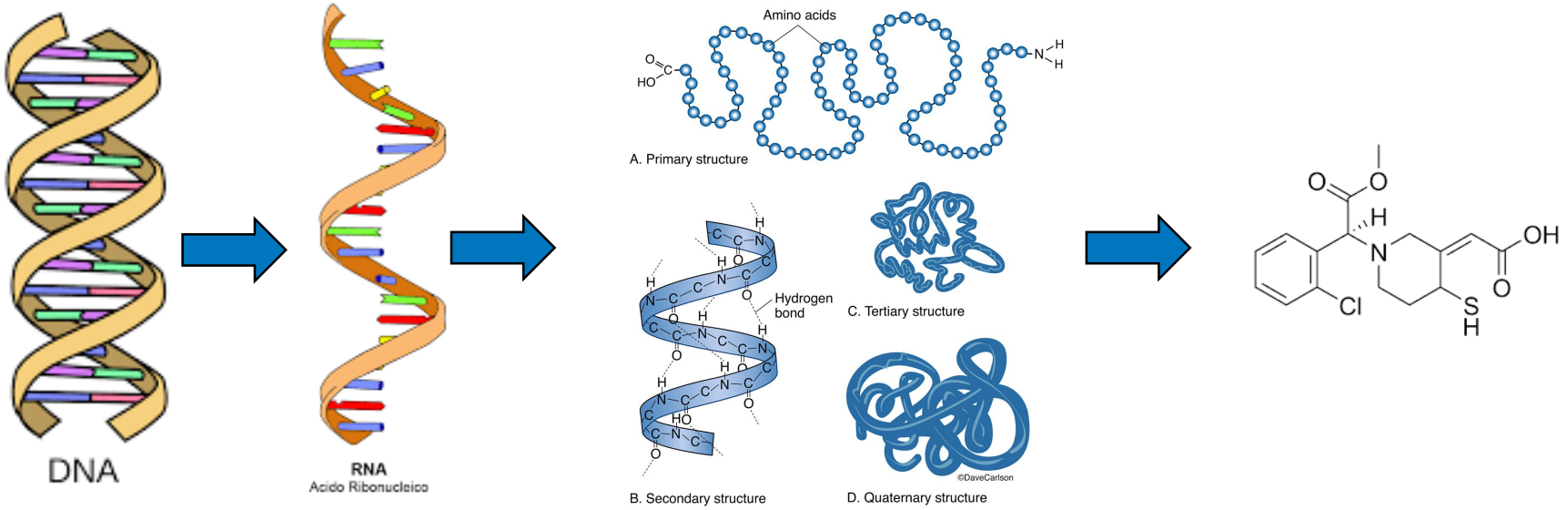
Application of Selective Reaction Monitoring (SRM) Proteomics to Quantify Reductive Dehalogenase Peptides (RDases) in Microbial Consortium SDC-9

K.H. Kucharzyk , J. Meisel, L. Mullins, M. Michelson, P. B. Hatzinger, F. E. Löffler, F. K. Murdoch, J. T. Wilson and J. Istock

April 2019
Bioremediation Symposium

BATTELLE

Relationship Between 'Omics'



GENOMICS

PROTEOMICS

METABOLOMICS

TRANSCRIPTOMICS

FUNCTIONAL GENOMICS

BATTELLE

Conventional and Advanced MBTs

A variety of MBTs are available to assist in the following:

- Determine if remediation is working
- Select the appropriate remediation strategy
- Transition from active treatment to MNA
- Include additional lines of evidence

Application of correct MBT to use must be carefully considered taking into account the following objectives:

- What are your goals?
- What type of data will you need?
- Which phase is the site in?
- What is the geochemical data saying (spatial vs temporal)?
- What are your other lines of evidence telling you?

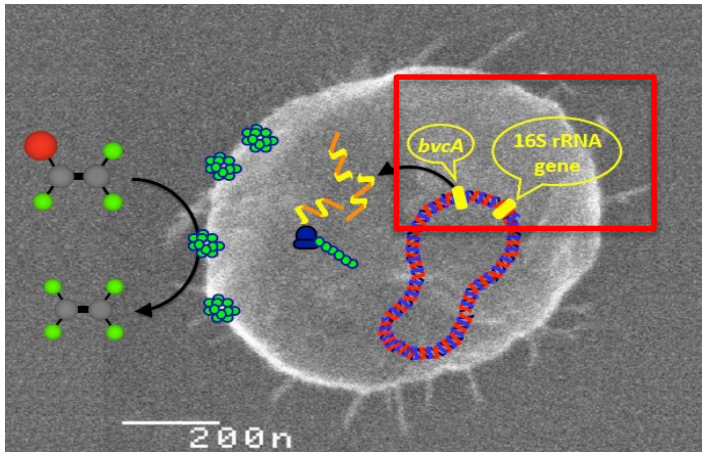
Understanding Biodegradation Rate Is Critical To Site Management

- Measuring biodegradation rates from field samples can be difficult due to:
  - Limited data (sample volume, frequency, replicates)
  - Challenging plume dynamics (insufficient delineation, lithologic heterogeneities, etc.)
- **Proteomic Benefits**:
  - Proteomics can provide additional lines of evidence to further support site recommendations and decision making (e.g., reduction in sampling frequency and/or sampling location(s), elimination of required analytes, site closure, etc..)

