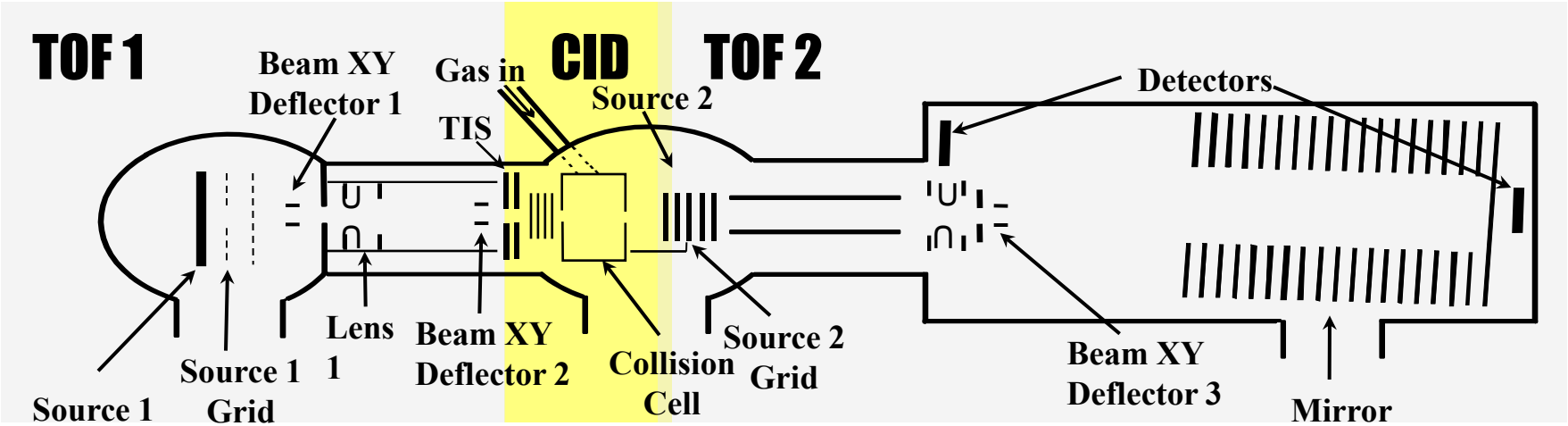
- **Tools for Sample Separation and Preparation**
  - **HPLC system:**
    - 2Dcap Ultimate LC, Agilent 1100, 2DnanoAcquity UPLC
  - **2D Gel System:**
    - Multiphor II IEF + 3 vertical 2nd PAGE: DALT Twelve, Six, SE600
    - Gel image system (2): Typhoon laser scanner, ImageScanner
  - **Gel Spot and Liquid Handling Robots (2)**
    - ProPic & ProPrep
- **Tools for Analytical Analysis**
  - MALDI-TOF/TOF 4700 Analyzer
  - ESI: 4000 Q Trap, Synapt HDMS, Esquire, LTQ-Orbitrap Velos*
- **Bioinformatics Tools (2)**
  - Gel image analysis software: DeCyder 6.5, ImageMaster 6.0
  - Search engines: Mascot 2.2, ProteinPilot 2.0, PLGS 2.3
  - Others: MarkerView 1.2, MRMPilot 1.0, MRMQuant 1.0
  - LIMS: for data storage, extracted data output and report

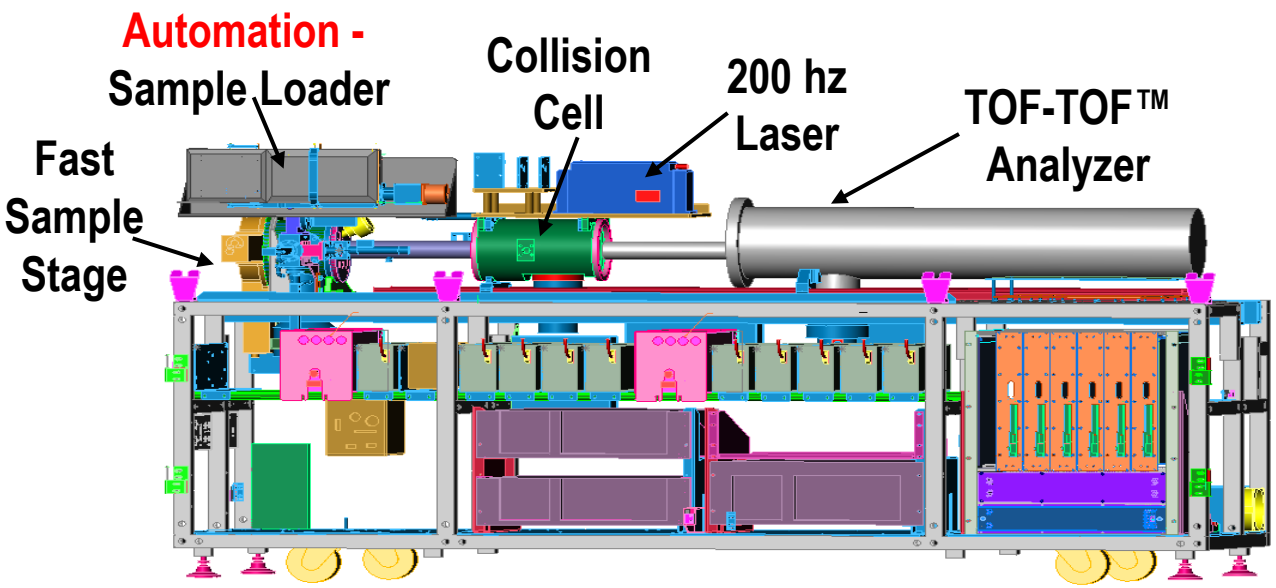
UltiMate 3000
3,700
732
R-83
5,000
1050
254



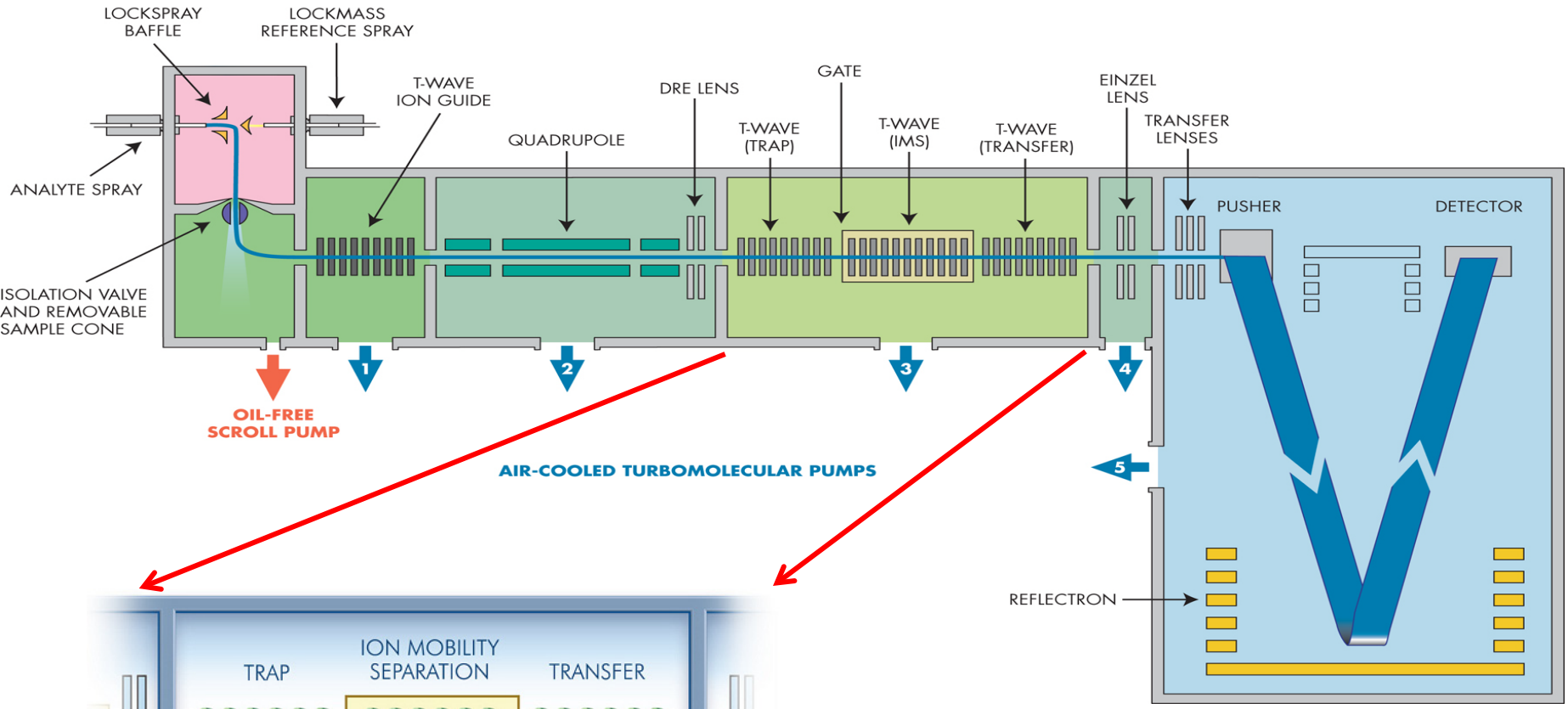
Applied Biosystems

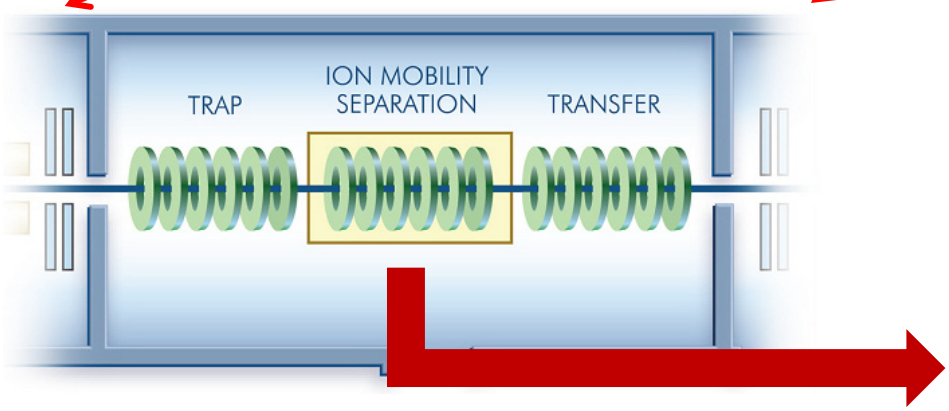
MALDI-TOF/TOF 4700 Proteomics Analyzer

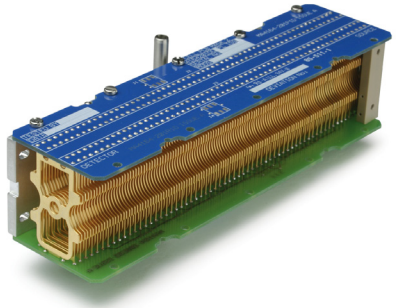




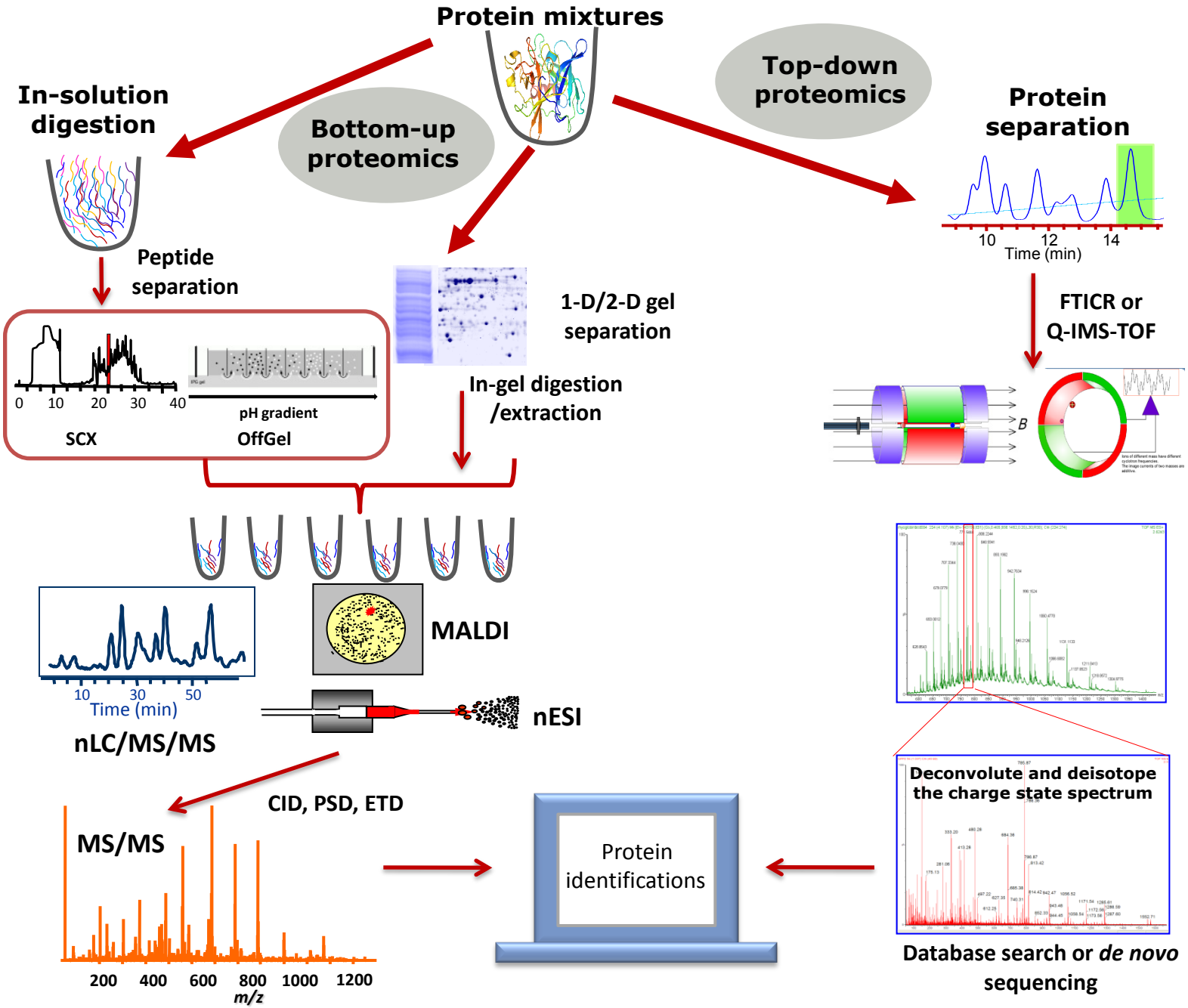
Synapt HDMS System
Hybrid Quadrupole / IMS / oa-ToF



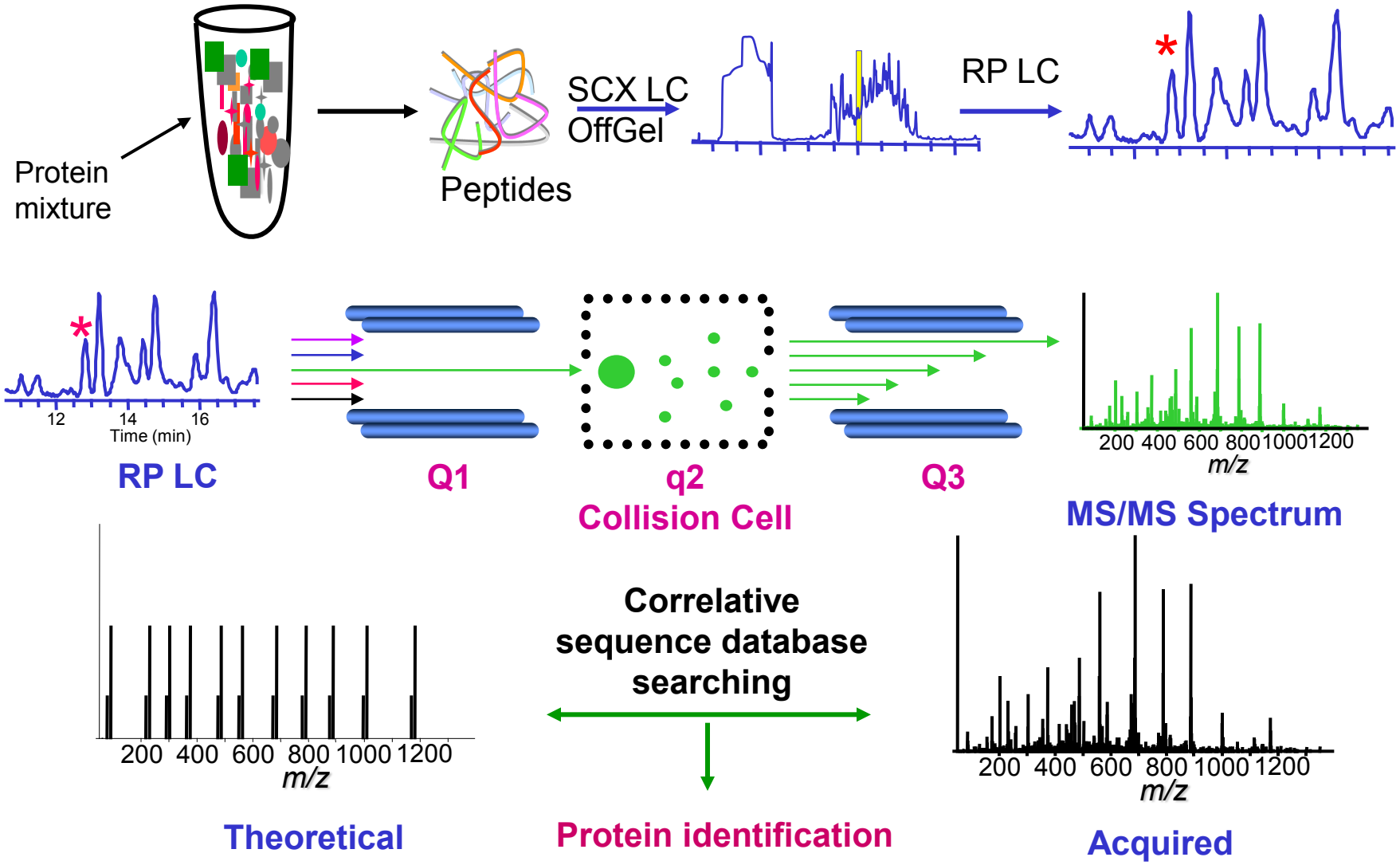




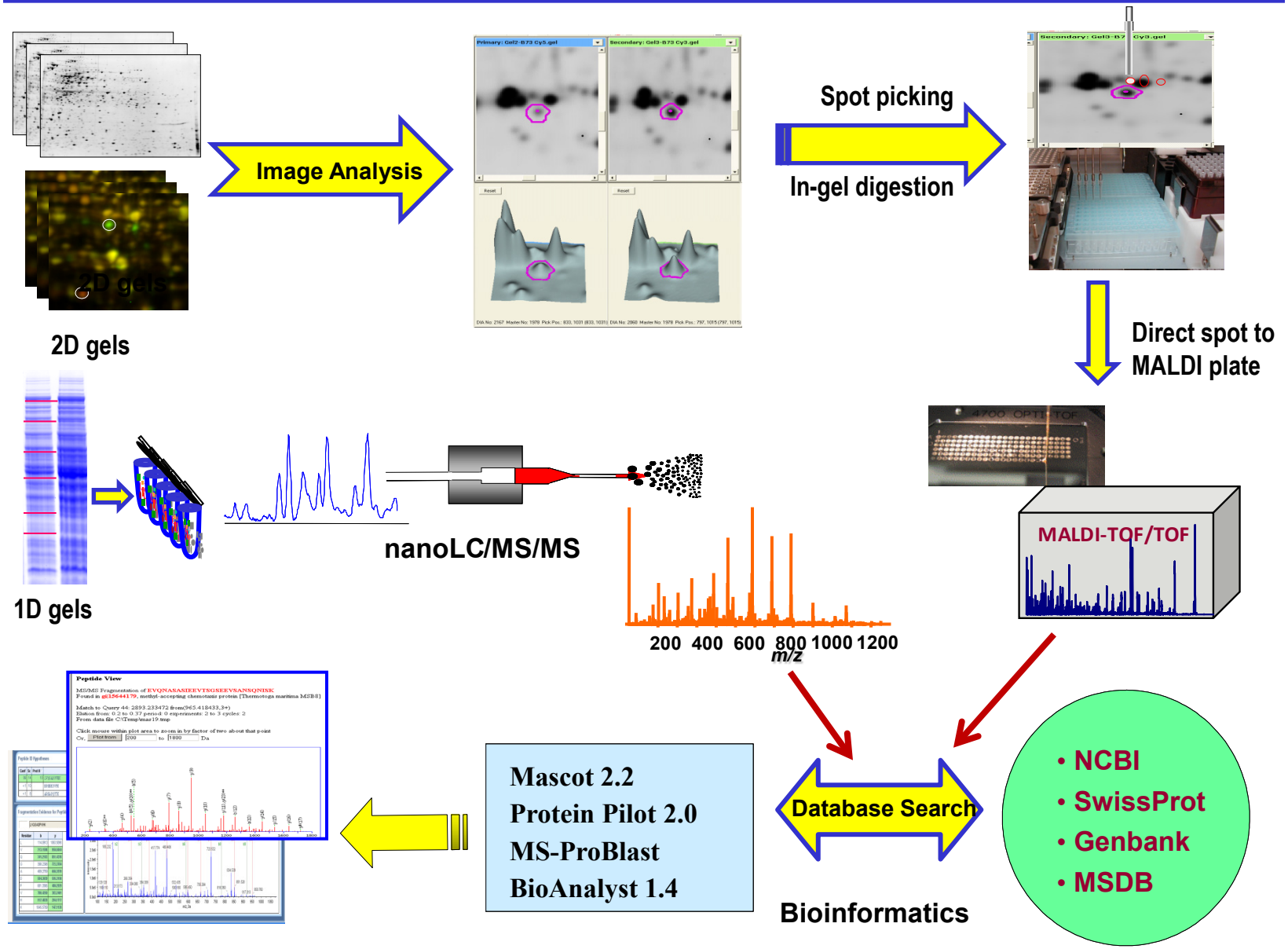
Overview of Proteomics Workflows



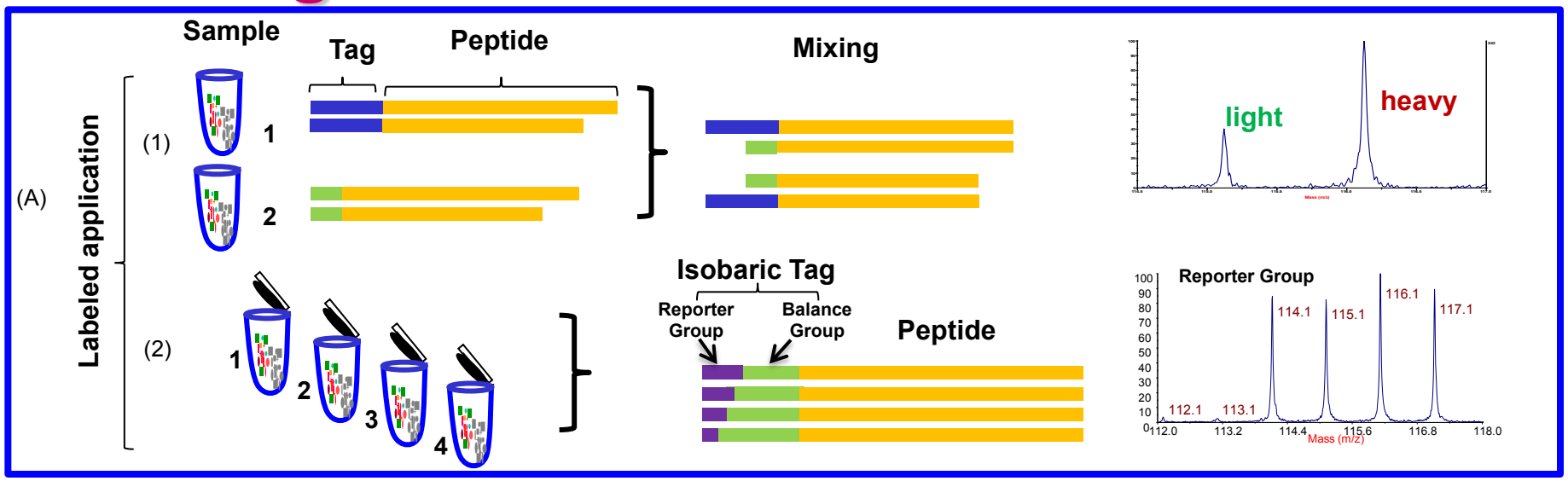
"Shotgun" Protein Identification by LC

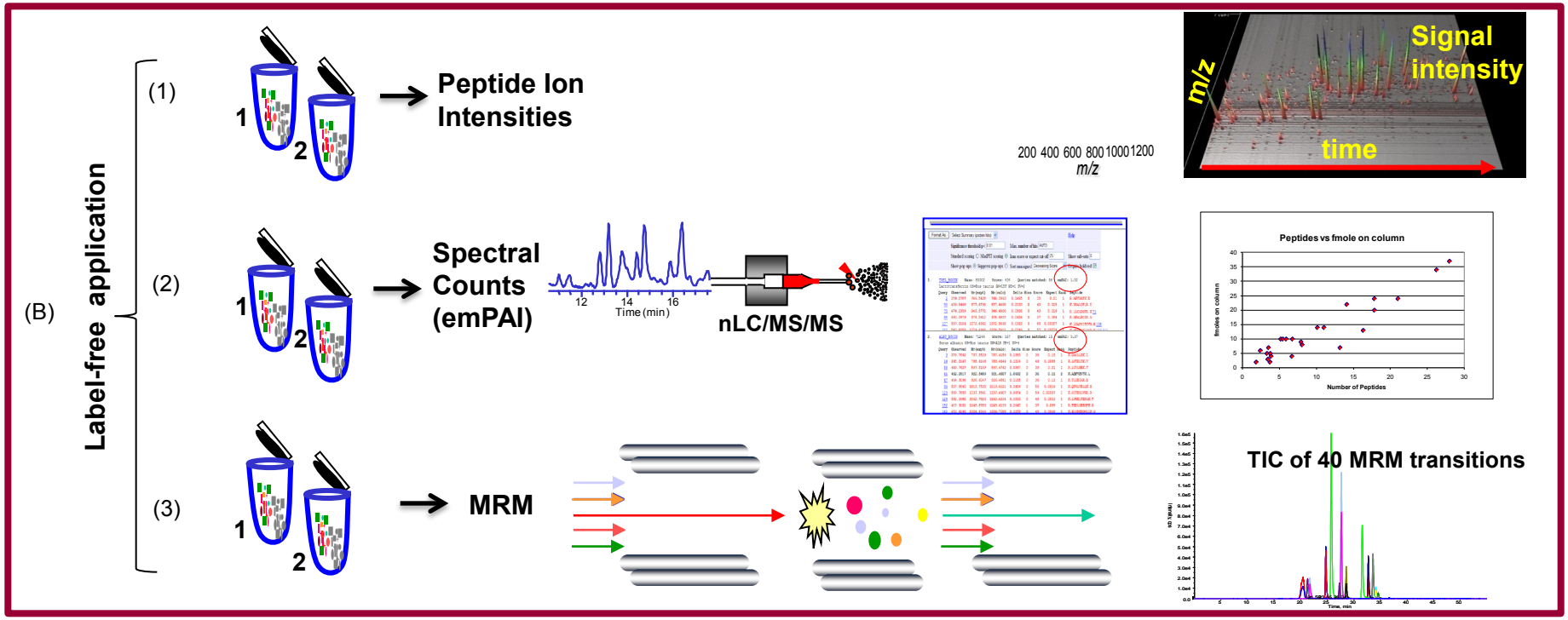


Gel-based Proteomics Workflow

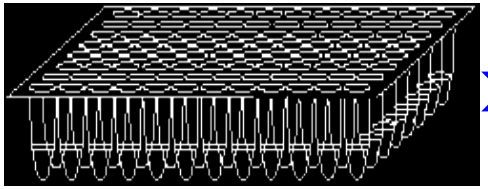


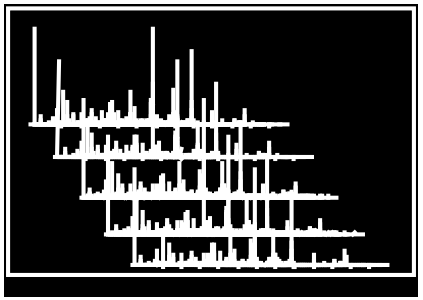
Shotgun-based Quantitative Proteomics

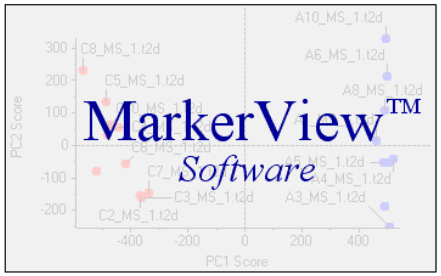




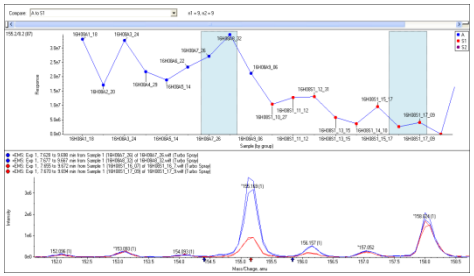
Small Molecule Profiling Workflow

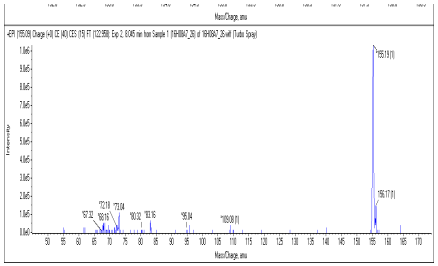


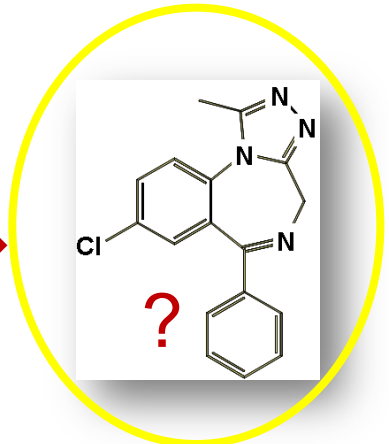




- Generate peaks files
- Normalize and align
- Data analysis:
  - ✓ T-test
  - ✓ PCA
  - ✓ PCA-DA







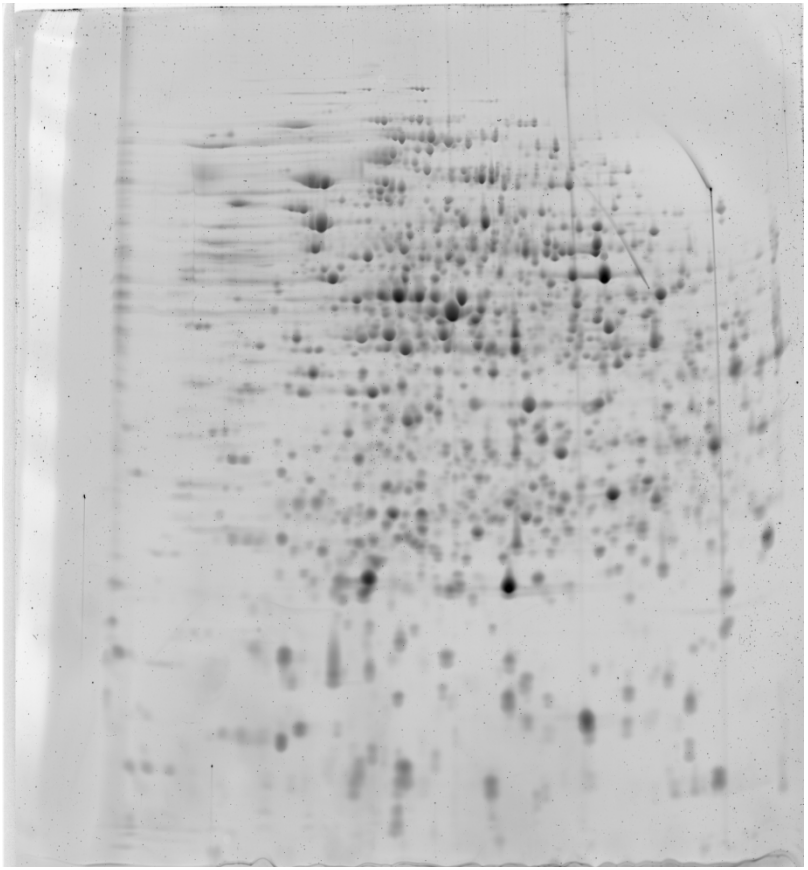
MS Applications to Proteomics and Protein Characterization

- 2D gel protein expression profiling
- Protein IDs by discovery and targeted techniques
- Quantitative proteomics for biomarker discovery
- Characterization of protein phosphorylation
- Peptide mapping for N- /C-terminal determination
- Protein complexes: protein-protein interactions

Comparison of 2D Gels of *E coli* Extracts (100 µg)

pH 4 → 7

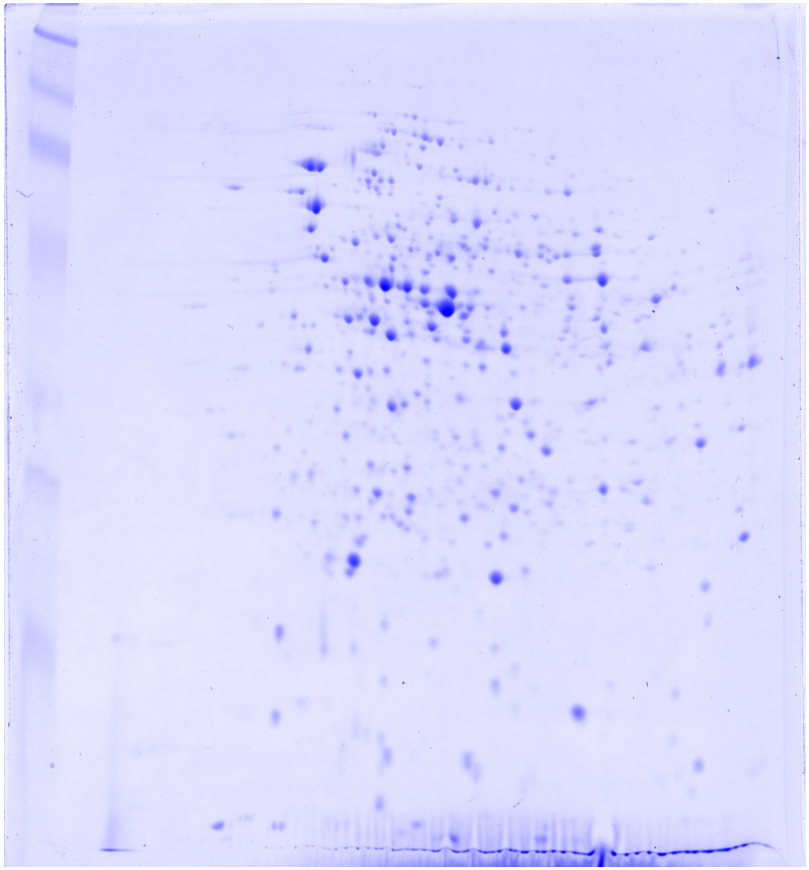
Sypro Ruby



1066 spots

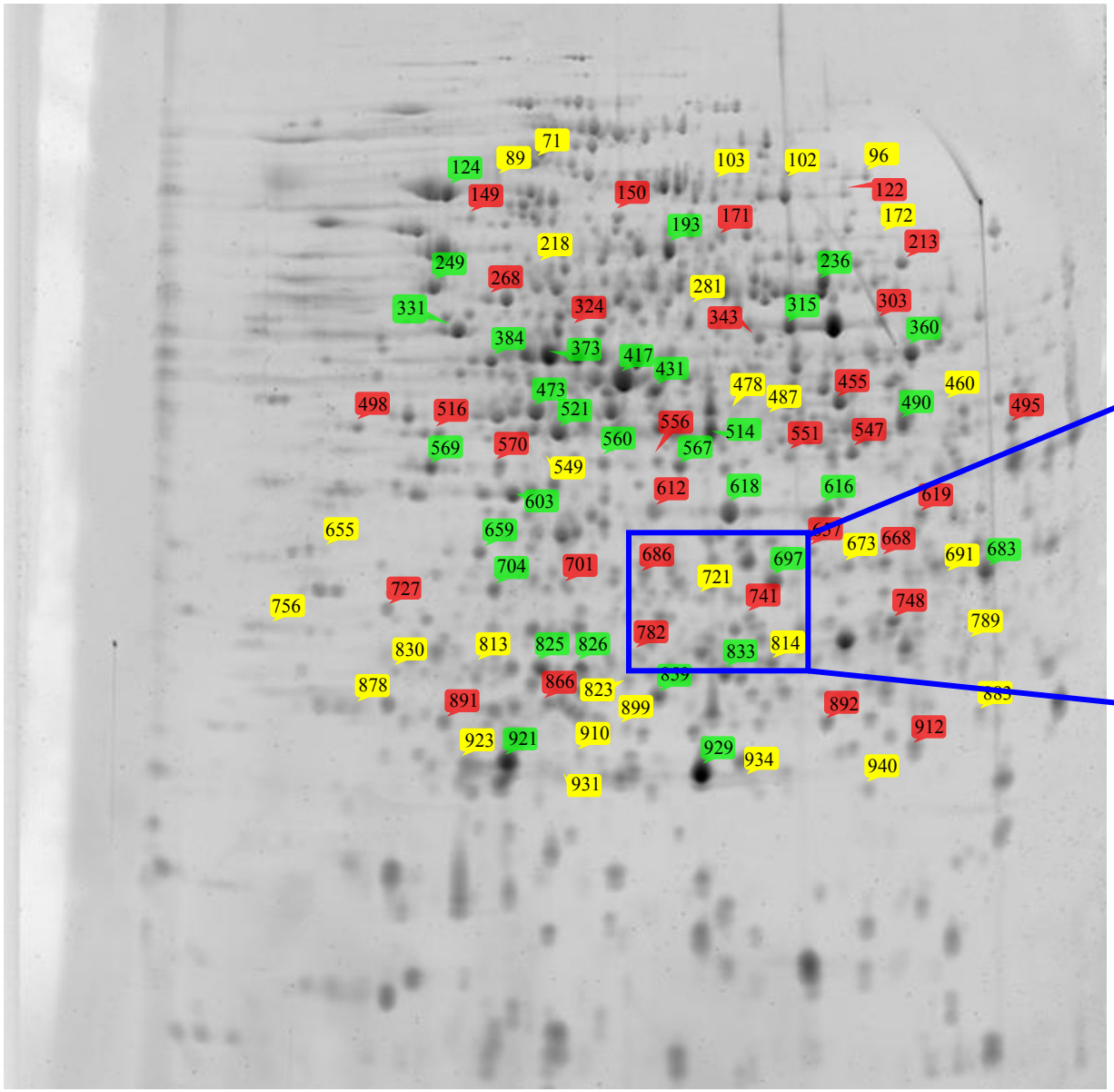
pH 4 → 7

Colloidal Coomassie Blue



626 spots

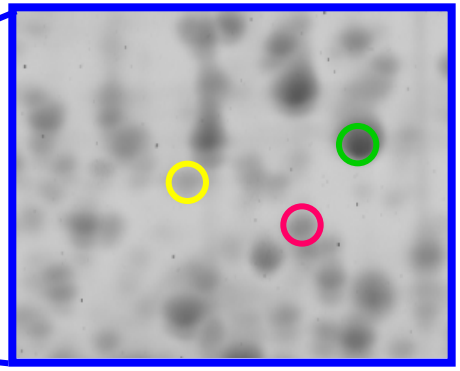
E coli Spots Selected for Robotic Picking up on Sypro Ruby Stained Gel

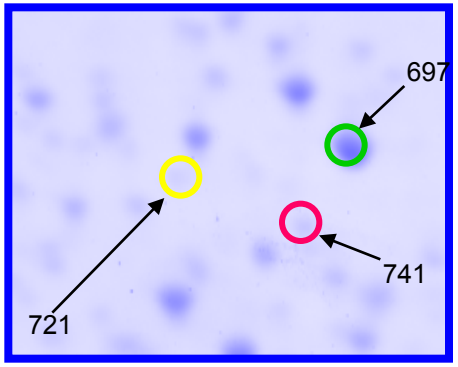


High intensity spots

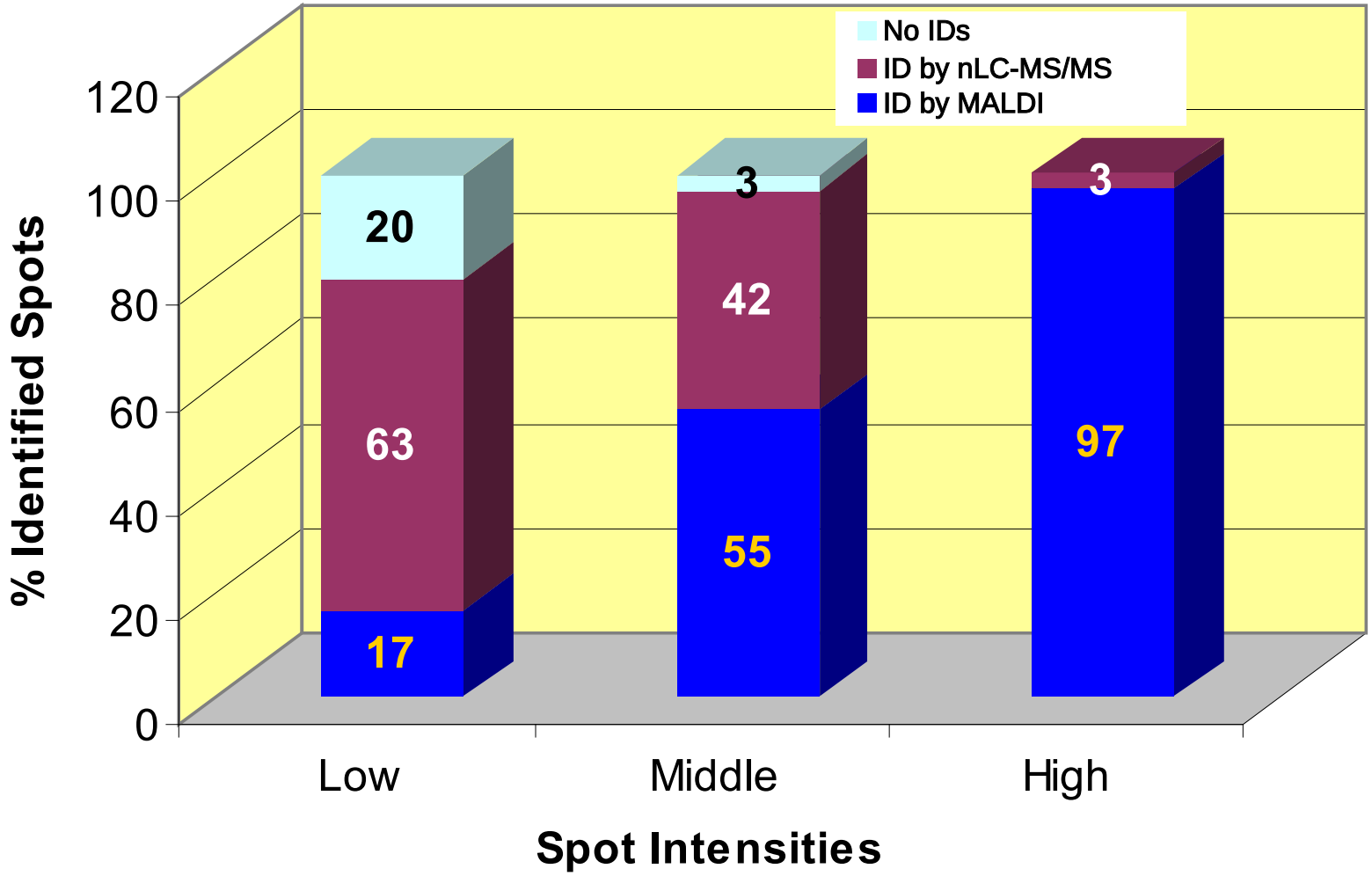
Middle intensity spots
(Barely visible in CCB)

Low intensity spots
(no detected in CCB)