DNA-based Molecular Biological Tools Do NOT Provide Quantitative Degradation Rate Information

ORGANISMS → GENES → TRANS-CRIPTS → **PROTEINS** → ACTIVITY (RATE)

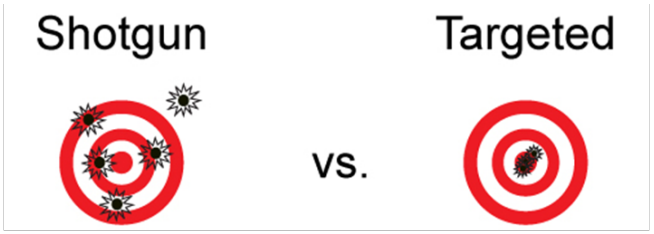
- **qPCR**: gene abundance
- **Metagenomics** (sequence or array)
  - Total microbial community composition
  - Relative gene and/or organism abundance



BATTELLE

Proteomics Provides Information on Functional Activity

ORGANISMS → GENES → TRANS-CRIPTS → **PROTEINS** → ACTIVITY (RATE)

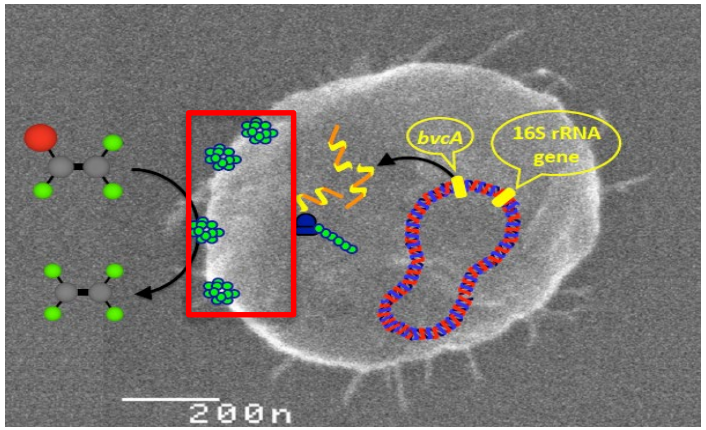


Discovery/Untargeted Shotgun Proteomics

- Ideal for identification of peptides and post translational modifications
- Global proteome mapping

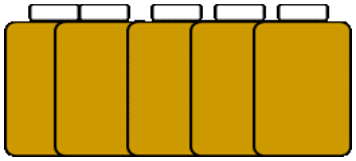
Targeted/Hypothesis Driven Proteomics

- Ideal technique for precise, accurate and reproducible quantification
- Important for large scale projects with large sample sets



Discovery Proteomics & Peptide Detection

1. Culture Growth and Sample Collection



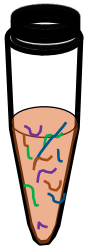
SDC-9 consortium cultivation
Sample at selected time points

Proteins

Cellular lysis
Protein extraction

Spike isotopic peptides
Reduction
Alkylation

2. Sample Processing

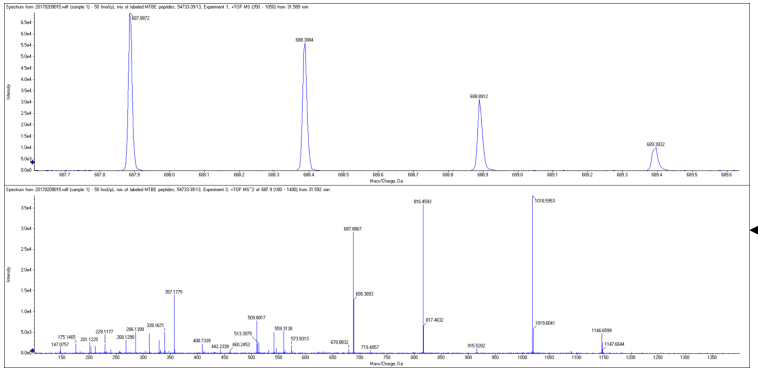


Peptide cleanup

Remove detergent
Desalt

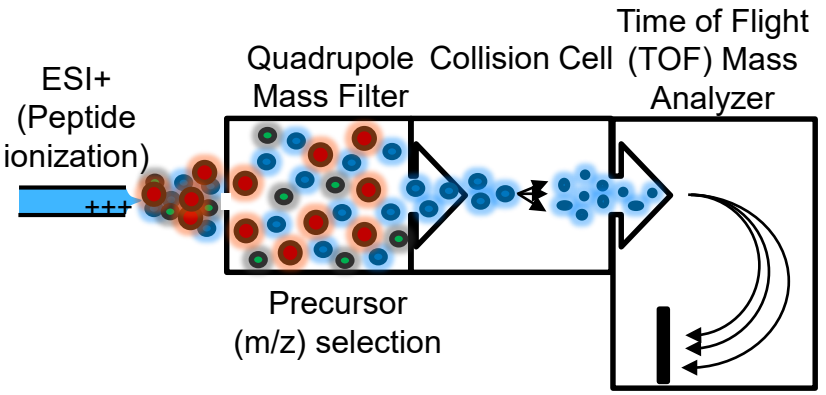
Peptides

4. MS/MS Search & Bioinformatic Analysis



- Identify RDase hits from MS/MS search
- Select peptides specific to RDases
- Build SRM method

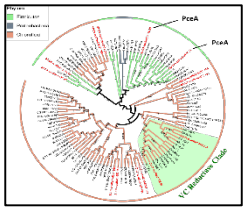
3. LC-MS/MS (Quadrupole Time of Flight, qTOF)