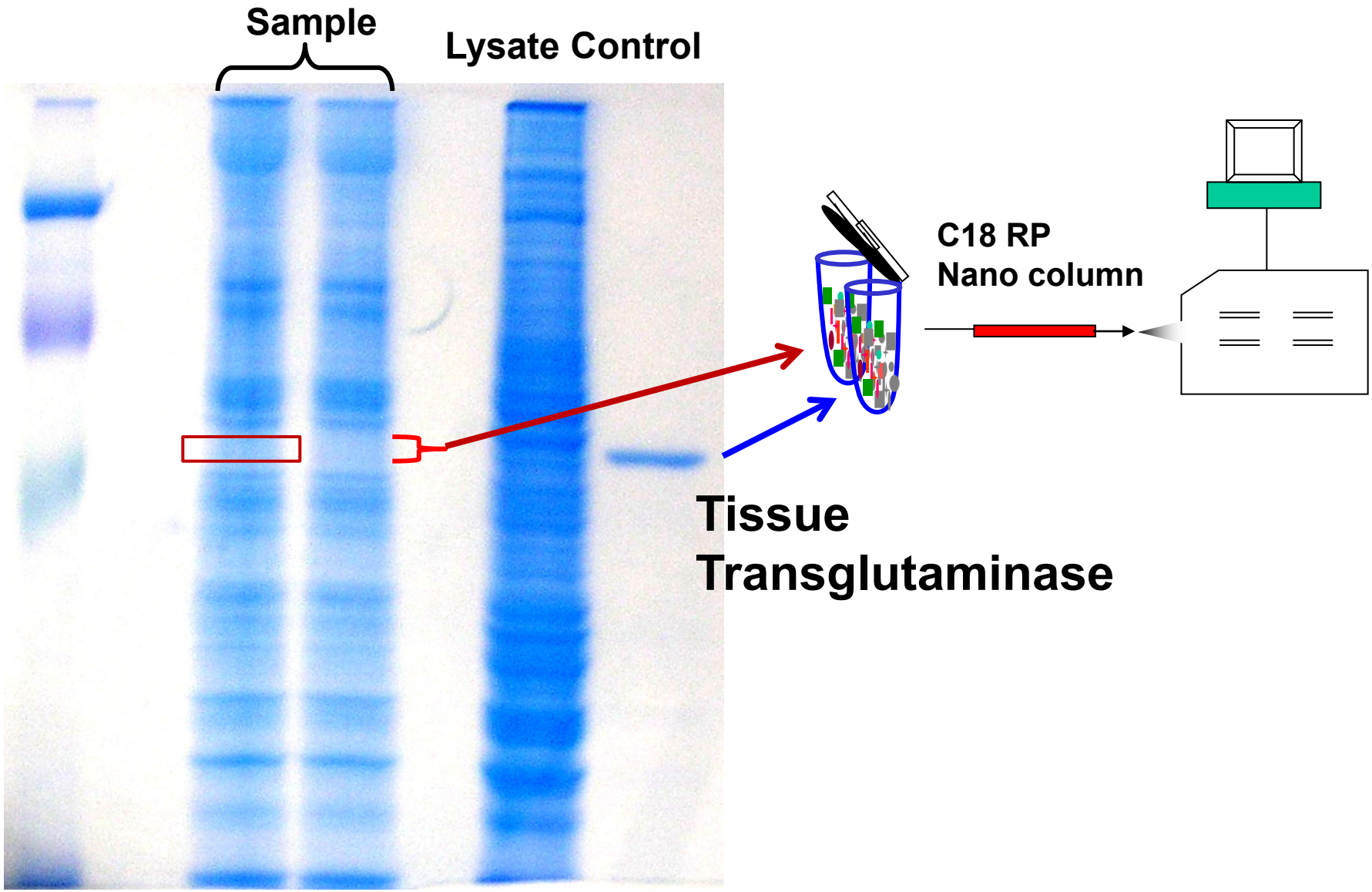




Spot Intensity on Protein ID Rates for Sypro Ruby Stained 2D-gel



Protein IDs by Non-targeted & Targeted Modes



No Targeted Protein Identified by DDA

```
User                    : Mascot Daemon
Email                   : sz14@cornell.edu
Search title            : NCBInrHuman (BoLi10198869re_08)
MS data file            : \\phoenix.biotech.cornell.edu\GROUP_SHARE\MS\Sheng backup\Analyst data since 24Jul06\Projects\2009_07_16Data\10Aug09\BoLi10198869re_8.wiff
Database                : NCBInr 20080728 (6833826 sequences; 2363426297 residues)
Taxonomy                : Homo sapiens (human) (185480 sequences)
Timestamp               : 9 Oct 2009 at 19:28:39 GMT
Enzyme                  : Trypsin
Variable modifications : Carbamidomethyl (C),Oxidation (M)
Mass values             : Monoisotopic
Protein Mass            : Unrestricted
Peptide Mass Tolerance : ± 1.5 Da
Fragment Mass Tolerance: ± 0.6 Da
Max Missed Cleavages    : 1
Instrument type         : Default
Number of queries       : 3314
Protein hits            :   gi|306891      90kDa heat shock protein
                            gi|154146191   heat shock protein 90kDa alpha (cytosolic), class A member 1 isoform 2 [Homo sapiens]
                            gi|11935049    keratin 1 [Homo sapiens]
                            gi|190074      lysyl hydroxylase
                            gi|5031863     galectin 3 binding protein [Homo sapiens]
                            gi|28317       unnamed protein product [Homo sapiens]
                            gi|5453914     phospholipid transfer protein isoform a precursor [Homo sapiens]
                            gi|7144648     oxidized protein hydrolase [Homo sapiens]
                            gi|55956899    keratin 9 [Homo sapiens]
                            gi|7022978     unnamed protein product [Homo sapiens]
                            gi|61744475    solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2 isoform a [Homo sapiens]
                            gi|10863945    ATP-dependent DNA helicase II [Homo sapiens]
                            gi|187387      myristoylated alanine-rich C-kinase substrate
                            gi|115298678   complement component 3 precursor [Homo sapiens]
                            gi|18677767    leucine rich repeat containing 15 [Homo sapiens]
                            gi|547754      Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (CK 2e) (K2e) (keratin-2)
                            gi|2506545     78 kDa glucose-regulated protein precursor (GRP 78) (Heat shock 70 kDa protein 5) (Immunoglobulin heavy chain-binding protein) (BiP)
                            gi|11276938    hypothetical protein DKFZp762H157.1 - human (fragment)
                            gi|29799       CD44R1 [Homo sapiens]
                            gi|38202255    threonyl-tRNA synthetase [Homo sapiens]
                            gi|6137581     Chain A, Crytal Structure Of The Ectodomain Of Human Transferrin Receptor
                            gi|182697      fibronectin
                            gi|4505257     moesin [Homo sapiens]
                            gi|35505       pyruvate kinase [Homo sapiens]
                            gi|6808243     hypothetical protein [Homo sapiens]
                            gi|186685      keratin type 16
                            gi|182855      80K-H protein
                            gi|30030       alpha-1 collagen VI (AA 574-1009) [Homo sapiens]
                            gi|7022784     unnamed protein product [Homo sapiens]
                            gi|4758236     extracellular matrix protein 1 isoform 1 precursor [Homo sapiens]
                            gi|55666251    phosphofructokinase, platelet [Homo sapiens]
                            gi|180392      alpha 1 (I) chain propeptide
                            gi|1381142     Hs-CUL-1
                            gi|5729877     heat shock 70kDa protein 8 isoform 1 [Homo sapiens]
                            gi|11321601    phosphofructokinase, platelet [Homo sapiens]
                            gi|1381150     Hs-CUL-4B
                            gi|5031631     scavenger receptor class B, member 2 [Homo sapiens]
                            gi|119617032   keratin 6B, isoform CRA_a [Homo sapiens]
                            gi|1478281     neutral amino acid transporter B
                            gi|6912482     leucine zipper-EF-hand containing transmembrane protein 1 [Homo sapiens]
                            gi|52545928    hypothetical protein [Homo sapiens]
                            gi|13603394    type VI collagen alpha 2 chain precursor [Homo sapiens]
                            gi|13325075    quiescin Q6 sulfhydryl oxidase 1 isoform a [Homo sapiens]
                            gi|2627129     polyubiquitin [Homo sapiens]
                            gi|6005942     valosin-containing protein [Homo sapiens]
                            gi|5107666     Chain A, Structure Of Importin Beta Bound To The Ibb Domain Of Importin Alpha
                            gi|190207      poliovirus receptor
                            gi|179832      calnexin
                            gi|4505891     procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 precursor [Homo sapiens]
                            gi|40674640    IQGAP1 protein [Homo sapiens]
                            gi|4503483     eukaryotic translation elongation factor 2 [Homo sapiens]
```

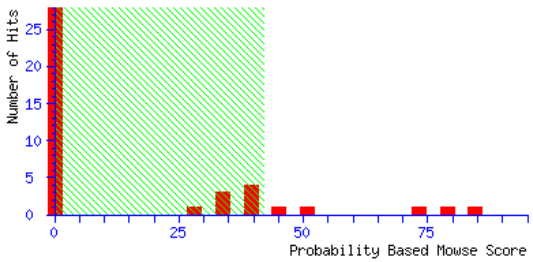
50 protein IDs

Targeted Protein Identified by MRM-Triggered DDA

```
Search title     : NCBInrHuman (BoLi10198869MRMIDA_26)
MS data file     : \\phoenix.biotech.cornell.edu\GROUP_SHARE\MS\Sheng backup\Analyst data since 24Jul06\Projects\2009_07_16Data\04Aug09\BoL
Database         : NCBInr 20080728 (6833826 sequences; 2363426297 residues)
Taxonomy         : Homo sapiens (human) (185480 sequences)
Timestamp        : 9 Oct 2009 at 19:18:59 GMT
Protein hits     : gi|39777597   transglutaminase 2 isoform a [Homo sapiens]
                   gi|177207     4F2 antigen heavy chain
                   gi|7020236    unnamed protein product [Homo sapiens]
                   gi|120659926  Hypothetical LOC284297 [Homo sapiens]
                   gi|6137581    Chain A, Crytal Structure Of The Ectodomain Of Human Transferrin Receptor
```

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 42 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.



Four peptides were detected for the targeted protein

Peptide Summary Report

Format As [Peptide Summary] Help

Significance threshold p< 0.05 Max. number of hits AUTO

Standard scoring ○ MudPIT scoring ◉ Ions score or expect cut-off 20 Show sub-sets 0

Show pop-ups ◉ Suppress pop-ups ○ Sort unassigned [Decreasing Score] Require bold red ☐

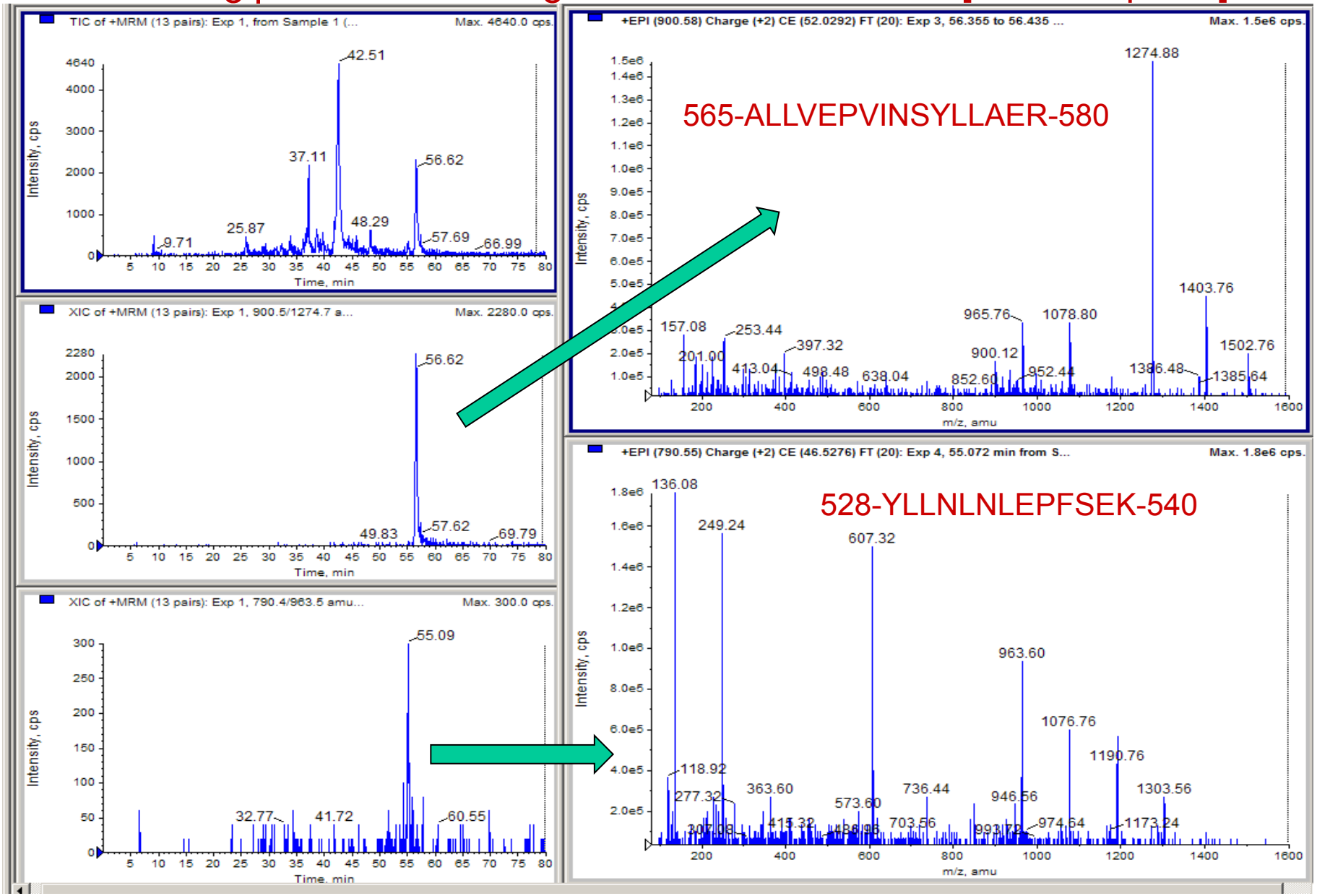
Select All | Select None | Search Selected | ☐ **Error tolerant** | Archive Report

1. gi|39777597 **Mass:** 77280 **Score:** 85 **Queries matched:** 4 **emPAI:** 0.15
transglutaminase 2 isoform a [Homo sapiens]
☐ Check to include this hit in error tolerant search or archive report

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
☑	146	790.4495	1578.8844	1578.8293	0.0550	0	68	0.00015	1	K.YLLNLNLEPFSEK.S
☑	162	806.4242	1610.8338	1610.8039	0.0299	0	22	5.7	1	K.TVEIPDPVEAGEEVK.V
☑	210	900.5830	1799.1515	1799.0192	0.1322	0	35	0.28	1	R.ALLVEPVINSYLLAER.D
☑	322	833.4035	2497.1886	2496.2033	0.9853	0	32	0.39	1	R.VVTNYNSAHDQNSNLLIEYFR.N

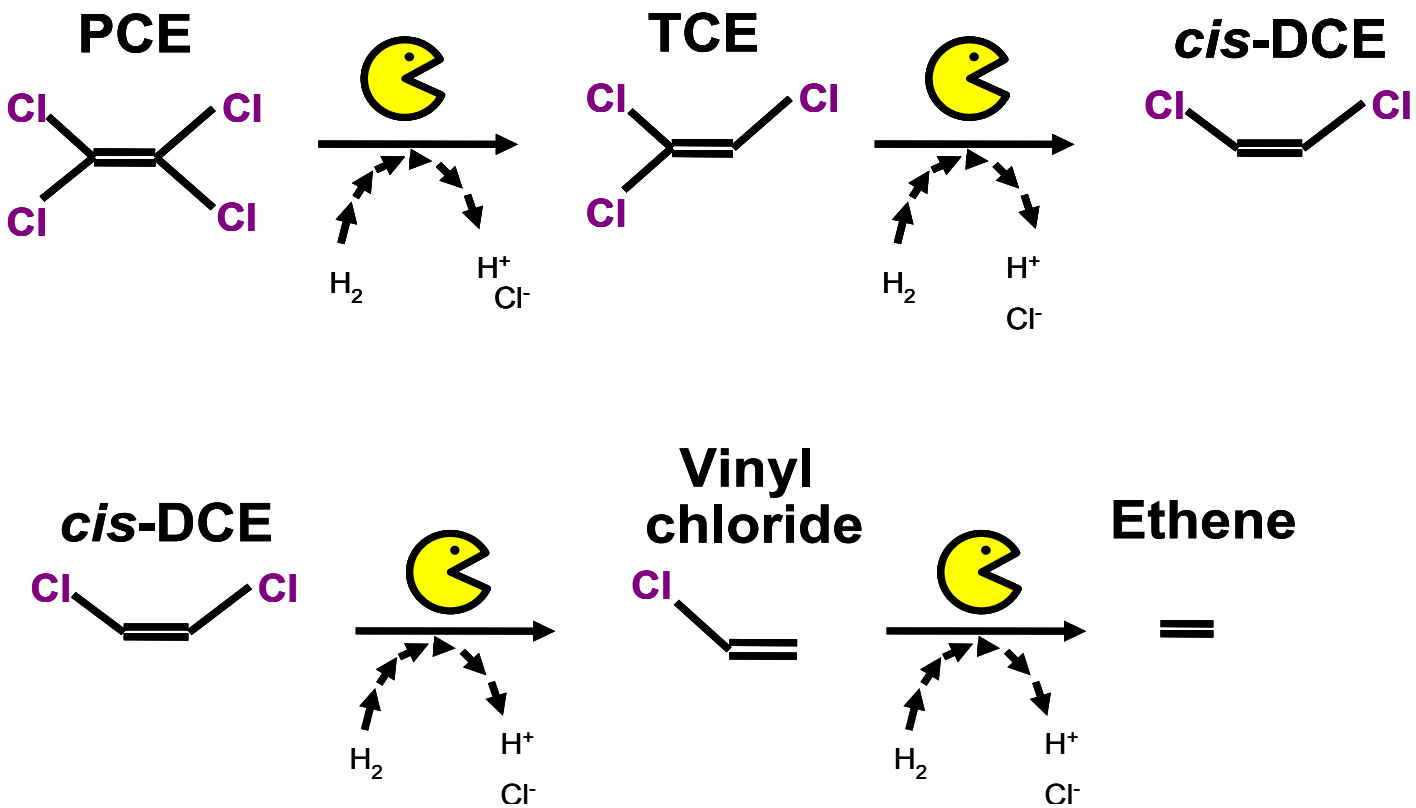
MRM-IDA Results:

Match to: gi|39777597 transglutaminase 2 isoform a [Homo sapiens]



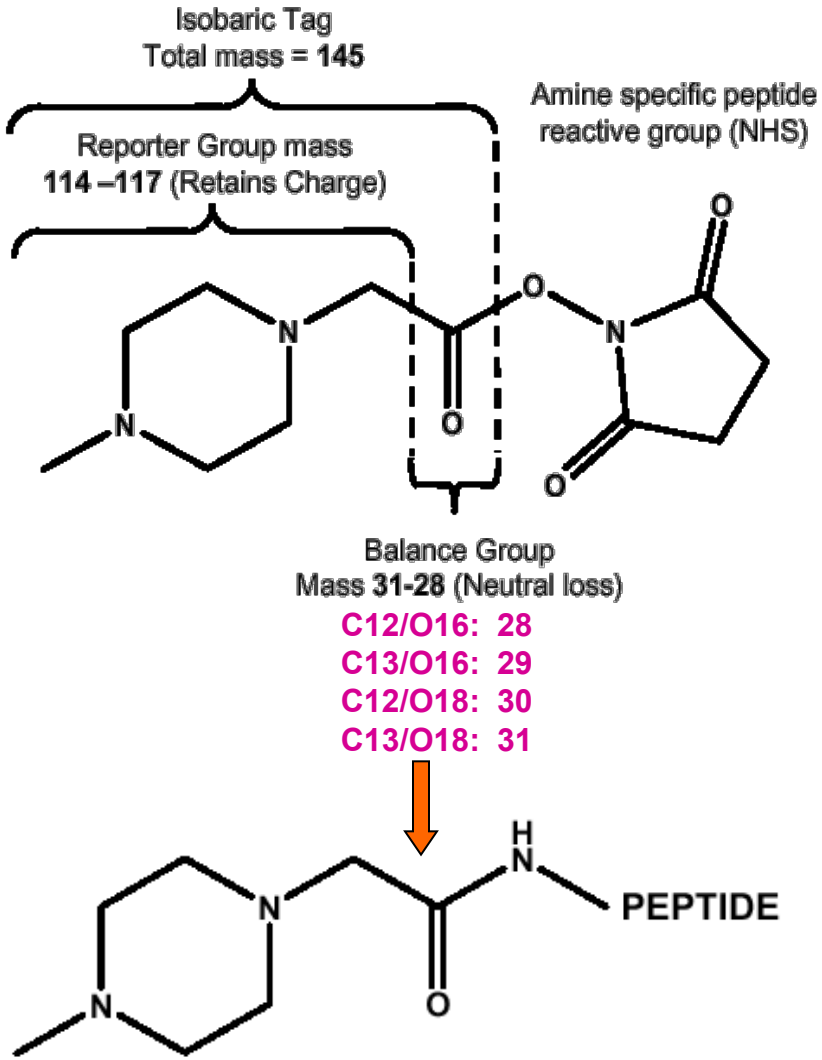
Protein Idenfications

✓ **GeLCMS/MS for Protein IDs of *Dehalococcoides* 195 membrane extraction**



Dehalococcoides spp. are the only organisms known to do all these steps metabolically.

Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) analysis

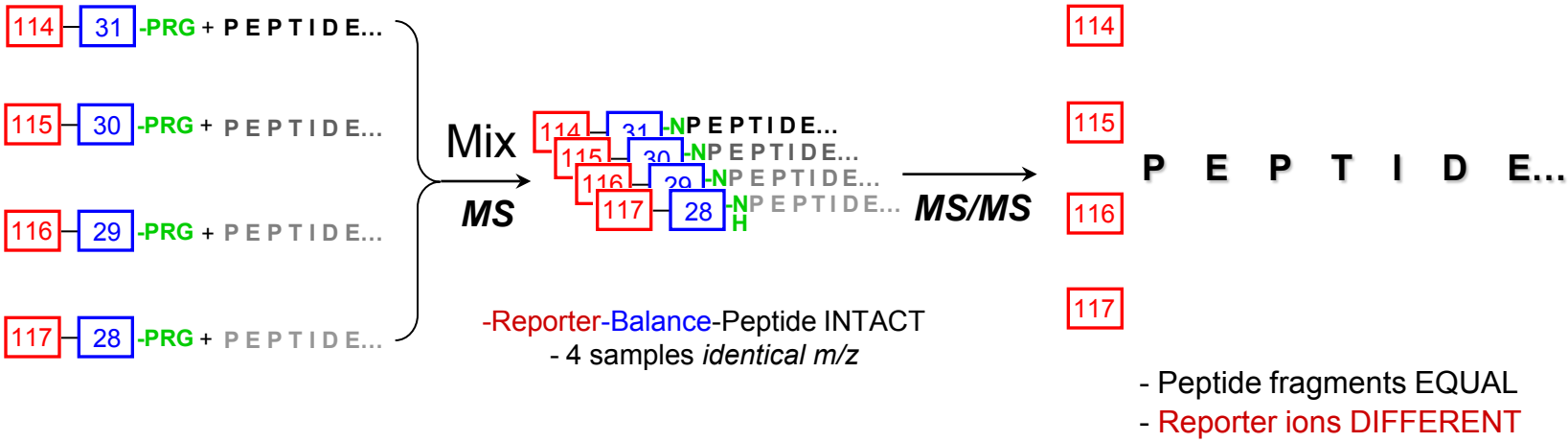


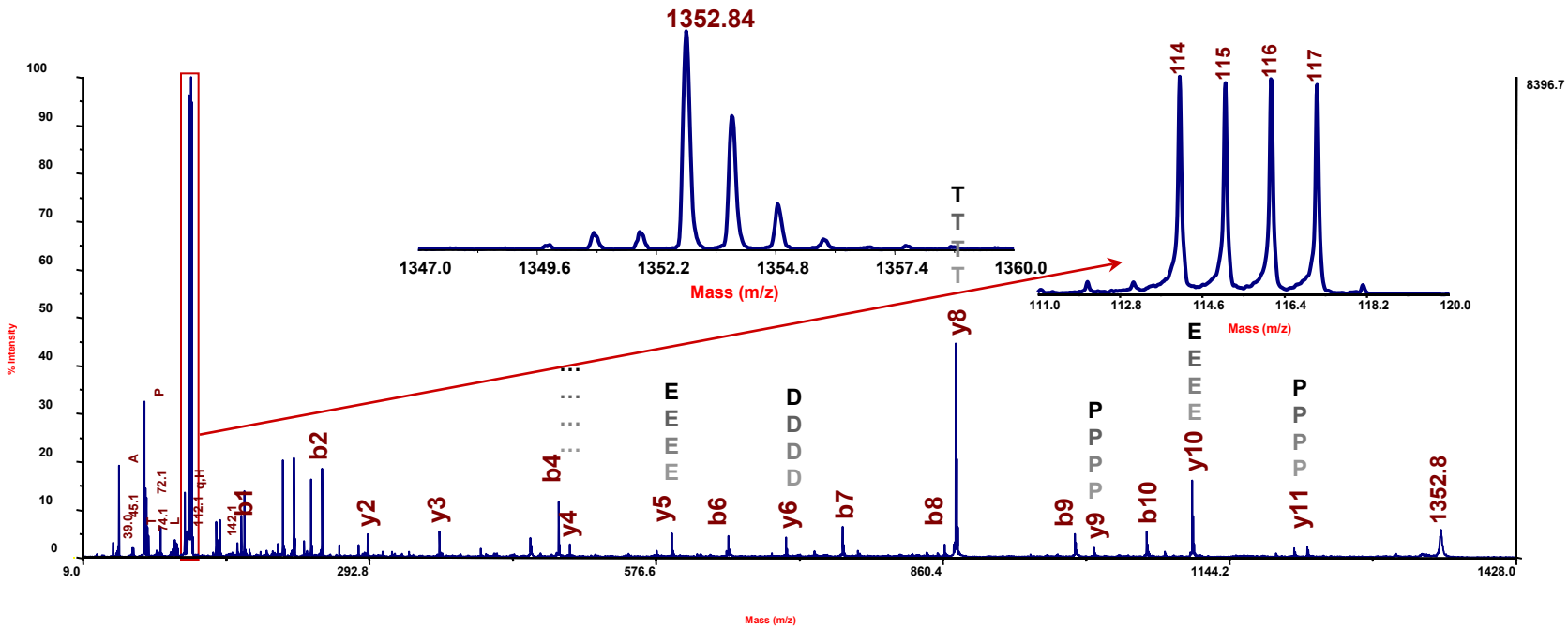
- Give strong MSMS signature ion masses at 114-117
- Generate good y and b-ion series
- Maintain ionization efficiency of labeled peptides by remaining charge state

- Balance the mass change of reporter
- Neutral loss in MSMS giving signature ions

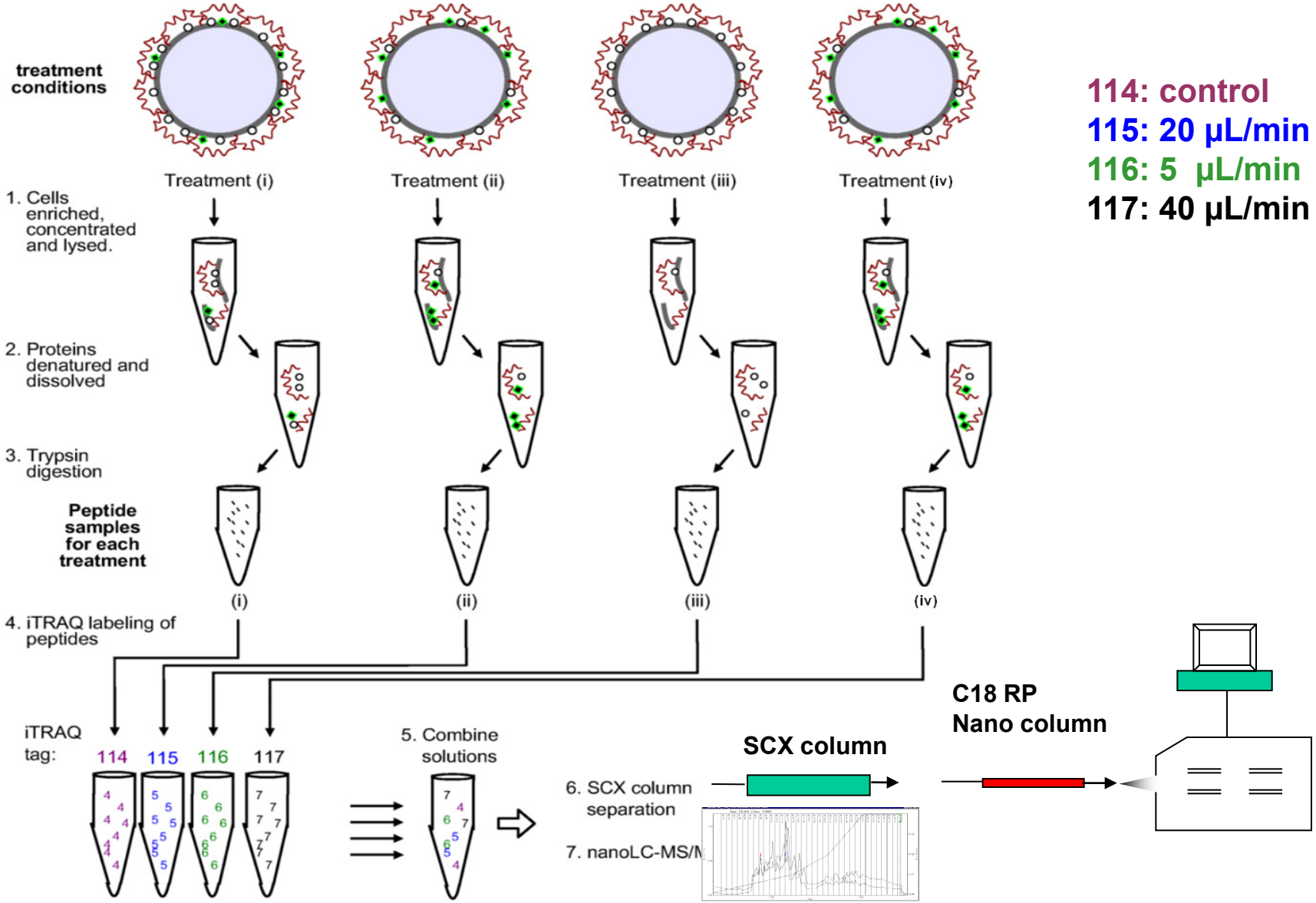
Ross PL *et al.* (2004) Mol. Cell. Proteomics. 3(12), 1154-1169

iTRAQ™ Reagent for Multiplexed Isobaric Tagging

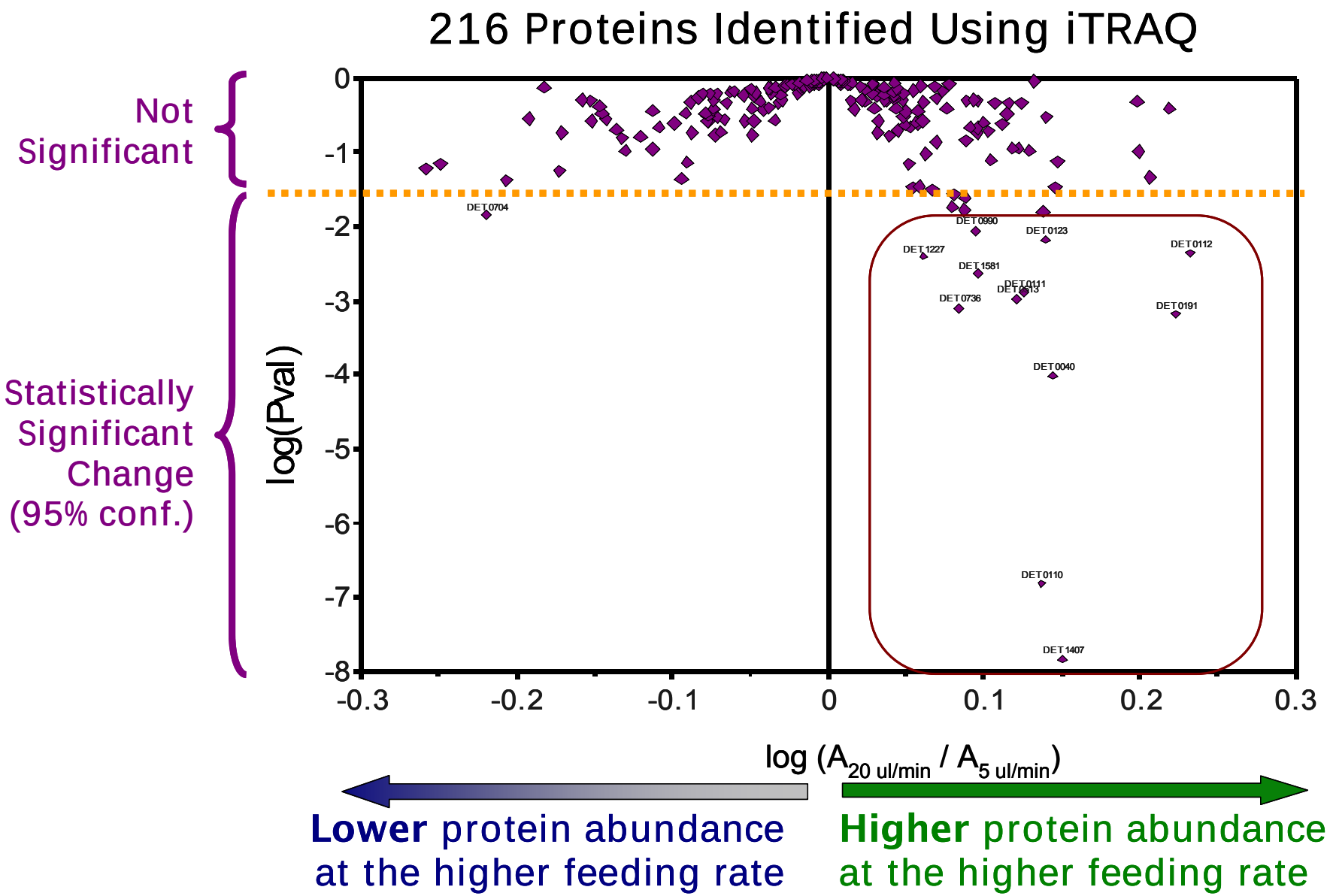




iTRAQ to measure protein changes of culture grown at slow and moderate PCE feeding rates



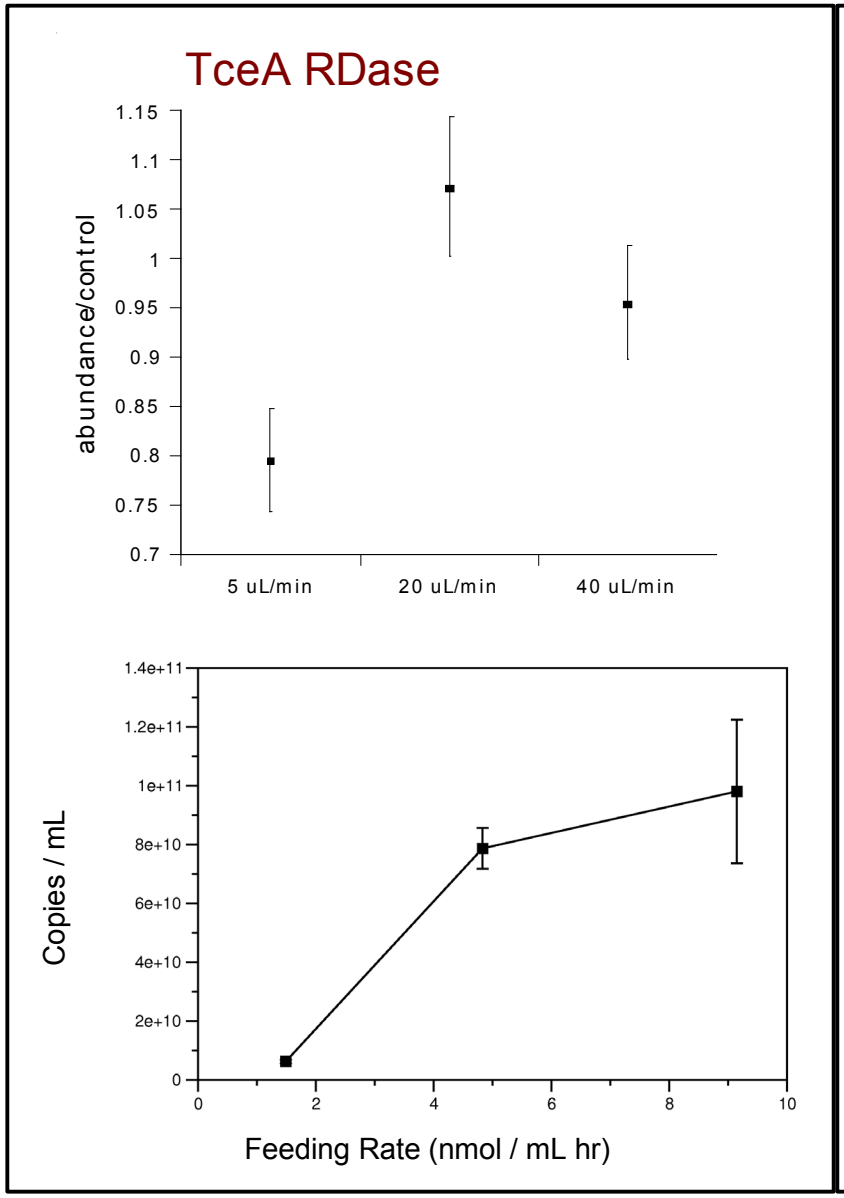
Dehalococcoides Mixed Culture Grown at Slow and Moderate PCE Feeding Rates (5µL/min, tag 116 vs 20µL/min, tag 115)

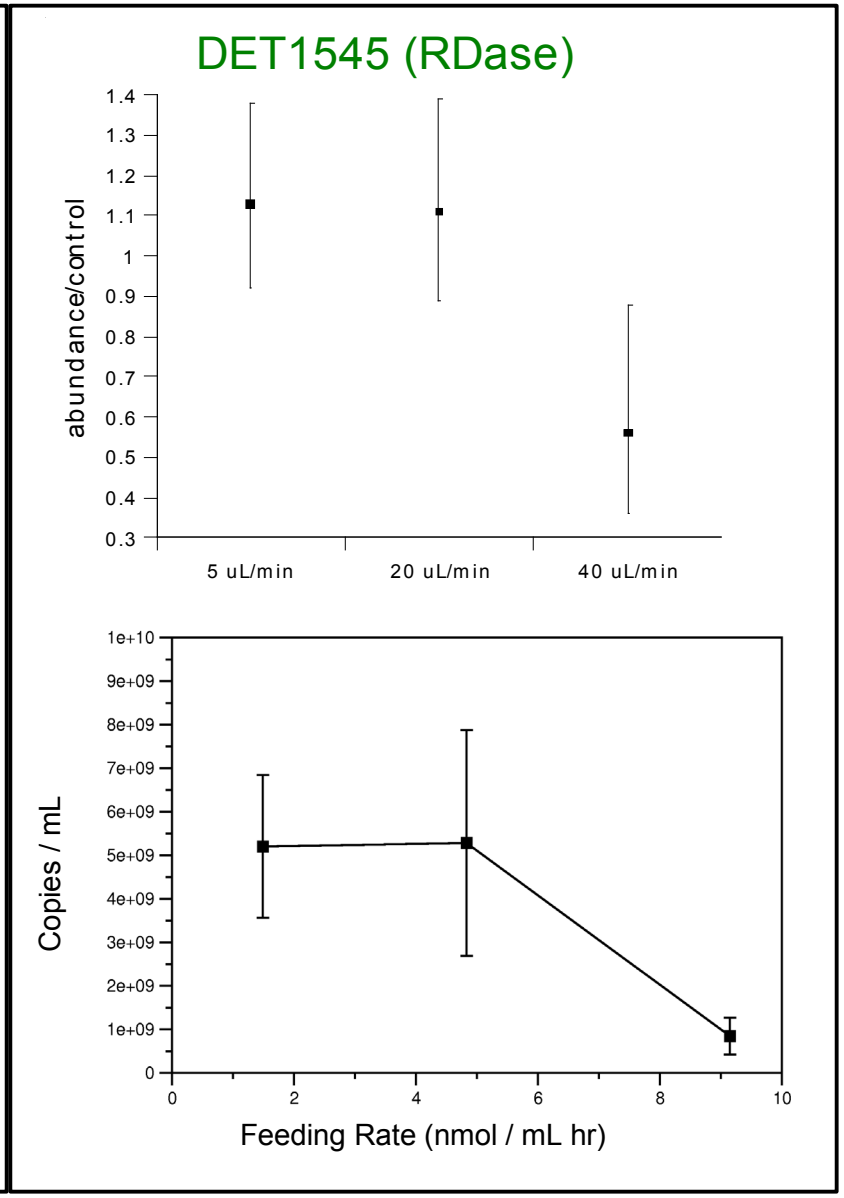


Comparison of RNA Expression and Protein expression Profiling in Response to Different PCE Feeding Rates

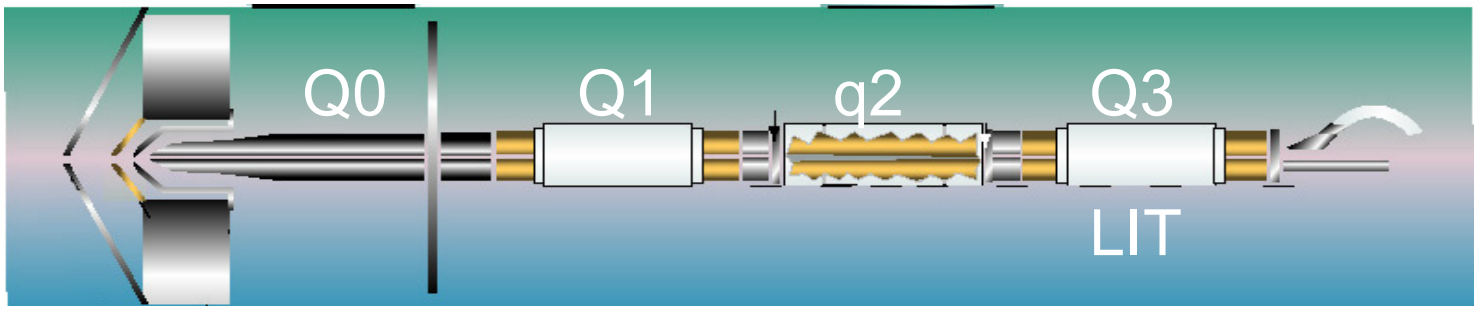
Protein

RNA





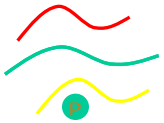
MRM: absolute quant of a specific protein in a complex sample



phosphopeptide

M^{2+} =784.0

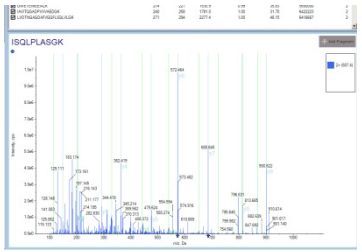
fragmentation



loss of phosphate, product ions

M^{2+} =735.0 M^{+}=847.0

MRMPilot 1.0

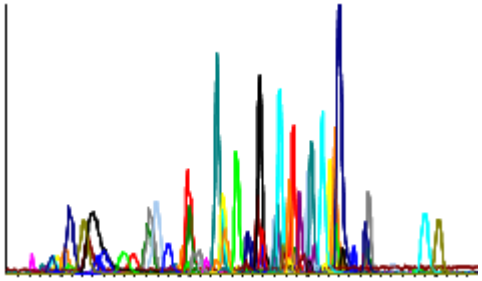


MGRKKIQITRIMDERNRQVTFTKRKFGLMKKAYELS
VLCDCEIALIIFNSSNKLFQYASTDMDKVLLKYTEYN
EPHESRTNSDIVEALNKKEHRGCDSPDPDTSYVLT
PHTEEKYKKINEEFDNMMRNHKIAPGLPPQNFSMS
VTVPVTSPNALSYTNPGSSLVSPSLAASSTLTDSSM
LSPPQTTLHRNVSPGAPQRPPSTGNAGGMLSTTD
LTVPNGAGSSPVGNGFVNSRASPNLIGATGANSLG
KVMPTKSPPPPGGGNLGMNSRKPDLRVVIPPSSK
GMMPPLSEEEELELNTQRISSSQATQPLATPVVSV
TTPSLPPQGLVYSAMPTAYNTDYSLTSADLSALQG
FNSPGMLSLGQVSAWQQHHLGQAALSSLVAGGQL
SQGSNLSINTNQNISIKSEPISPPRDRMTPSGFQQQ

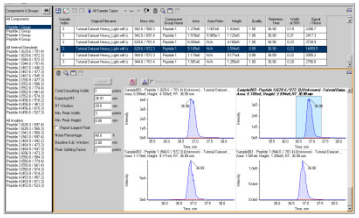
sequence

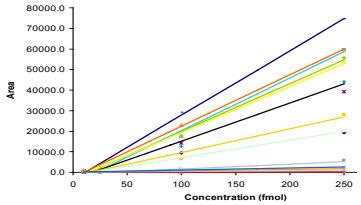
Q1	Q3	Sequence
537.3	488.3	NRQVTFTK
545.2	496.2	MRMDAWVT
571.8	522.8	NFIAVSAANR
617.3	568.3	SEPISPPRDR
635.8	586.8	KNFIAVSAANR
642.3	593.3	TNSDIVEALNK
651.8	602.8	ISSSGALDDDDK
677.8	628.8	IQITRIMDER
699.8	650.8	LFQYASTDMDK
706.3	657.3	TNSDIVEALNKK
709.4	660.4	NFIAVSAANRFK

MRM transitions

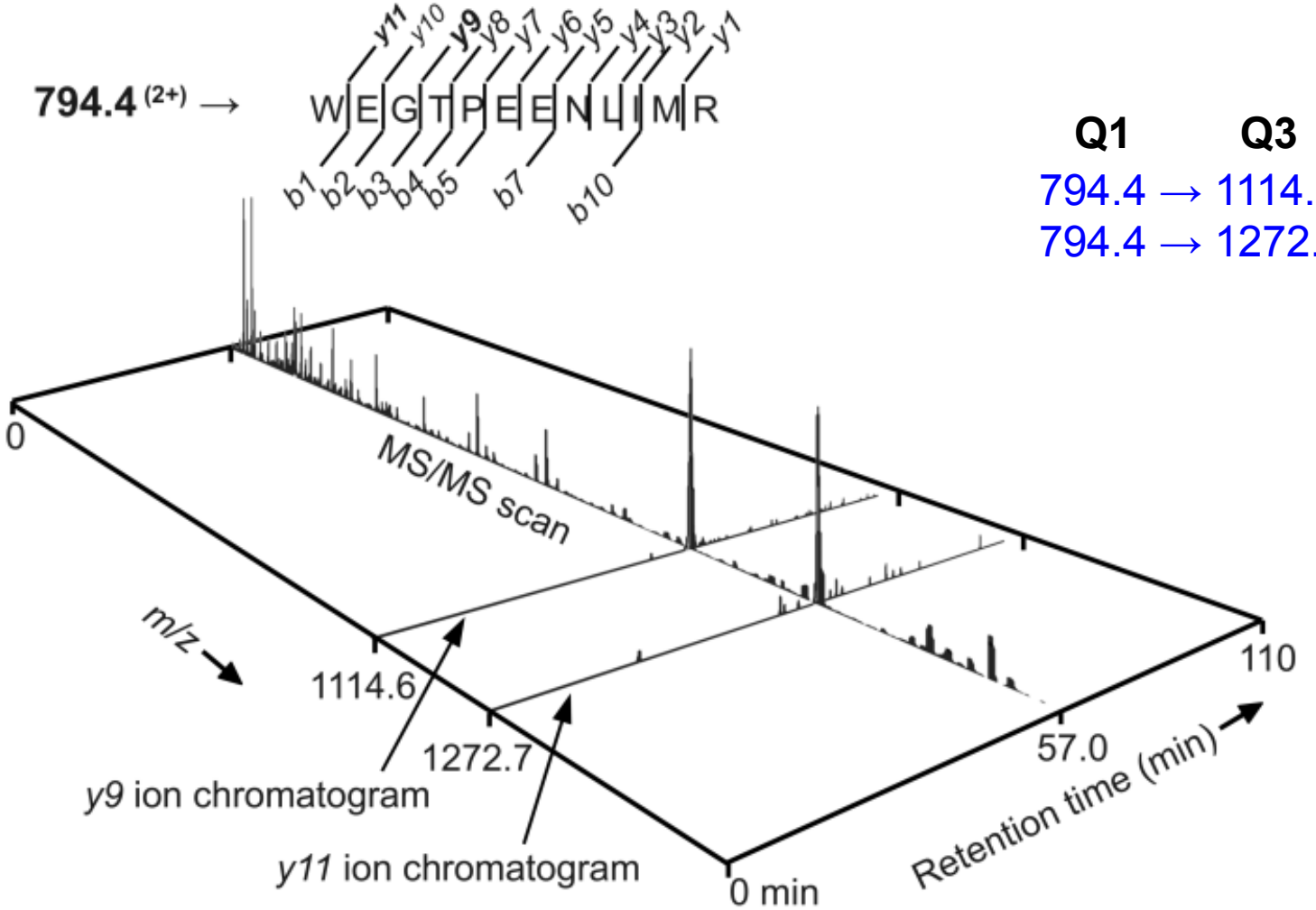


MultiQuant 1.0





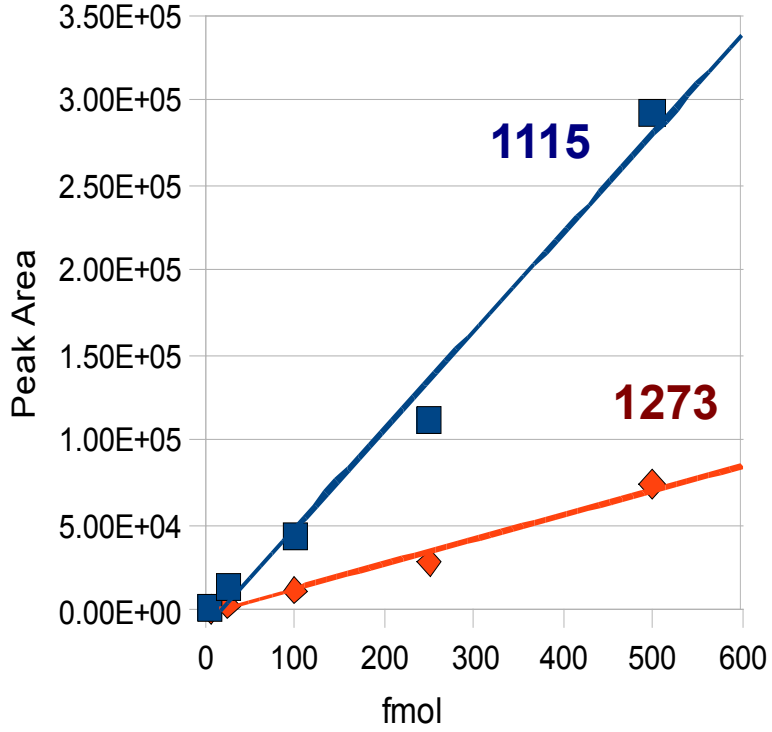
Peptide Quantification *via* Multiple Reaction Monitoring



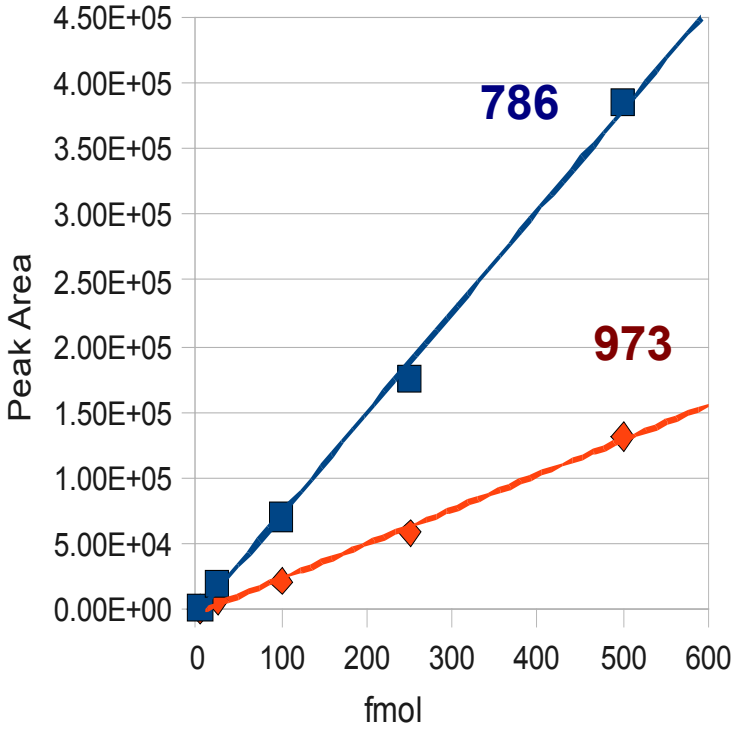
Q1	Q3
794.4 →	1114.6
794.4 →	1272.7

Synthetic Peptide Standards Spiked Into a Background Sediment Digest

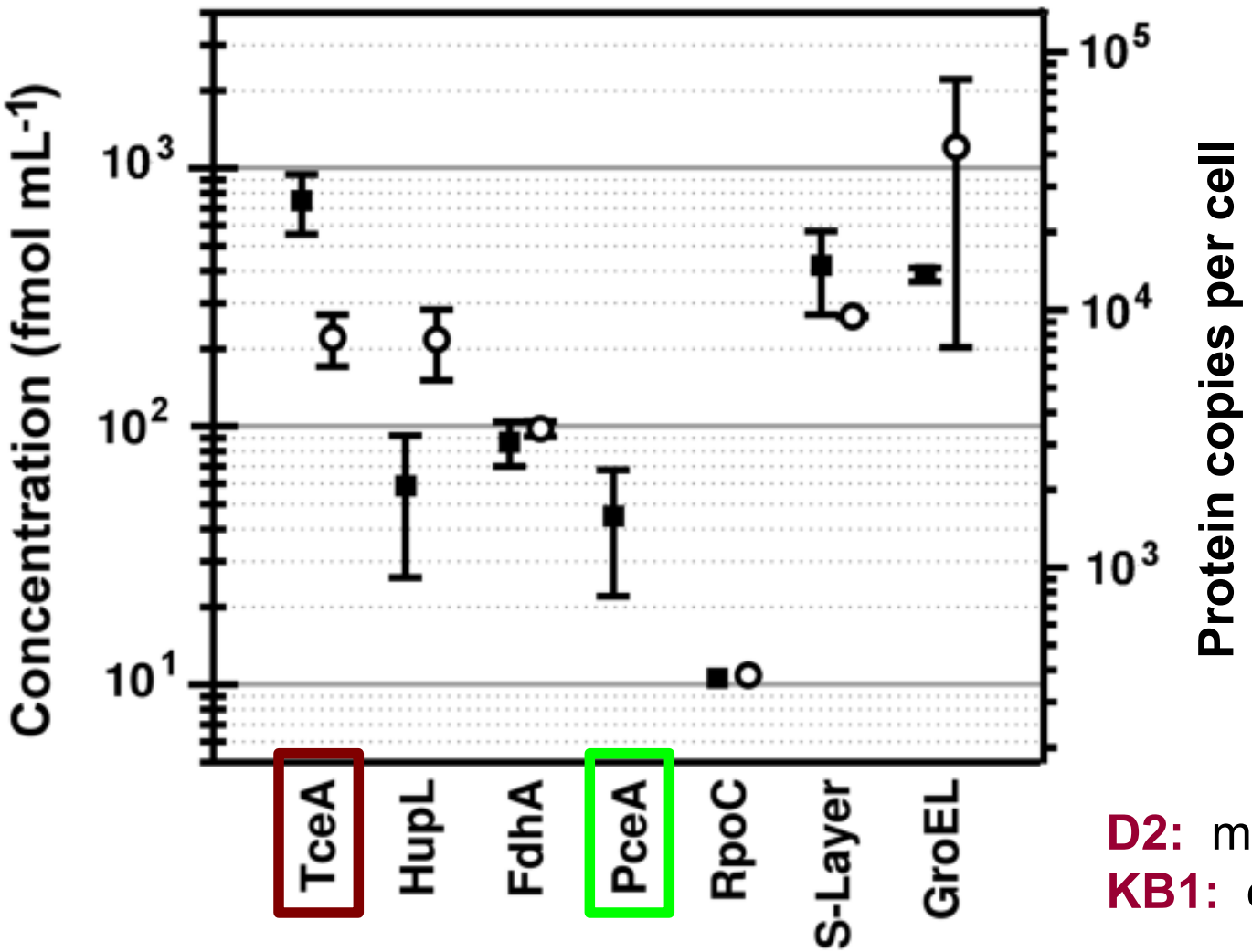
TceA: WEGTPEENLLIMR 794 →



TceA: DEWWASENPIR 702 →



Concentrations in Mixed Cultures



Cultures

■ D2

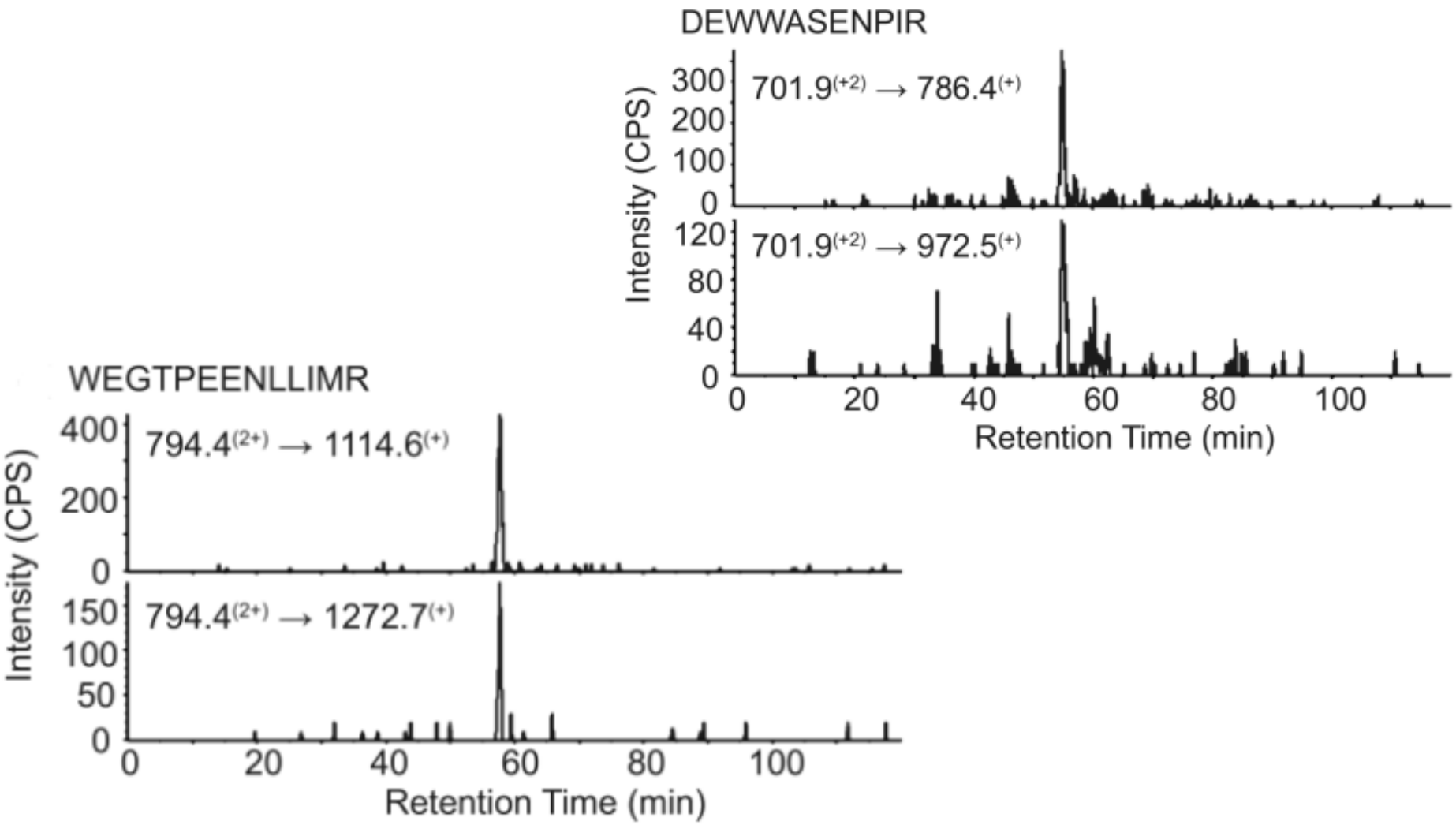
○ KB1

D2: maintained in PCE

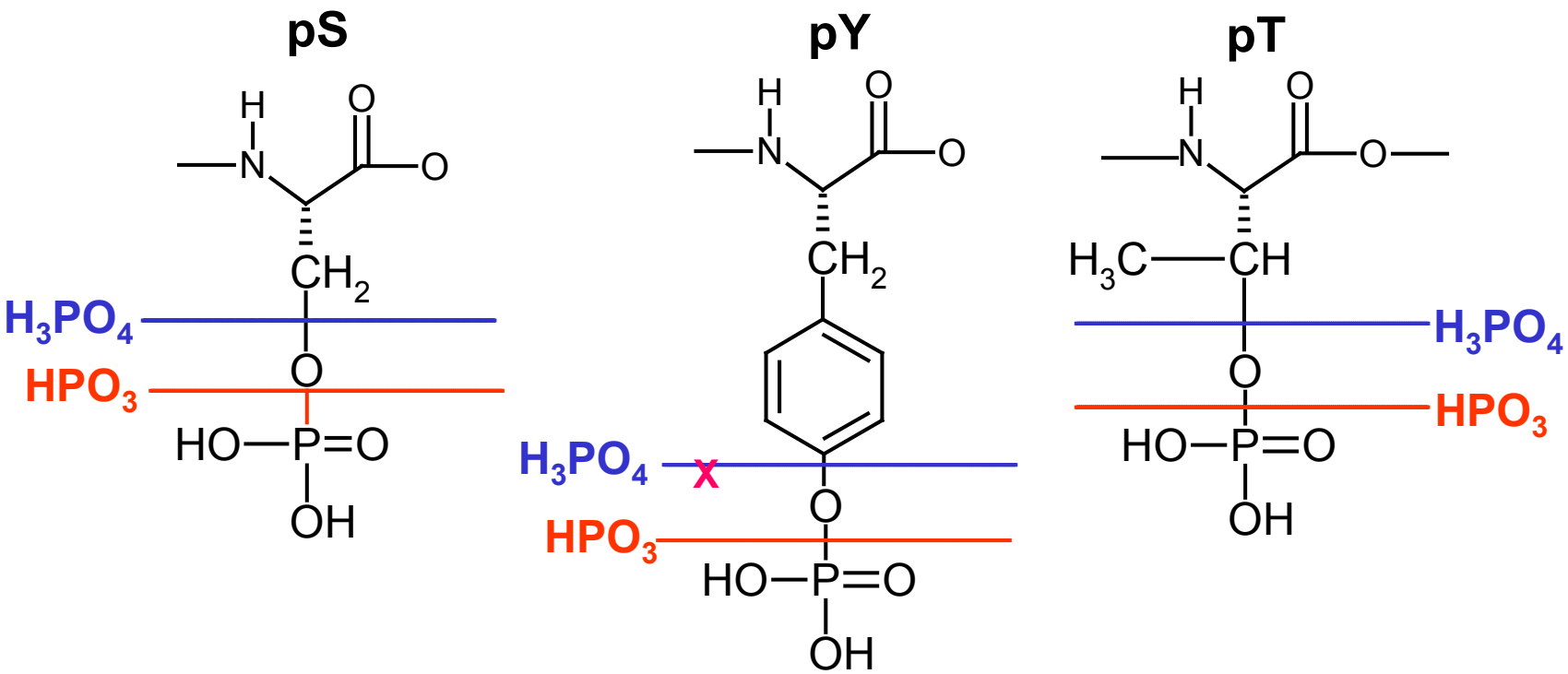
KB1: enriched on TCE

MRM Confirming TceA Enzyme

Previously Thought to be Missing from Commercial Culture KB1™

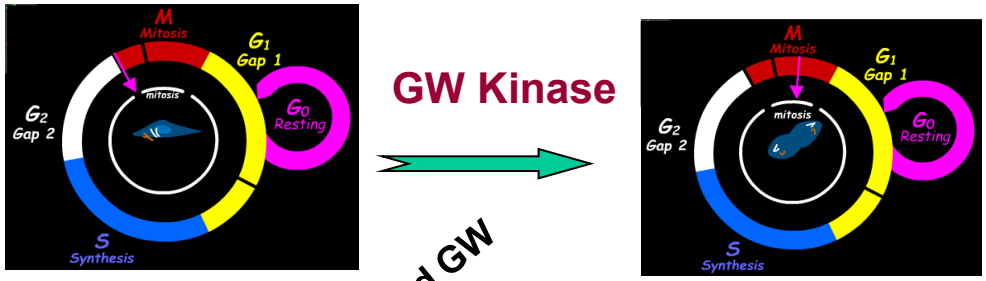


Protein Phosphorylation



	pS	pY	pT
Mass change	+80	+80	+80
Actual change	-18	+80	-18
Marker ions:	m/z -79 (PO_3^-)	m/z -79 m/z +216.04 (pY immonium)	m/z -79

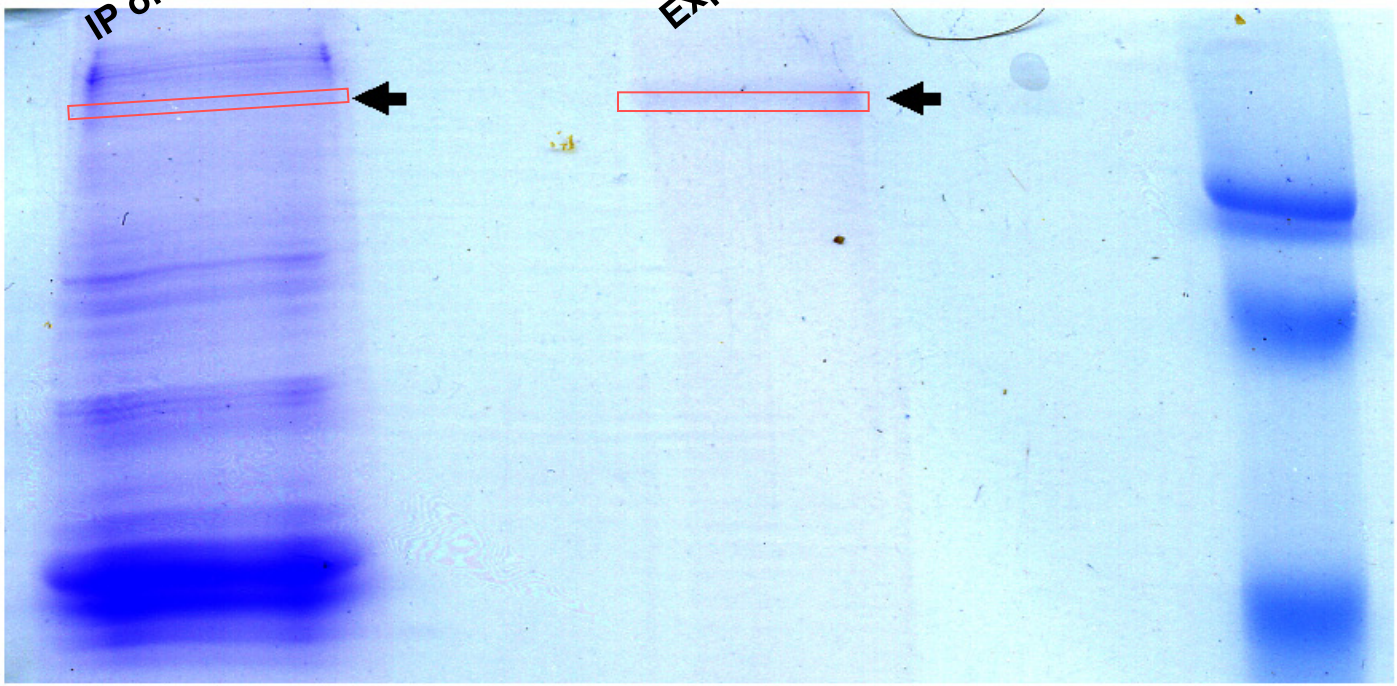
SDS-PAGE of in vivo/in vitro treated Greatwall kinase samples