BATTELLE

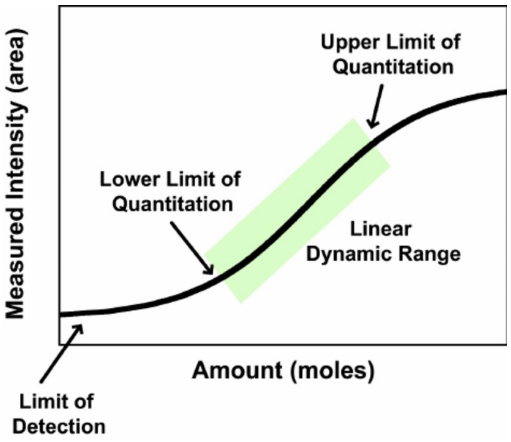
Targeted Proteomics & Peptide Detection

RDase Identification and Down selection



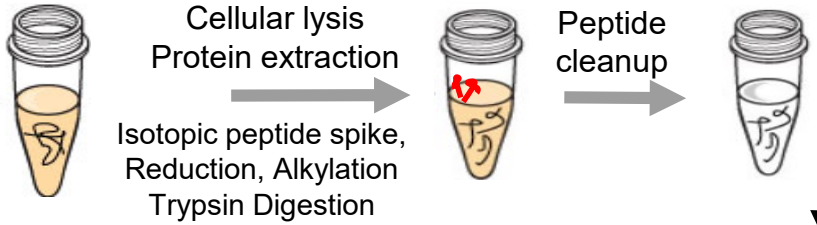
- SDC-9 metagenome sequencing
- Identified a total of 14 RDase gene sequences
- Identified at least 39 unique RDase peptides from the SDC-9 protein digests

Establish IDL and MDL for Selected Peptide Targets



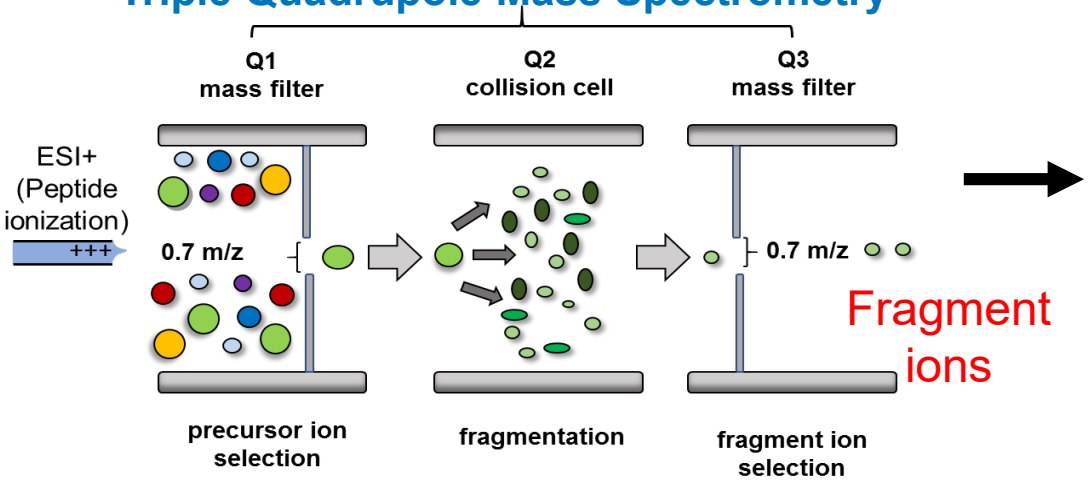
Proteins

Protein Extraction



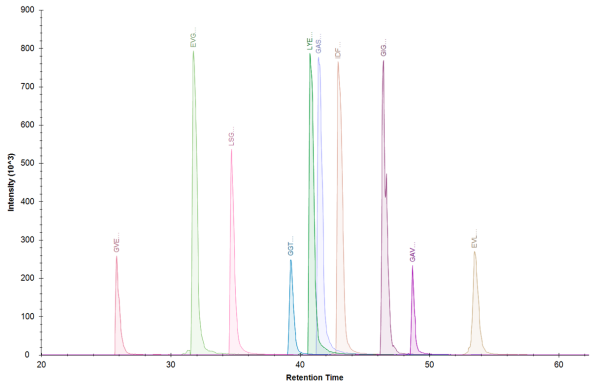
Peptides

Triple Quadrupole Mass Spectrometry



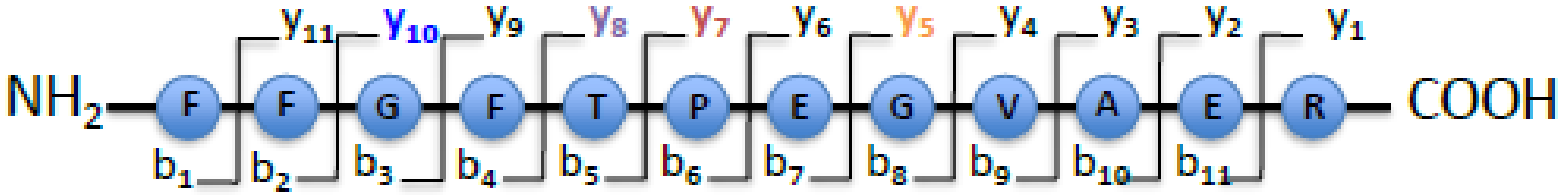
Fragment ions

Bioinformatic Analysis



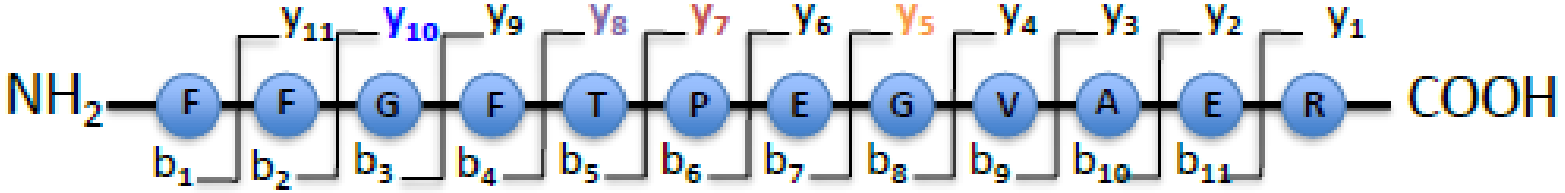
BATTELLE

Principle of Data Acquisition



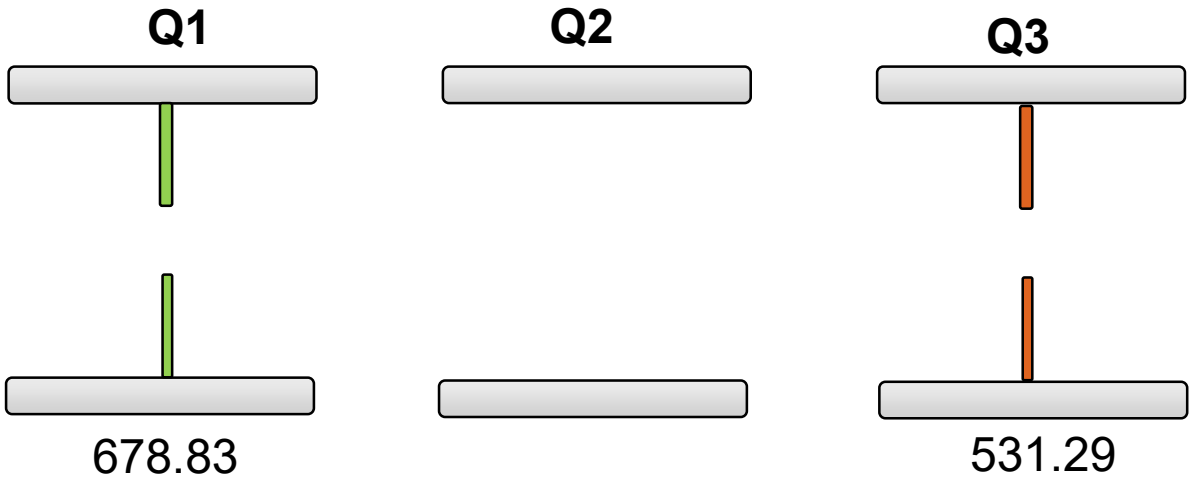
Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250