GW Kinase

IP of Cell Lysate

Expressed GW

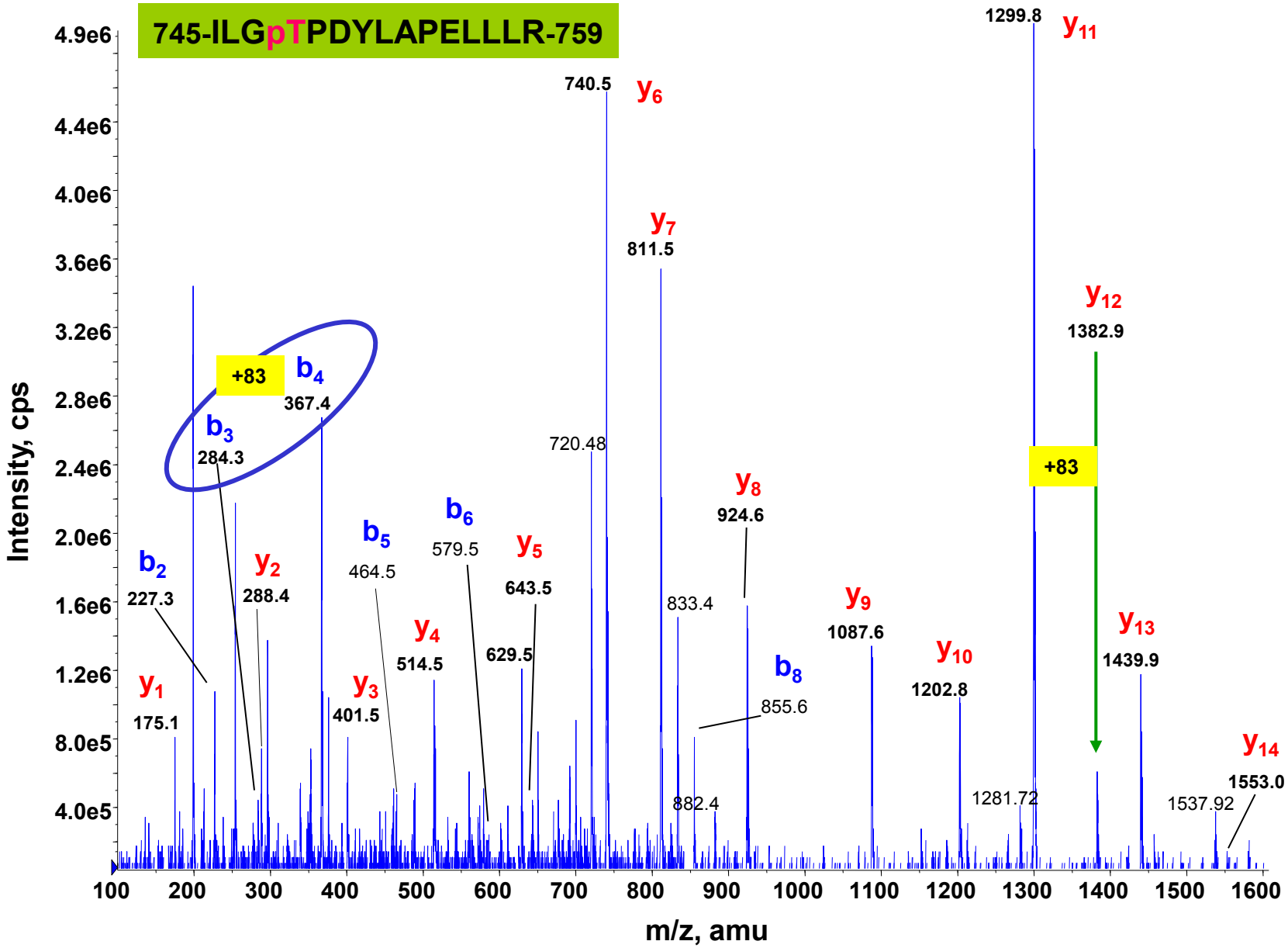


118k

85k

47k

MS/MS Spectrum identifying a Phosphopeptide ion at *m/z* 882.4^{2+}

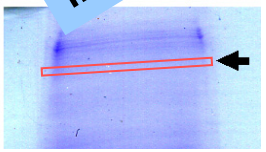


LC/MS/MS of Greatwall Kinase from Cell Lysate

24 Phospho-sites Identified

Expressed GW

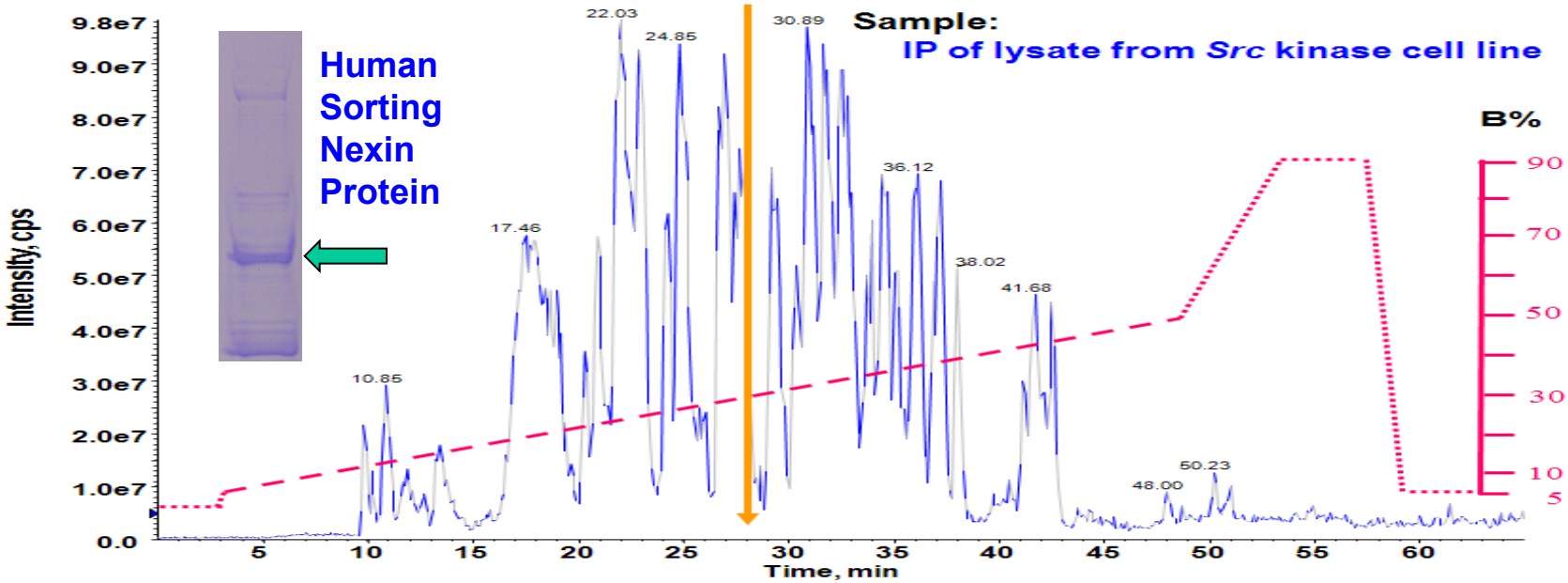
IP of Cell Lysate



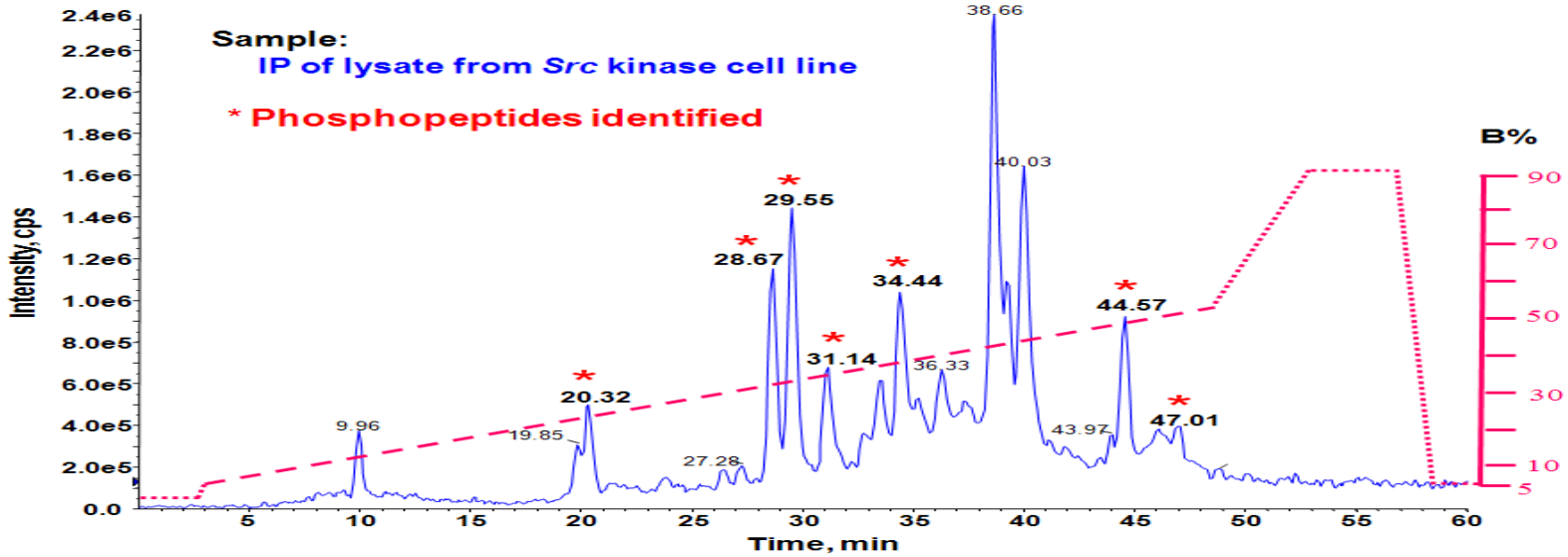
21 proteins identified

Residue	Peptide Sequence	+	Score	Extra (IP)
184 - 200	ELCMMDILpTTPSMAKPK	3	21	√
206 - 226	TPGQVLSLISSLGFNpTPAGGR	3	30	
206 - 226	pTPGQVLSLISSLGFNpTPAGGR	2	57	√
227 - 239	TQGpSLNQQTEGMR	2	69	√
240 - 249	GNASpTPLLMK	2	39	
260 - 277	LMISCPEASLpSSPSIPVK	2	28	√
278-285	CLpTPNLLK	2	26	
288 - 300	TQFATSSTSpSQSR	2	39	√
357 - 368	RLEFAFpSPIVDR	2	48	
374-394	AGFQDETGELSDpTPLATLNAK	2	59	
374-394	AGFQDETGELpSDTPLATLNAK	2	68	√
374-394	AGFQDEpTGELpSDTPLATLNAK	2	49	√
374-394	AGFQDEpTGELSDpTPLATLNAK	2	57	√
374-394	AGFQDEpTGELpSDpTPLATLNAK	2	27	√
424-449	FTSSPDpSPPWLANGSVAPIQFNDEEK	3	32	
458 - 475	NYDLVEKpSPEQELLQDKK	2	46	
532 - 544	TLEKLDpSGNPVAK	2	47	√
637 - 665	YEPNTSILPDSLQNVLApSPAPASAMTNPR	3	53	*
637 - 665	YEPNTpSILPDSLQNVLApSPAPASAMTNPR	3	44	√
674 - 689	SYNpSPINVSNVSEPSK	2	88	*
719 - 728	VQGLIEpTPYR	2	43	
719 - 731	VQGLIEpTPYRpTPK	2	36	
745 - 759	ILGpTPDYLAPELLLR	2	81	
875 - 887	NNAQHLKVpSGFpSL	2	59	√

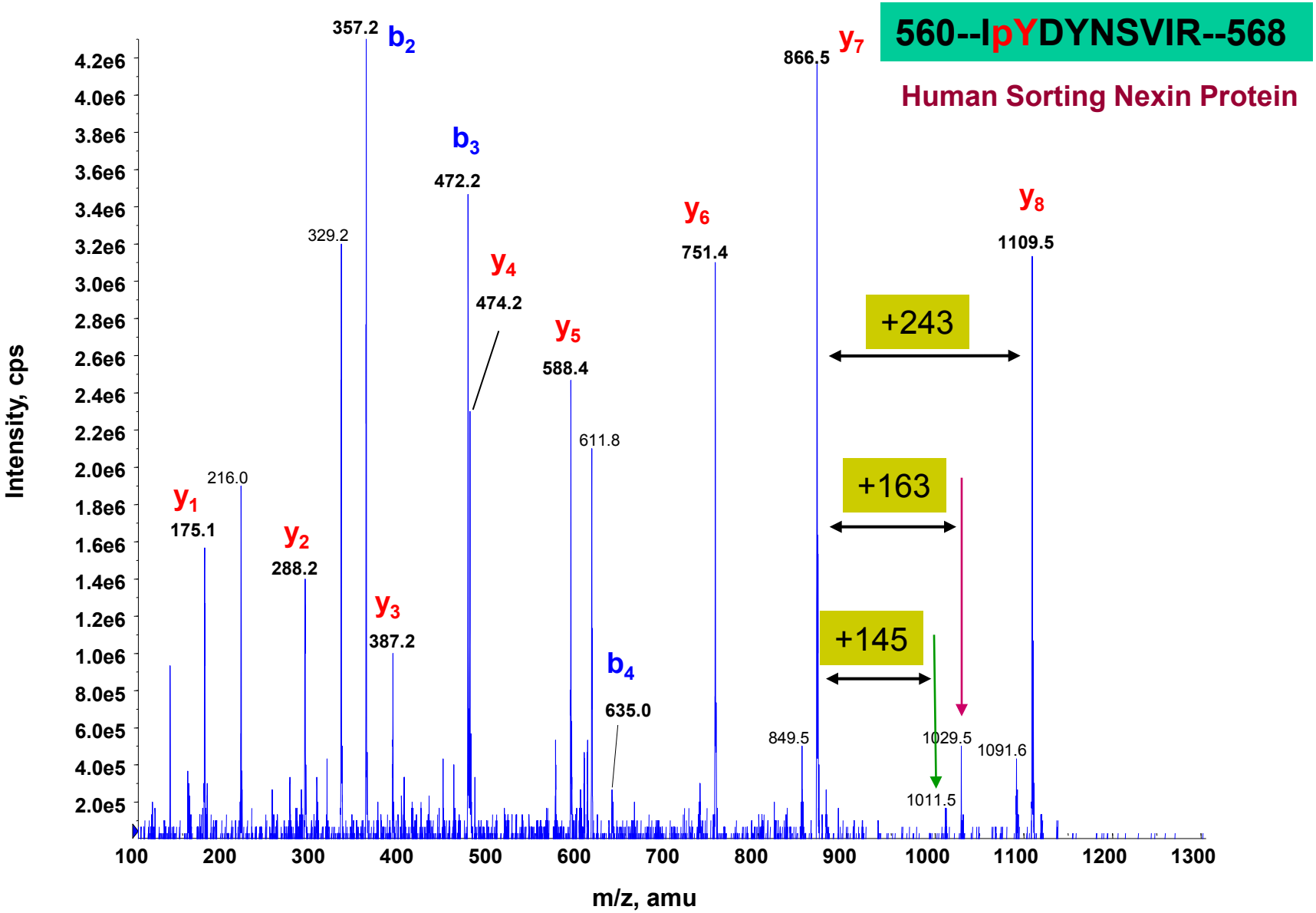
Base Peak Chromatogram of NanoLC-MS/MS in EMS-IDA Mode



Base Peak Chromatogram by Targeted pY Detection in PI Scan (+216.04)



MS/MS Spectrum identifying a Phosphopeptide ion at *m/z* 611.8^{2+}



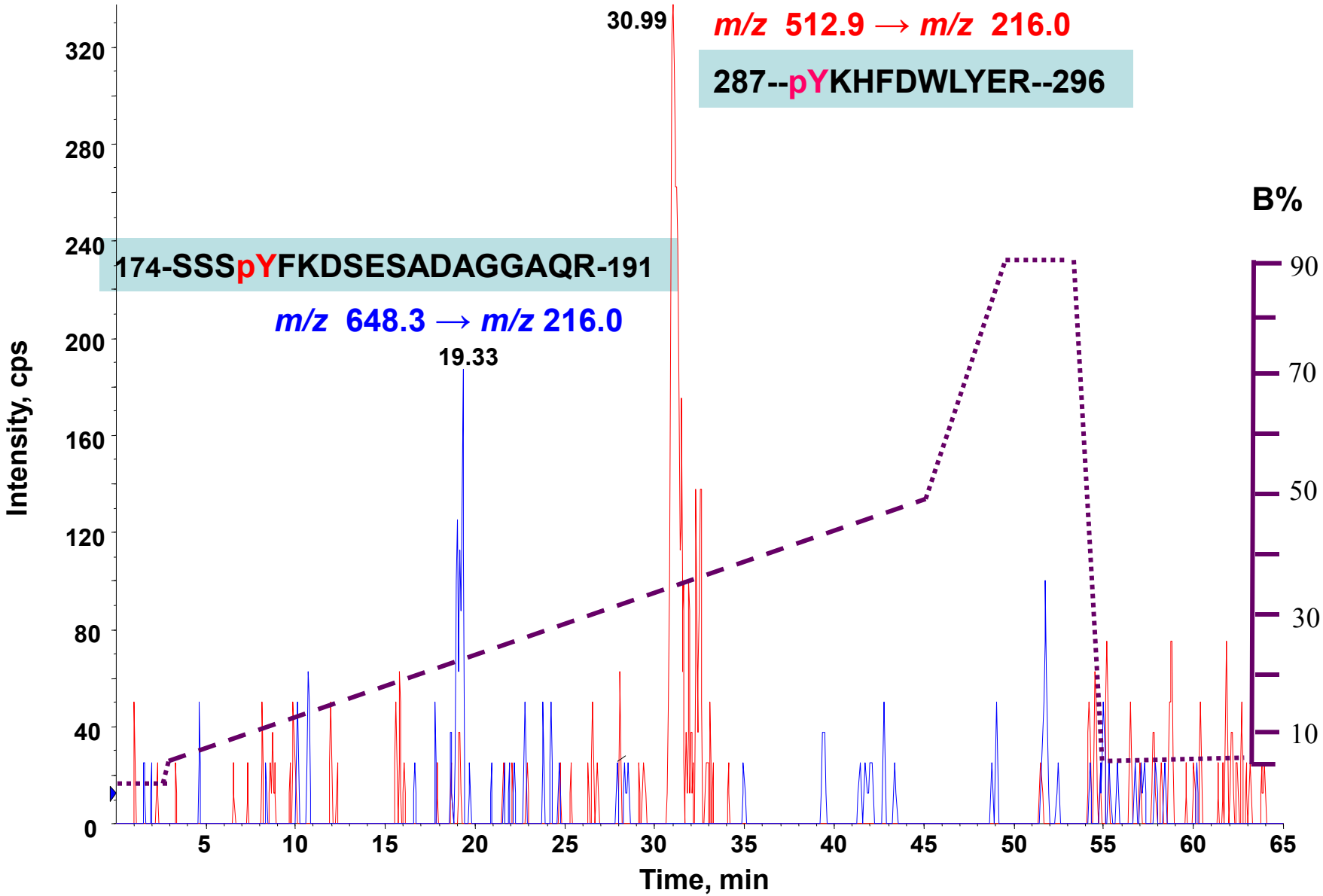
Five pY Sites of Human SNX9 Protein (*Src* Kinase) Identified by nLC-IDA-MS/MS*

Residue	Peptide Sequence	+	*M/Z*	Mascot Score	% phospho-rylation
174-191	SSSpYFKDSESADAGGAQR	3	648.3	73	6.0
230-258	EKIPIIVGDpYGPMWVYPTSTFDCVV ADPR	3	1136.15	48	25.2
232 - 258	IPIIVGDpYGPMWVYPTSTFDCVVAD PR	3	1050.1	95	23.2
268 - 281	SpYIEYQLTPTNTNR	3	890.4	46	2.7
287-296	YKHFDWLpYER	3	521.9	44	10.0
560 - 568	IpYDYNSVIR	2	611.8	58	4.7