Principle of Data Acquisition

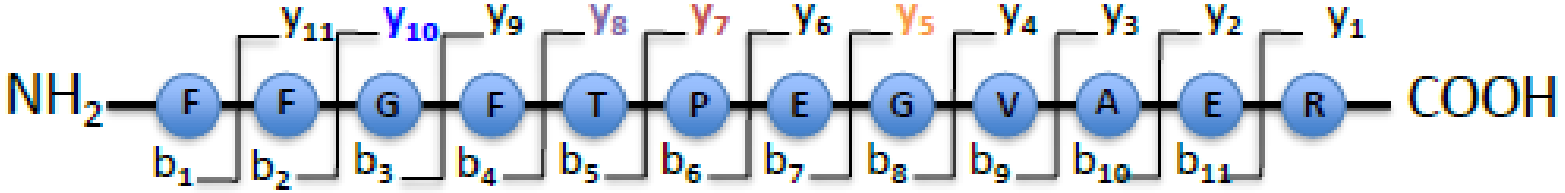


Transition 1

Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250

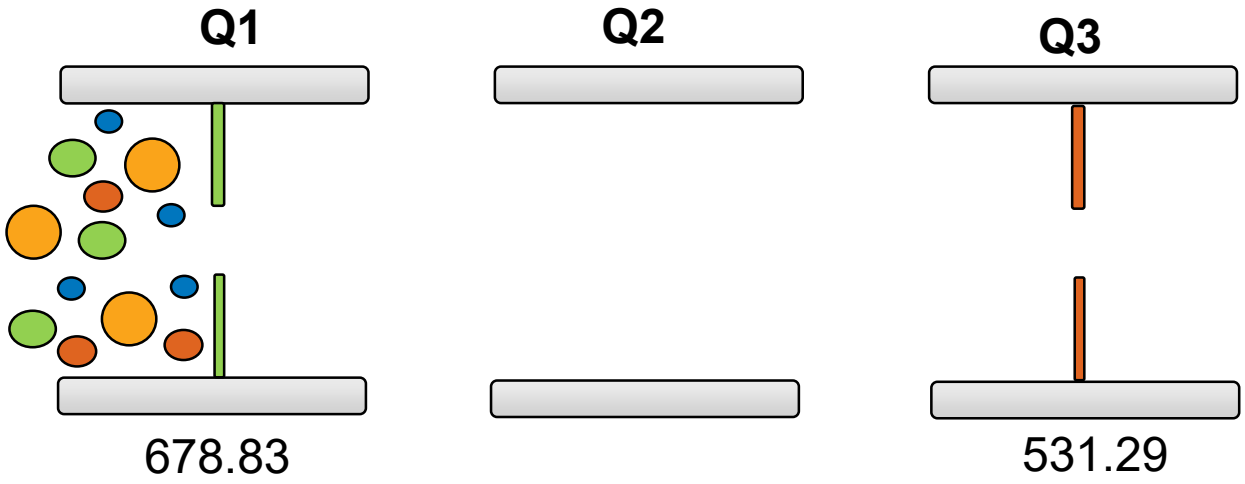


Principle of Data Acquisition

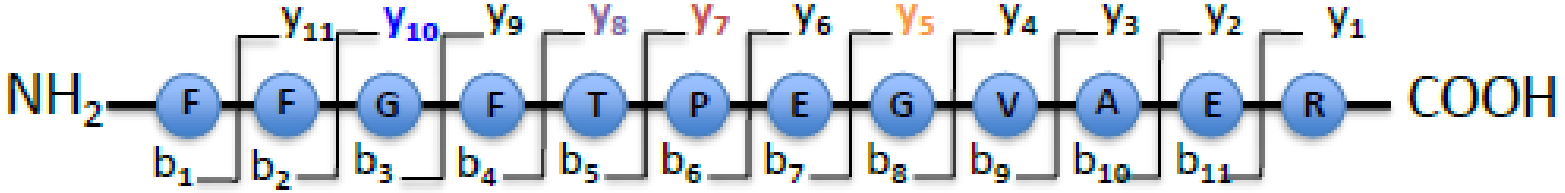


Transition 1

Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250

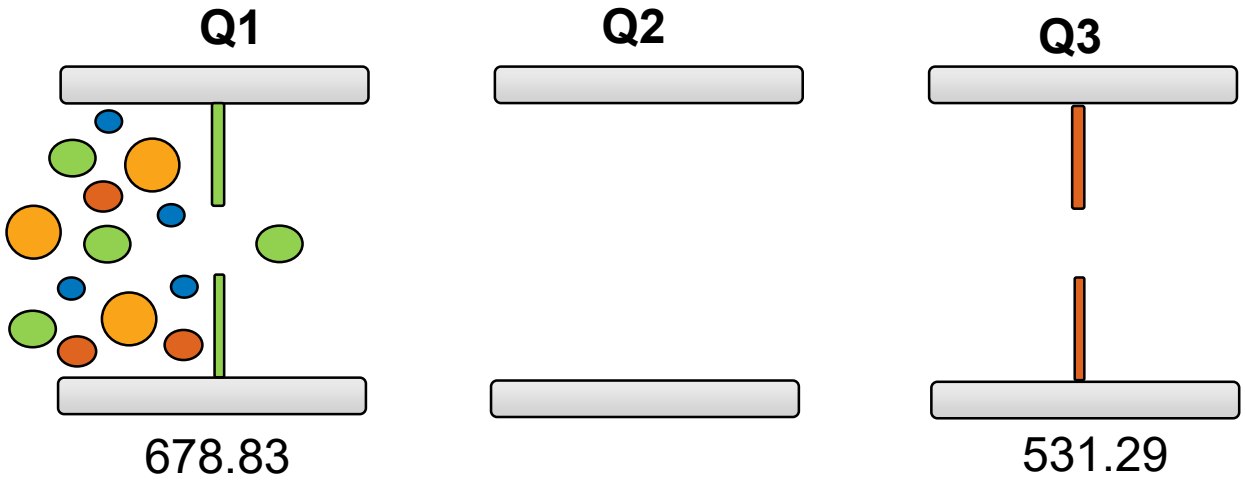


Principle of Data Acquisition

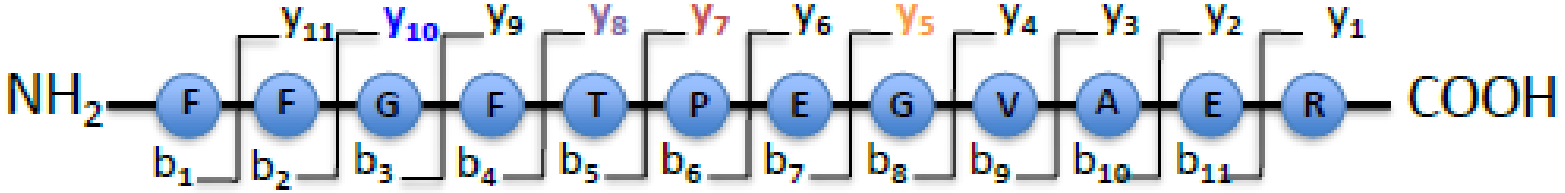


Transition 1

Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250

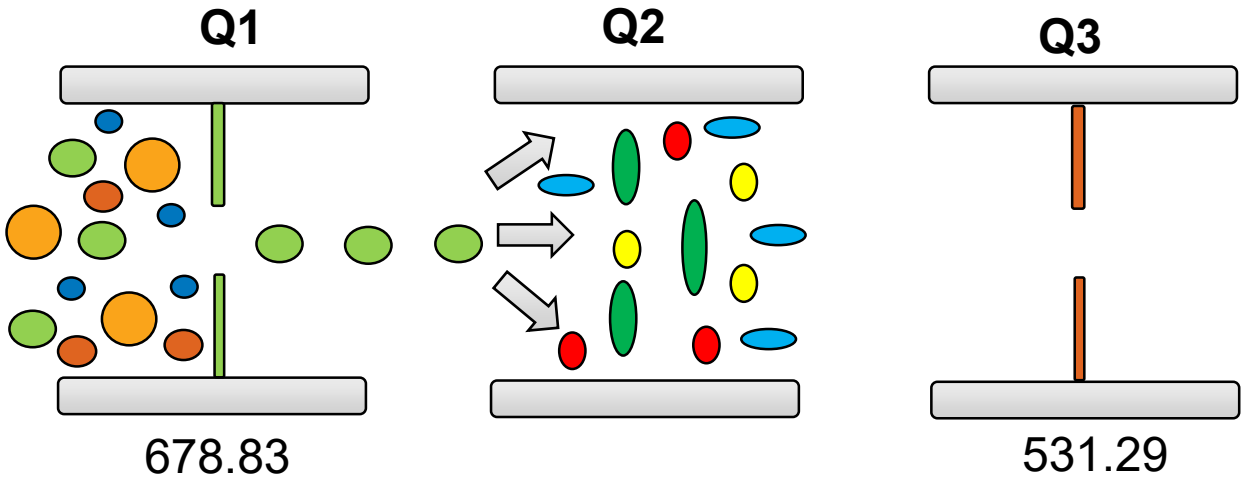


Principle of Data Acquisition

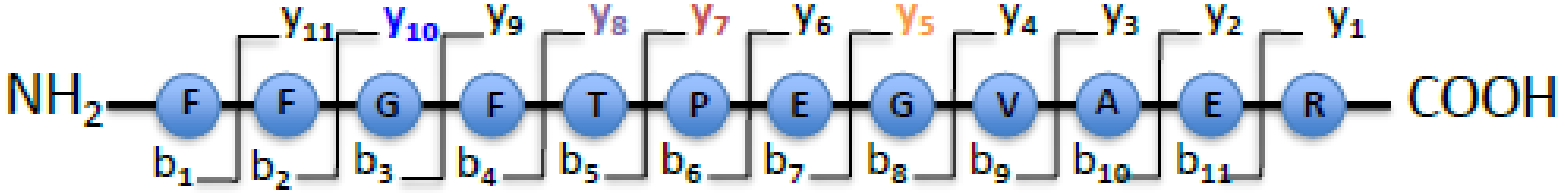


Transition 1

Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250

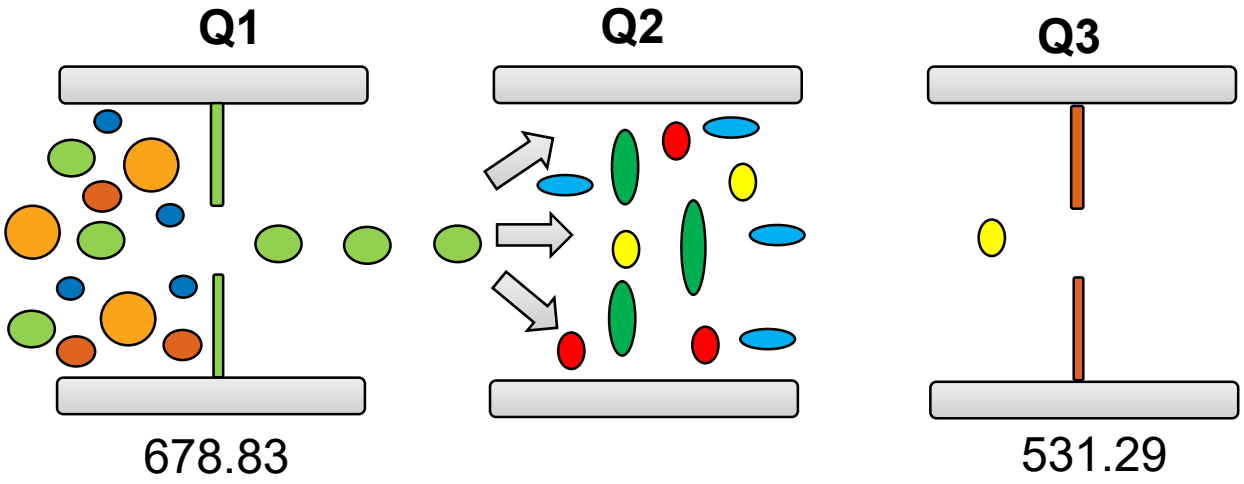


Principle of Data Acquisition

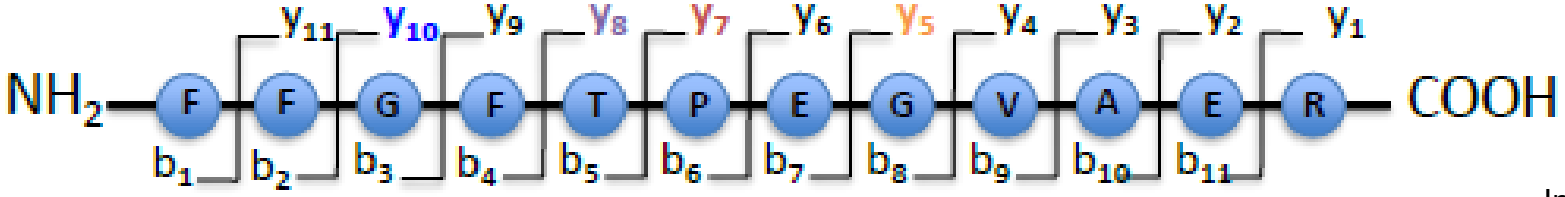


Transition 1

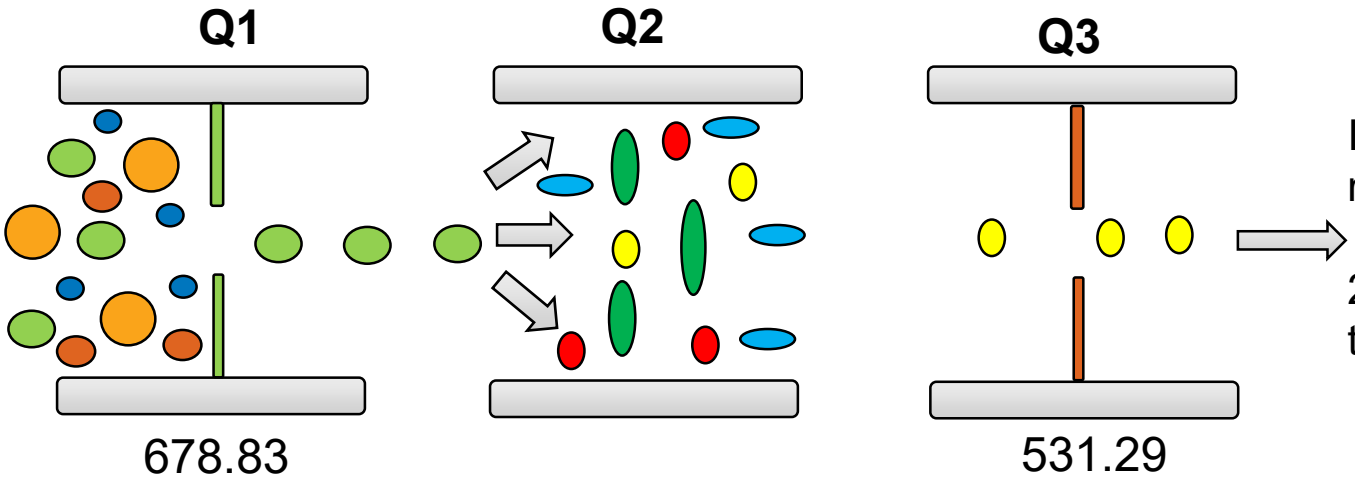
Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250