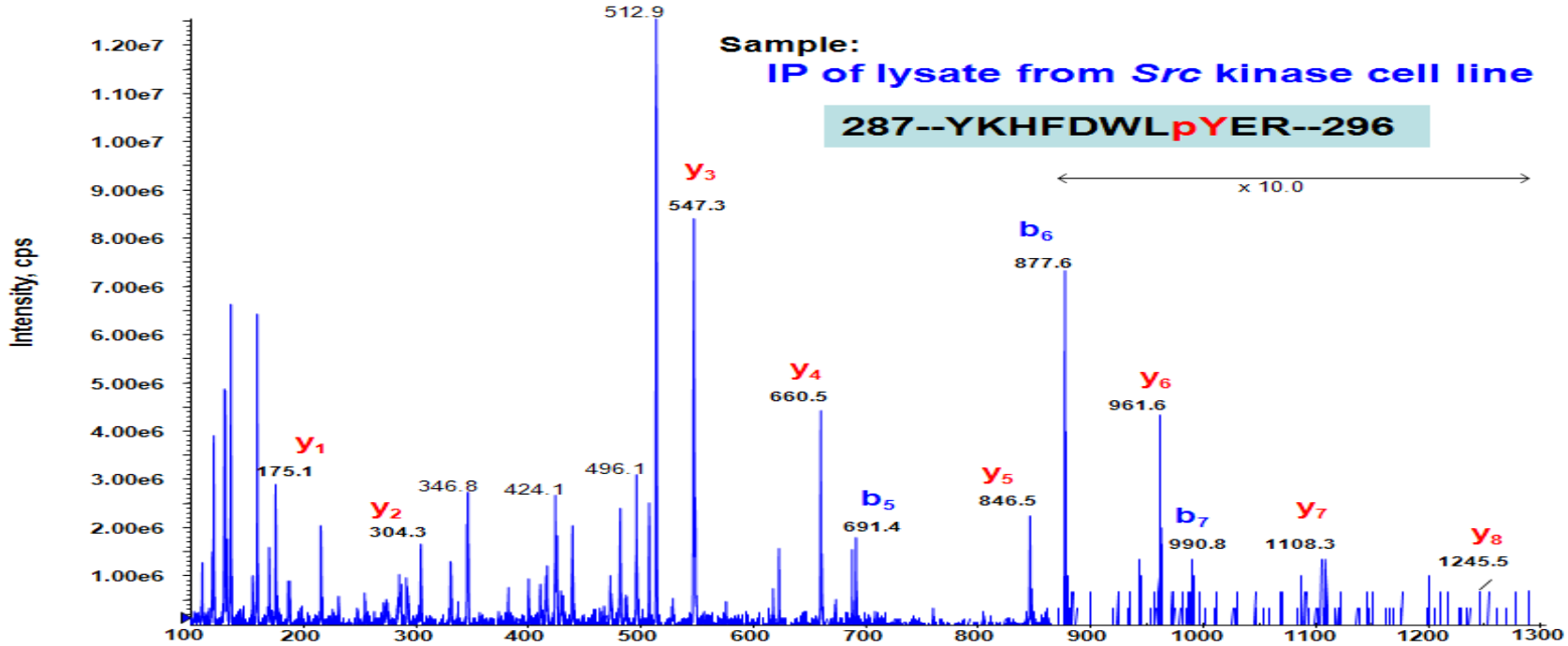
Two pY Sites (*ACK2* Kinase) Identified by nLC-IDA-MS/MS

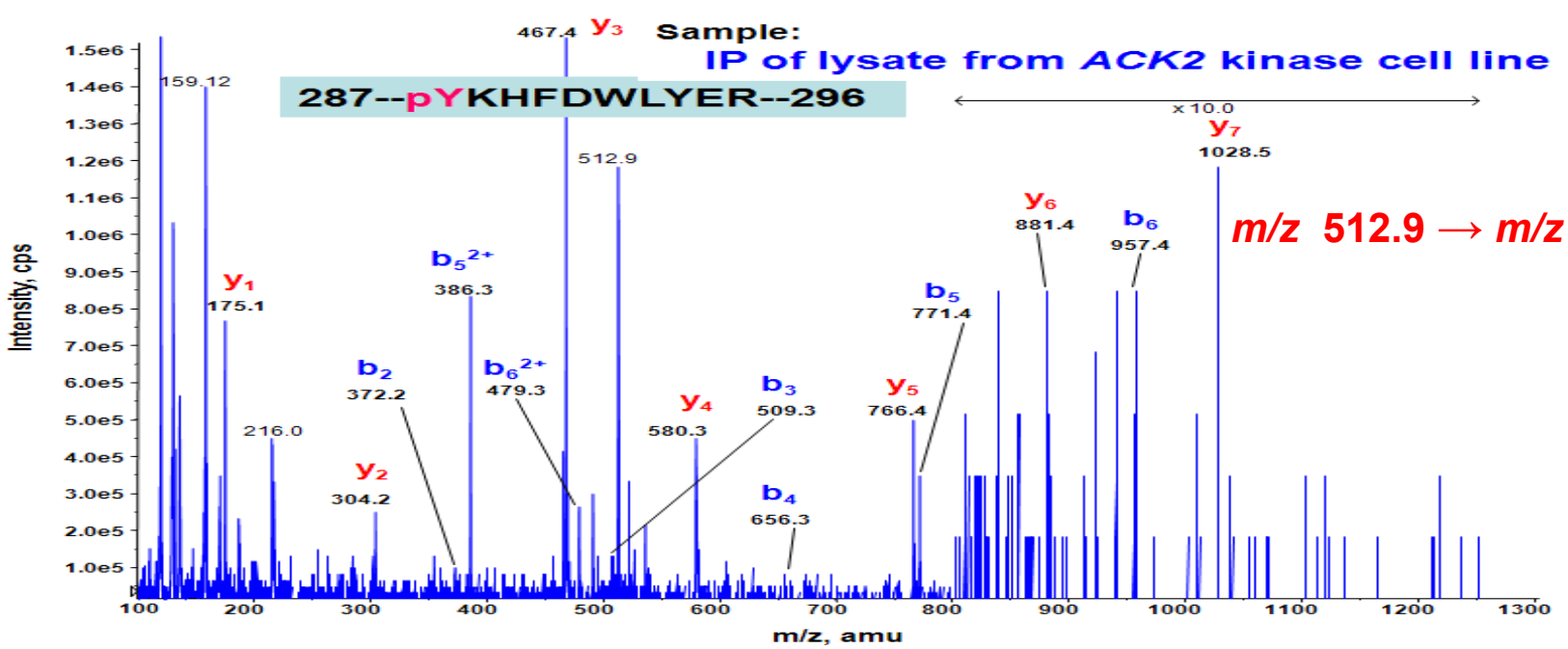
Residue	Peptide Sequence	+	*M/Z*	Mascot Score	MS-IDA	MRM-IDA (X/216)	Prec-IDA (+216)
174-191	SSSpYFKDSESADAGGAQR	3	648.3	56	No	Yes	Yes
287-296	pYKHFDWLYER	3	521.9	26	No	Yes	Yes

XIC of MRM for Two-pair Transition Ions of Identified pY-peptides



MS/MS Spectrum Identifying a Phosphopeptide Ion at *m/z* 512.9^{3+}

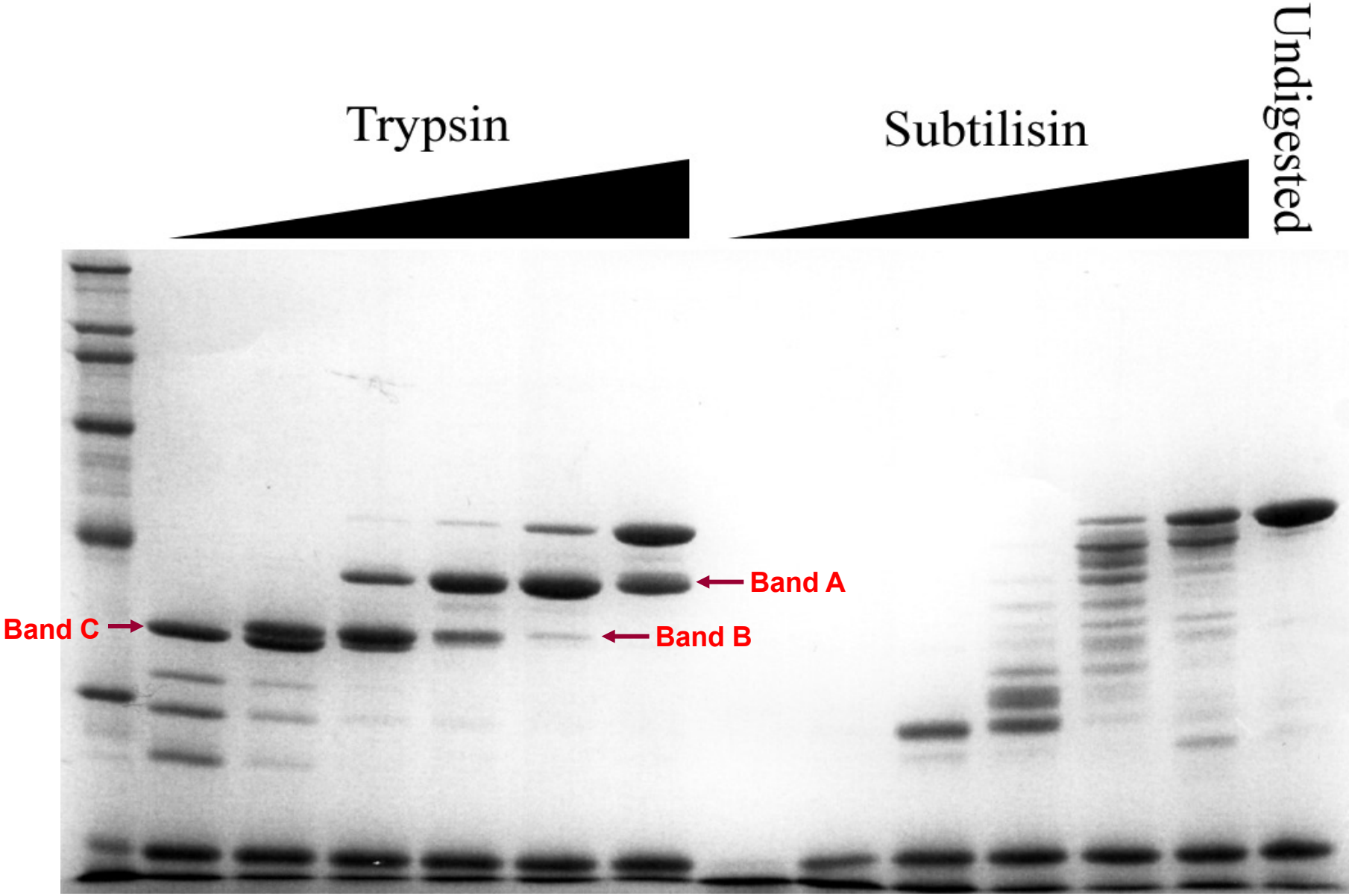




m/z 512.9 → *m/z* 216.04

Peptide mapping for N/C-terminal determination

Limited Proteolysis of Yeast Chitin Biosynthesis Protein



Chymotrypsin digestion

Band A

GAMDPEFSSVDVLLTVGKLDASLALLTTQDHHVIEFPTVLLPENVKAGSIIKMQVSQNLEE
EKKQRNHFKSIQAKILEKYGTHKPESPVLKIVNVTQTSCVLAWDPLKLGSAKLKSLILYRK
GIRSMVIPNPFKVTTTKISGLSVDTPYEFQLKLITTSGTLWSEKVILRTHKMTDMSGITVC
LGPLDPLKEISDLQISQCLSHIGARPLQRHVAIDTTHFVCNDLDNEESNEELIRAKHNNIP
IVRPEWVRACEVEKRIVGVRGFYLDADQSILKSYTFPPVNEEELSYSKENEPVAEVADENK
MPEDTTDVEQVASHNDNEGNPSEAKEQGEKSGHEAAPVSPAEDPLHASTALENETTIETVN
PSVR

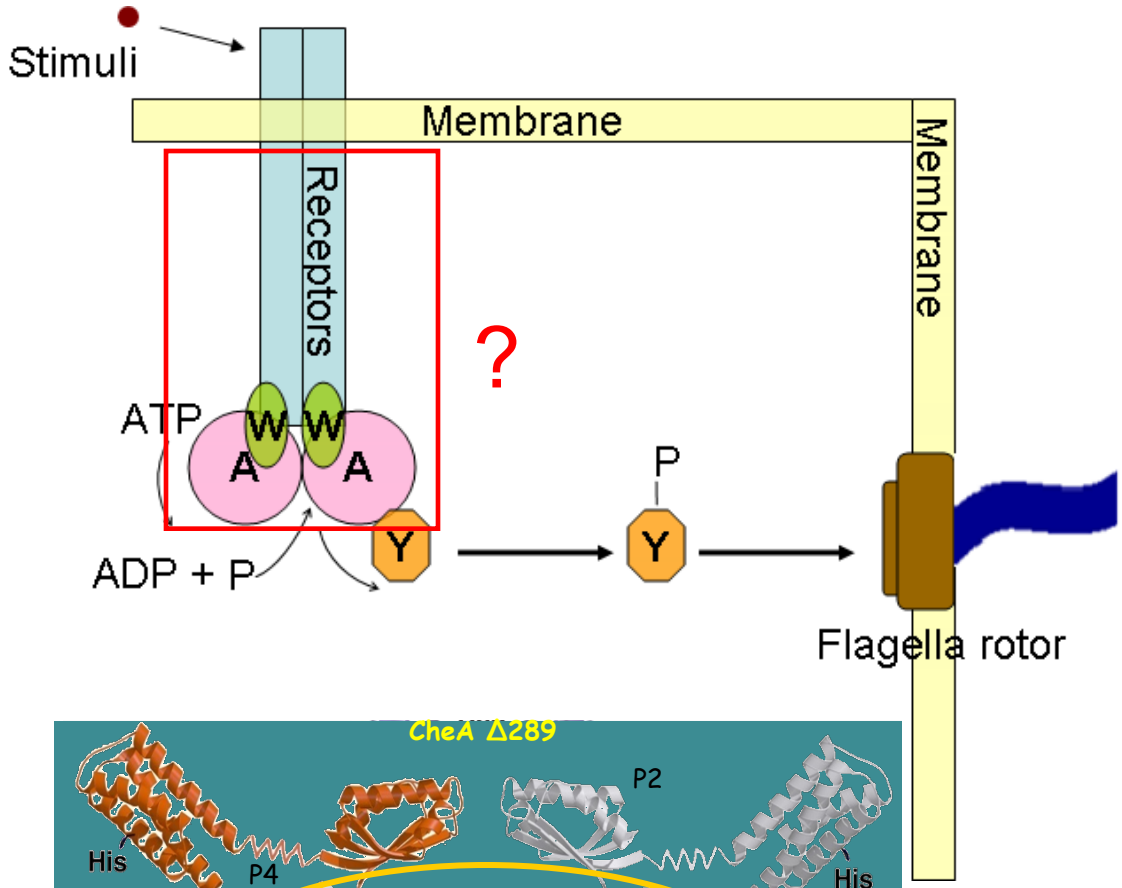
Band B

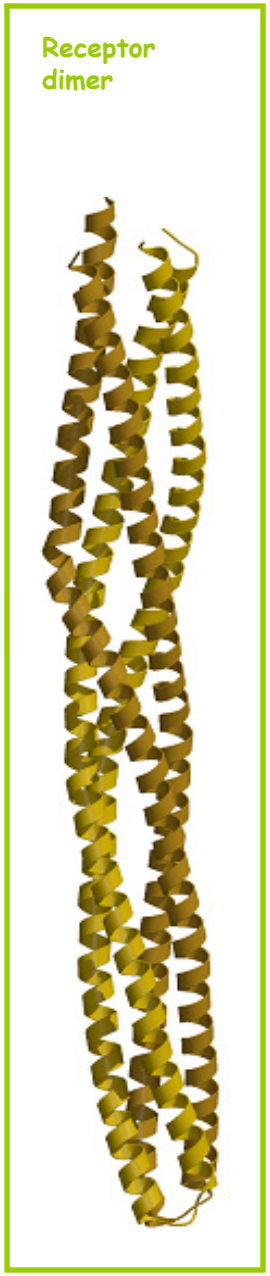
GAMDPEFSSVDVLLTVGKLDASLALLTTQDHHVIEFPTVLLPENVKAGSIIKMQVSQNLEE
EKKQRNHFKSIQAKILEKYGTHKPESPVLKIVNVTQTSCVLAWDPLKLGSAKLKSLILYRK
GIRSMVIPNPFKVTTTKISGLSVDTPYEFQLKLITTSGTLWSEKVILRTHKMTDMSGITVC
LGPLDPLKEISDLQISQCLSHIGARPLQRHVAIDTTHFVCNDLDNEESNEELIRAKHNNIP
IVRPEWVRACEVEKRIVGVRGFYLDADQSILKSYTFPPVNEEELSYSKENEPVAEVADENK
MPEDTTDVEQVASHNDNEGNPSEAKEQGEKSGHEAAPVSPAEDPLHASTALENETTIETVN
PSVR

Band C

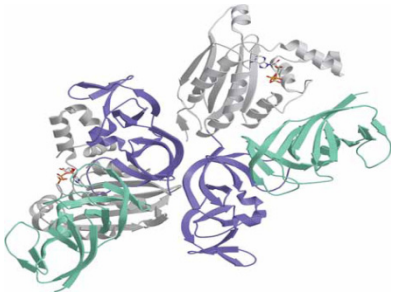
GAMDPEFSSVDVLLTVGKLDASLALLTTQDHHVIEFPTVLLPENVKAGSIIKMQVSQNLEE
EKKQRNHFKSIQAKILEKYGTHKPESPVLKIVNVTQTSCVLAWDPLKLGSAKLKSLILYRK
GIRSMVIPNPFKVTTTKISGLSVDTPYEFQLKLITTSGTLWSEKVILRTHKMTDMSGITVC
LGPLDPLKEISDLQISQCLSHIGARPLQRHVAIDTTHFVCNDLDNEESNEELIRAKHNNIP
IVRPEWVRACEVEKRIVGVRGFYLDADQSILKSYTFPPVNEEELSYSKENEPVAEVADENK
MPEDTTDVEQVASHNDNEGNPSEAKEQGEKSGHEAAPVSPAEDPLHASTALENETTIETVN
PSVR

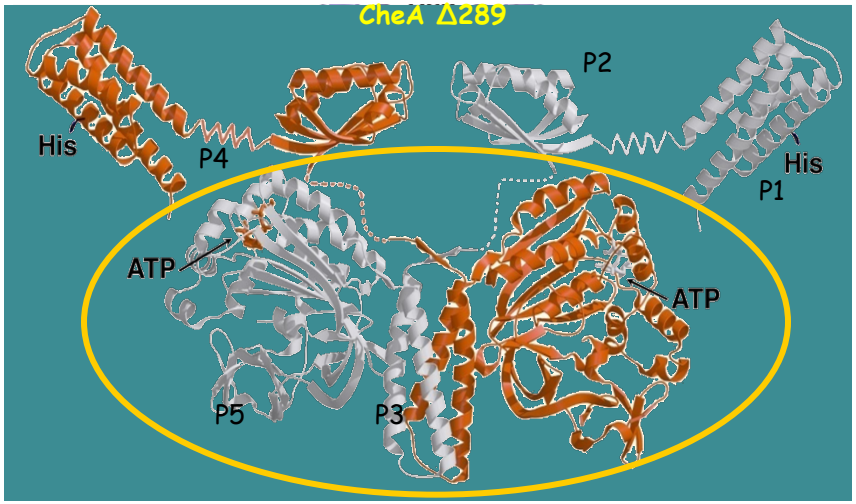
Components of ternary complex





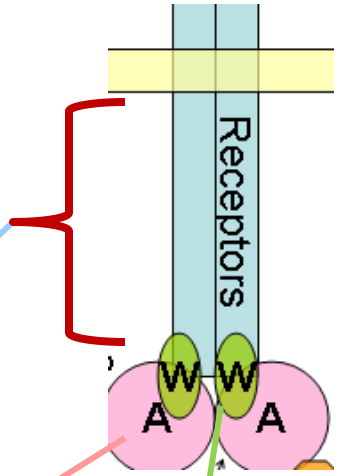
A:W complex





Protein-Protein Complex

A ternary mixture (transmembrane complexes that mediate bacterial chemotaxis) consists of:



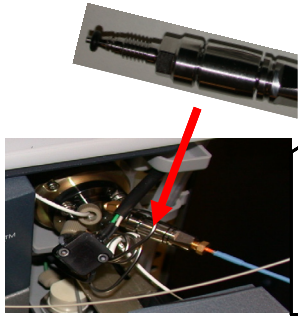
C2 (a cytoplasmic fragment of a chemoreceptor)

D289 (a truncated version of CheA, the chemotaxis signaling kinase)

- CheW (an adaptor protein that mates the receptor to CheA)

A ternary mixture: C2-D289-CheW

LC MS TOF



100 µg injected on desalting column

A: 0.1% Formic Acid; B: ACN-0.1%FA

Gradient: 0 → 0.5 → 8 → 8.5 min,
10% → 20% → 45% → 80% B

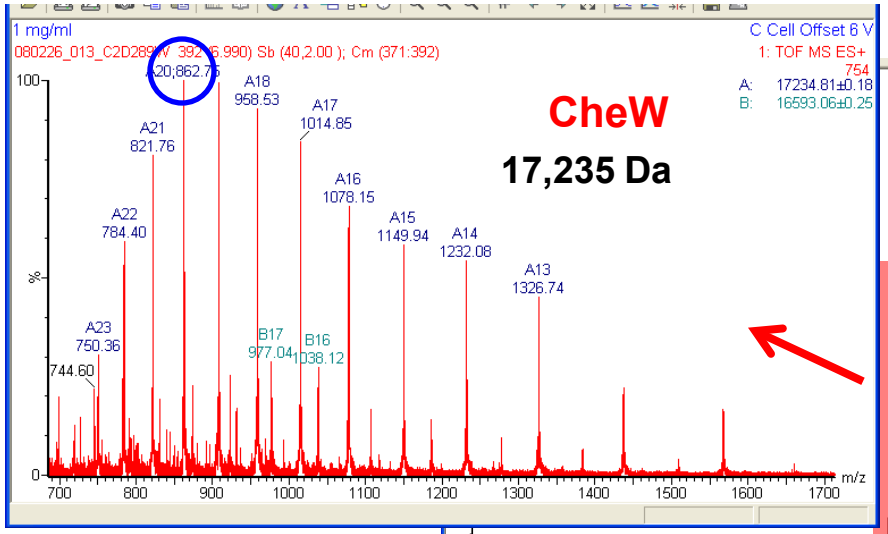
Cone Voltage = 40 V

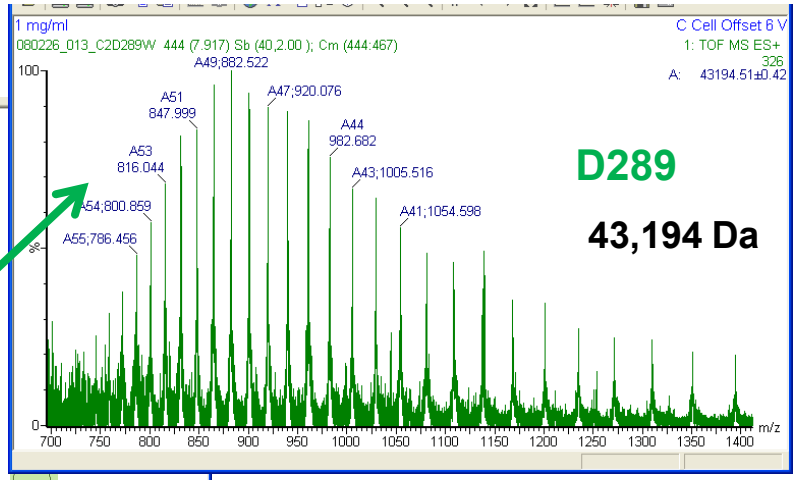
Collision Energy = 6 V

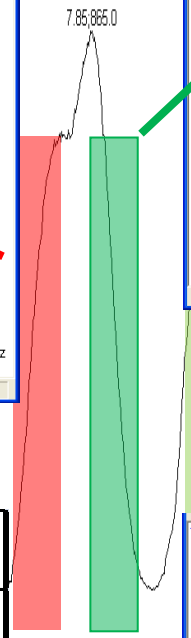
Source Temperature = 105 ° C

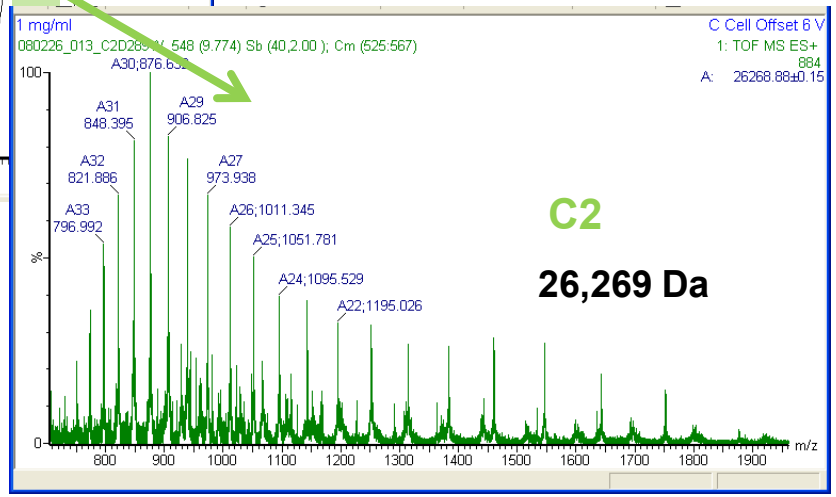
Desolvation Temperature = 150 ° C

LC-MS analysis (under denaturing conditions)