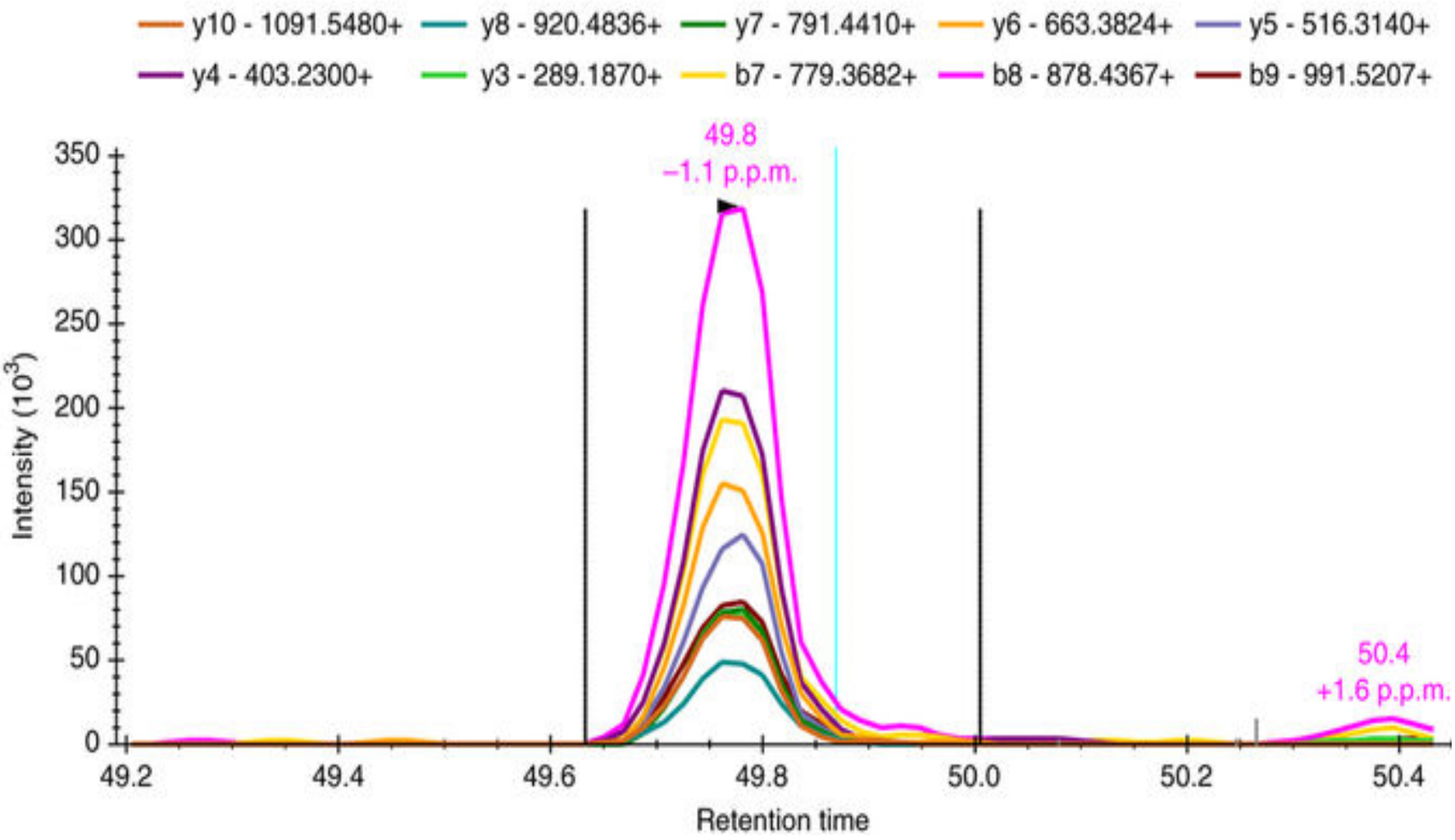
Principle of Data Acquisition



Sequence	Q1 [m/z]	Q3 [m/z]	fragmentation	CE [eV]	Dwell Time [ms]	Intensity time point x
Transition 1						
FFGFTPEGVAER	678.83	531.29	Y5	26.4	250	2676
FFGFTPEGVAER	678.83	757.38	Y7	26.4	250	
FFGFTPEGVAER	678.83	858.43	Y8	26.4	250	
FFGFTPEGVAER	678.83	1062.52	Y10	26.4	250	