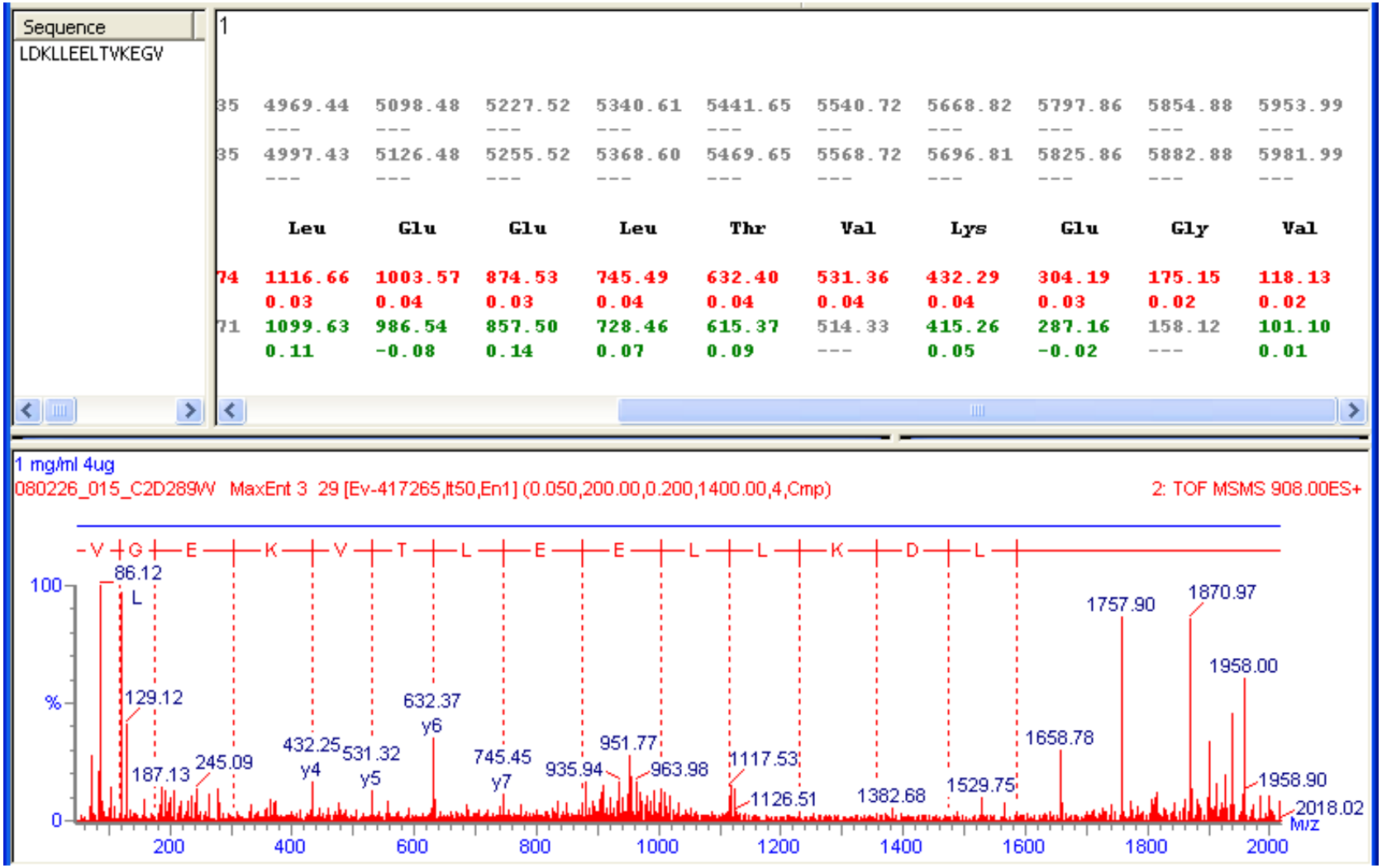






Protein	expected	measured	difference
Chew -W	16954	16593.3	-360.7
Chew -W		**17234.8**	280.8
C	26419	24518.0	-1901.0
C		**26268.9**	-150.1
D289	42785	**43194.3**	409.3

MS/MS Spectrum of m/z 908^{19+} ion in CheW Sample



MADALKE**FEV LSFEIDEQAL AFD**VDNIEMV IEKSDITPVP KSRHFVEGVI NLRGRIIPVV NLAKILGISF DEQKMKSIIV ARTKDVEVGF LVDRVLGVLR ITENQLDLTN VSDKFGKKSK GLVKTDGRLI IY**LDIDKIIE EITVKEGV** 17,235 Da

MS/MS Spectrum of m/z 908^{19+} ion in CheW Sample

Sequence
FEVLSFELDEQALAFD

bserved MW: 5998.9922 Precursor ion charge state: 1
/z tolerance: 0.30 Intensity threshold: 47 (0.750%)

1354.69	1501.76	1630.80	1729.87	1842.95	1929.98	2077.05	2206.09	2319.18	2434.20	2563.25	2691.3
---	---	0.02	0.09	-0.01	---	0.01	-0.05	---	0.04	---	0.00
1382.68	1529.75	1658.79	1757.86	1870.95	1957.98	2105.05	2234.09	2347.17	2462.20	2591.24	2719.3
0.00	-0.00	0.01	-0.04	-0.02	-0.02	-0.04	-0.05	0.02	0.05	-0.02	-0.04
1381.67	Phe	Glu	Val	Leu	Ser	Phe	Glu	Leu	Asp	Glu	Gln

Should be 1101.6 from sequence (281.1 Da extra)

1 mg/ml 4ug
080226_015_C2D289W MaxEnt 3 29 [Ev-417265,It50,En1] (0.050,200.00,0.200,1400.00,4,Cmp) 2: TOF MSMS 908.00ES+

(1381.67) F E V L S F E L D E Q A L A F D

86.12 L; 129.12; 187.13; 432.25; 632.37; 935.94; 951.77; 956.36; 1117.53; 1126.51; 1529.75 b1; 1658.78 b2; 1757.90 b3; 1870.97 b4; 1958.00 b5; 2105.09 b6; 2234.14 b7; 2347.15 b8; 2462.15 b9; 2790.30 b12; 2903.41 b13; 2974.39 b14; 3349.95; 3564.41; 3582.89

100 % 0

200 400 600 800 1000 1200 1400 1600 1800 2000 2200 2400 2600 2800 3000 3200 3400 M/z

thrombin

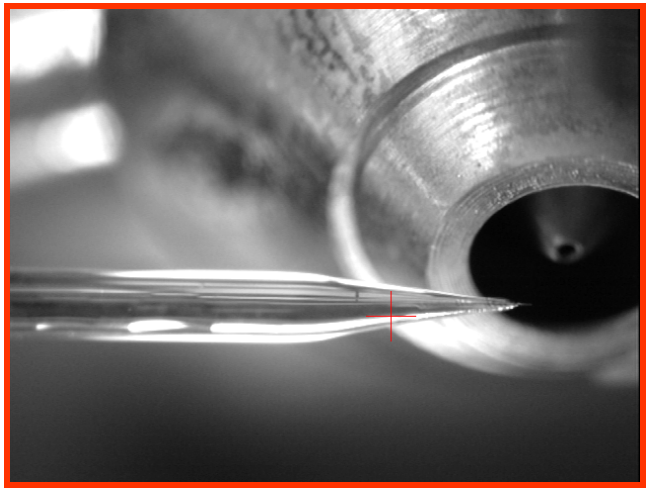
MGSSHHHHHHSSGLVPRGSH

MADALKE**FEV LSFEIDEQAL AFD**VDNIEMV IEKSDITPVP KSRHFVEGVI NLRGRIIPVV NLAKILGISF DEQKMKSIIV
ARTKDVEVGF LVDRVLGVLR ITENQLDLTN VSDKFGKKSK GLVKTDGRLI IY**LDIDKIIE EITVKEGV** 16,954 Da

GSH = 281 Da, + 16,954 = 17,235 Da

A ternary mixture: C2-D289-CheW

Nanospray MS



C2D289W 1 μg/μL 100mM NH_4Ac

Cone Voltage = 150 V

Trap Energy = 20 V

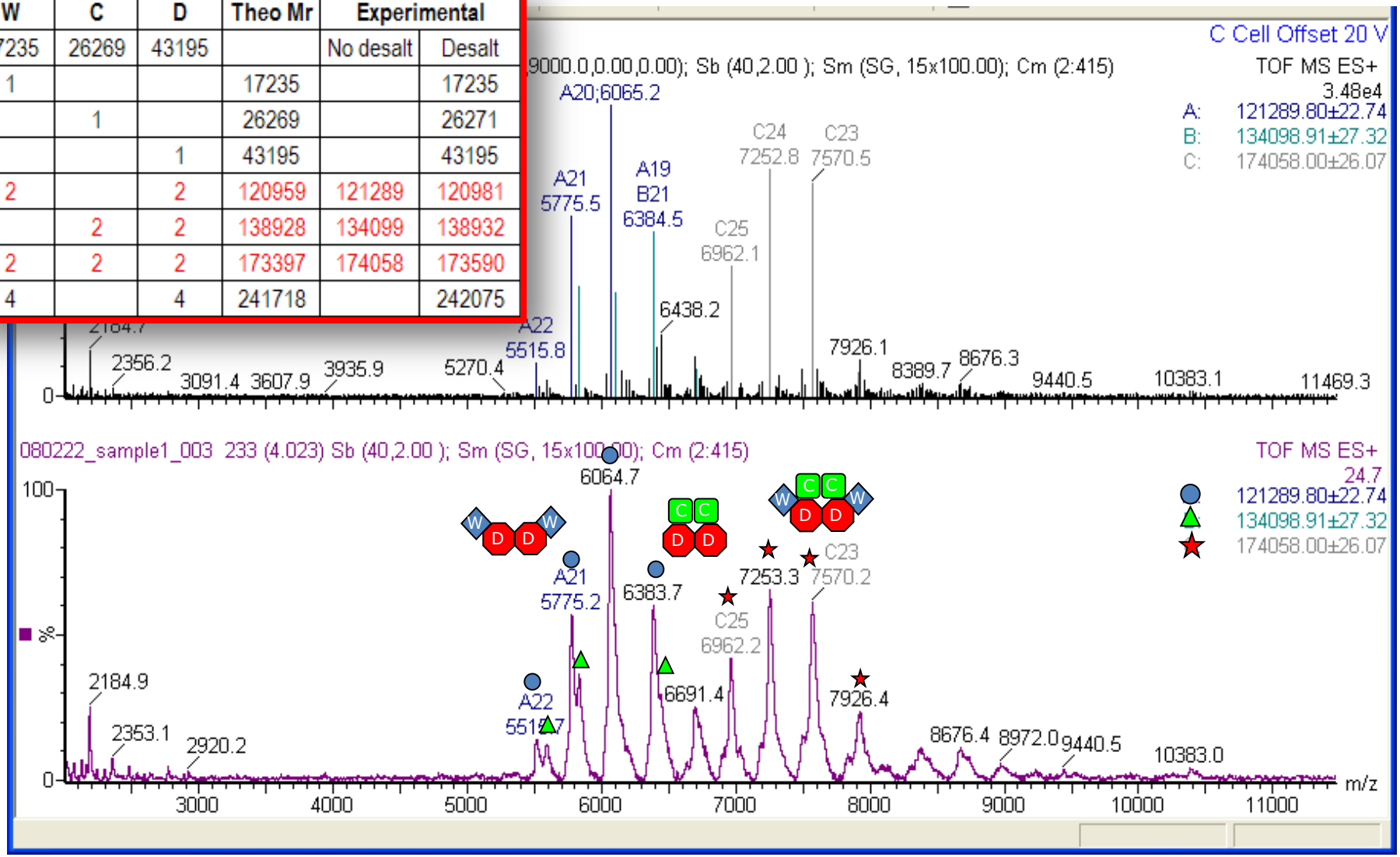
Transfer Energy = 20 V

Source Temperature = 60 ° C

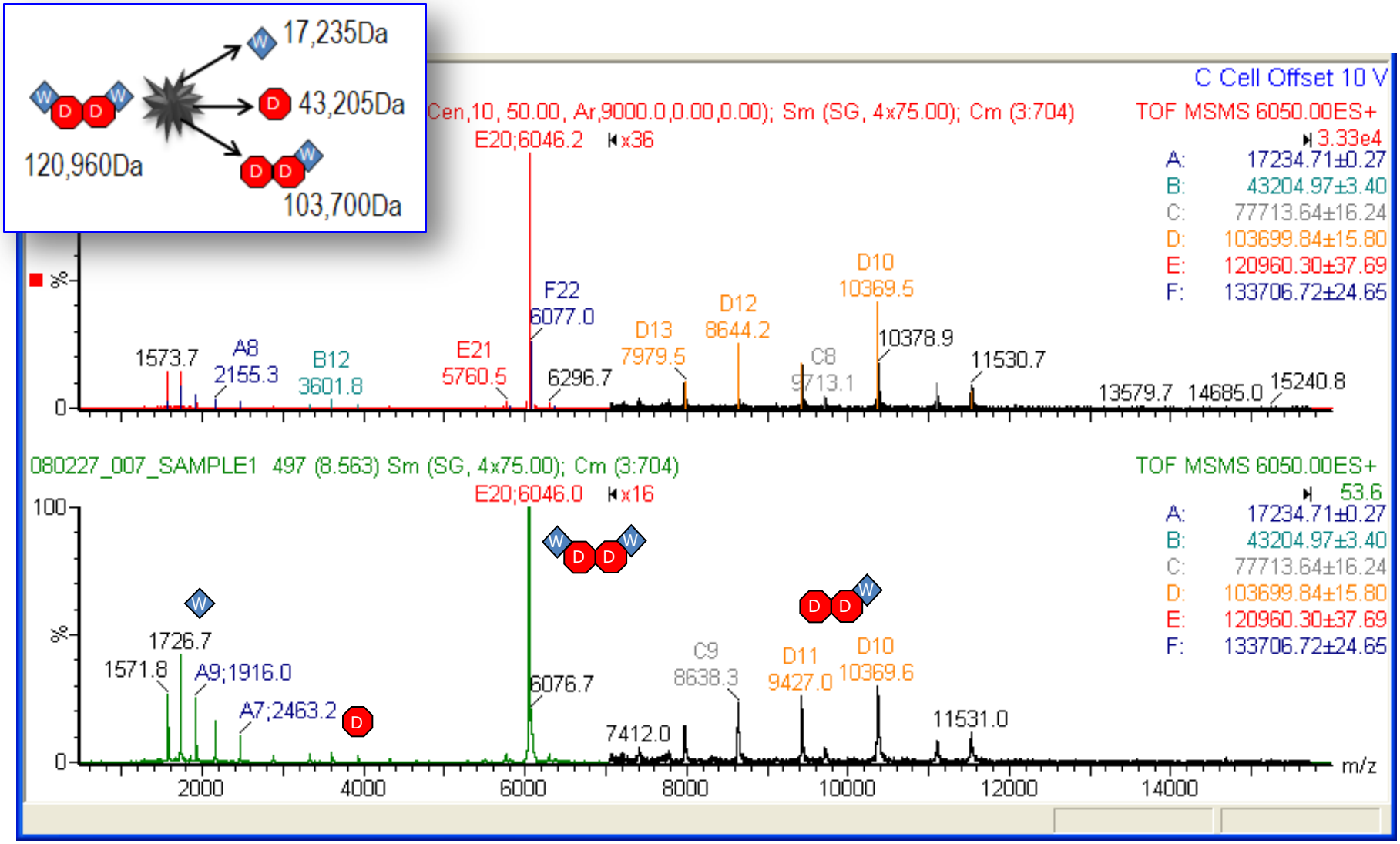
Infusion analysis (under native conditions)

MS Spectra for C2-D289-CheW Complex

W	C	D	Theo Mr	Experimental	
17235	26269	43195		No desalt	Desalt
1			17235		17235
	1		26269		26271
		1	43195		43195
2		2	120959	121289	120981
	2	2	138928	134099	138932
2	2	2	173397	174058	173590
4		4	241718		242075

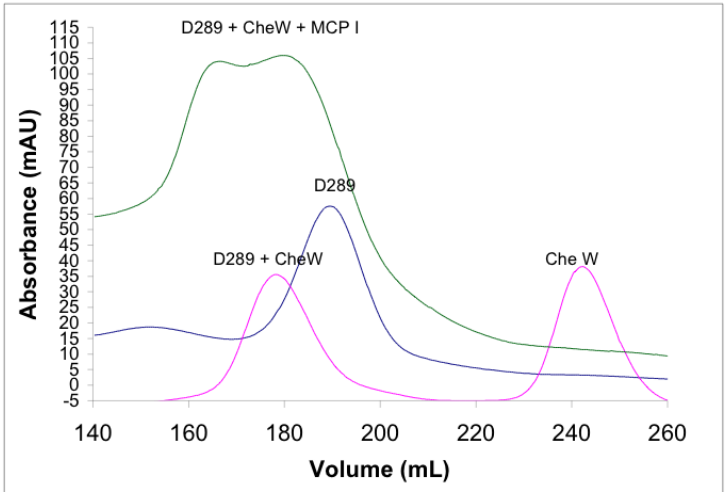


MS/MS Spectra of m/z 6050^{20+} (2xCheW-2XD289)



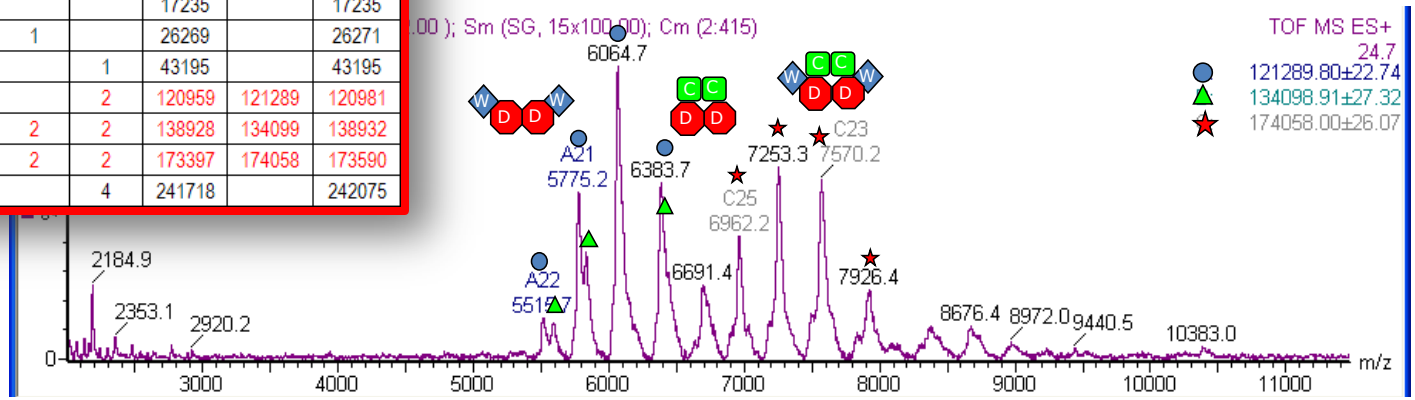
T. maritima MCP fragments bind CheA:CheW

Size-exclusion chromatography



Electrospray mass spectrometry

W	C	D	Theo Mr	Experimental	
17235	26269	43195		No desalt	Desalt
1			17235		17235
	1		26269		26271
		1	43195		43195
2		2	120959	121289	120981
	2	2	138928	134099	138932
2	2	2	173397	174058	173590
4		4	241718		242075



Equilibrium ultracentrifugation

Sample	Model	Fixed Molecular Weights (kD)	Predicted Molecular Weight	Variance	Number of Runs
CheW	One component	17.86	16.95	1.83E-05	37.71%
CheA (P3-P4-P5)	One component	78.87	42.78 (85.56)	3.99E-05	44.08%
Receptor fragment	One component	42.77	24.11(48.22)	2.26E-05	35.15%
CheA (P3-P4-P5) + CheW	Two component	78.34, 117.5	119.46	1.53E-05	33.61%
Receptor fragment CheA(P3-P4-P5) + CheW	Three component	116.9, 146.2, 175.4	119.46, 133.78, 167.68	1.76E-05	34.99%

D289 + CheW

CheA+C2

D289+CheW+C2

Acknowledgments

- **Prof. Steve Zinder (Microbiology):** GeLC-MS/MS
- **Prof. Ruth Richardson (CEE):** iTRAQ analysis
- **Prof. Michael Goldberg (MBG):** GW phosphorylation
- **Prof. Richard Cerione (Vet):** Targeted analysis, pY of SN
- **Prof. Chris Fromme (Weill Institute):** N/C-terminal
- **Prof. Brian Crane (Chemistry):** Protein complex

Take-home Messages

✓ Ionization is required in MS Analysis and describe how ESI works.

✓ MS becomes an invaluable technique in the proteomics field, why?

✓ Why proteomics relies heavily on genomics information for the targeted species?

Reference books:

1. "Fundamentals of Contemporary Mass Spectrometry" 2007 by Chhabil Dass
2. "Protein Sequencing and Identification Using Tandem Mass Spectrometry" 2000 by Michael Kinter and Nichols E. Sherman