Principle of Data Acquisition SRM Method



Peptide Query Parameters for Targeted Proteomic Experiment

Target Proteins

Target Peptides

Protein Extraction

Cellular lysis
Protein extraction

Isotopic peptide spike,
Reduction, Alkylation
Trypsin Digestion

Peptide
cleanup

***Parameters strictly required to develop targeted proteomic method**

Triple Quadrupole Mass Spectrometry

Precursor m/z

Fragment Ion m/z

Q1
mass filter

Q2
collision cell

Q3
mass filter

ESI+
(Peptide
ionization)
+++

0.7 m/z

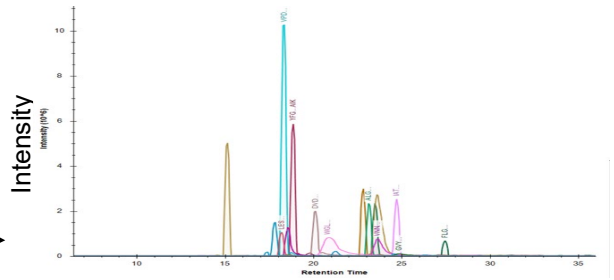
0.7 m/z

precursor ion
selection

fragmentation

fragment ion
selection

Total Ion Chromatogram

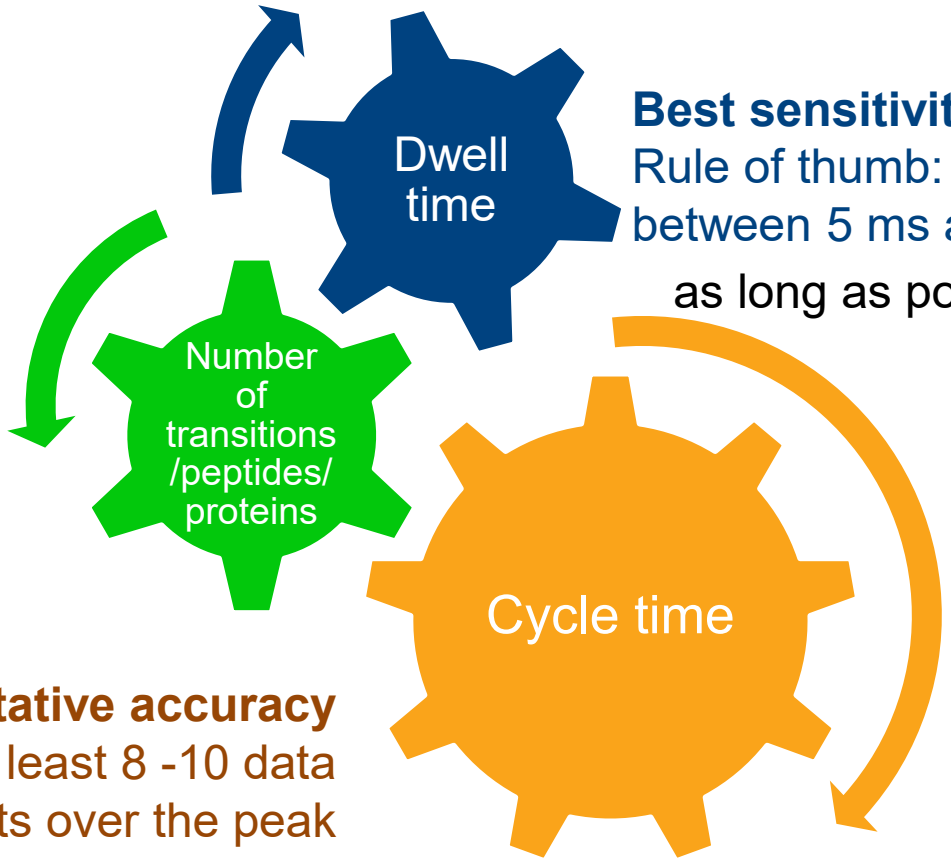


Retention time (min)

BATTELLE

SRM Methods is a Balancing Act

Goal: Achieve the best sensitivity at a high quantitative accuracy with as many proteins as possible!



Best sensitivity
Rule of thumb: Dwell times between 5 ms and 200 ms

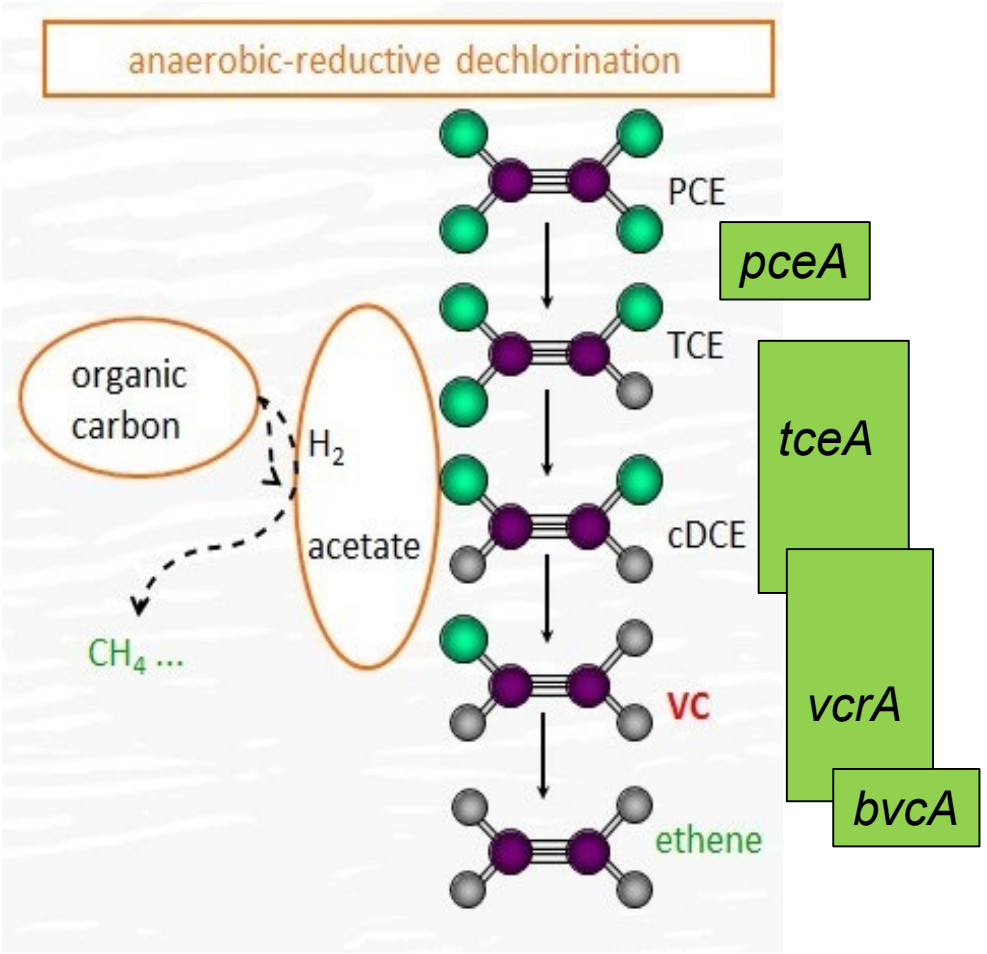
as long as possible

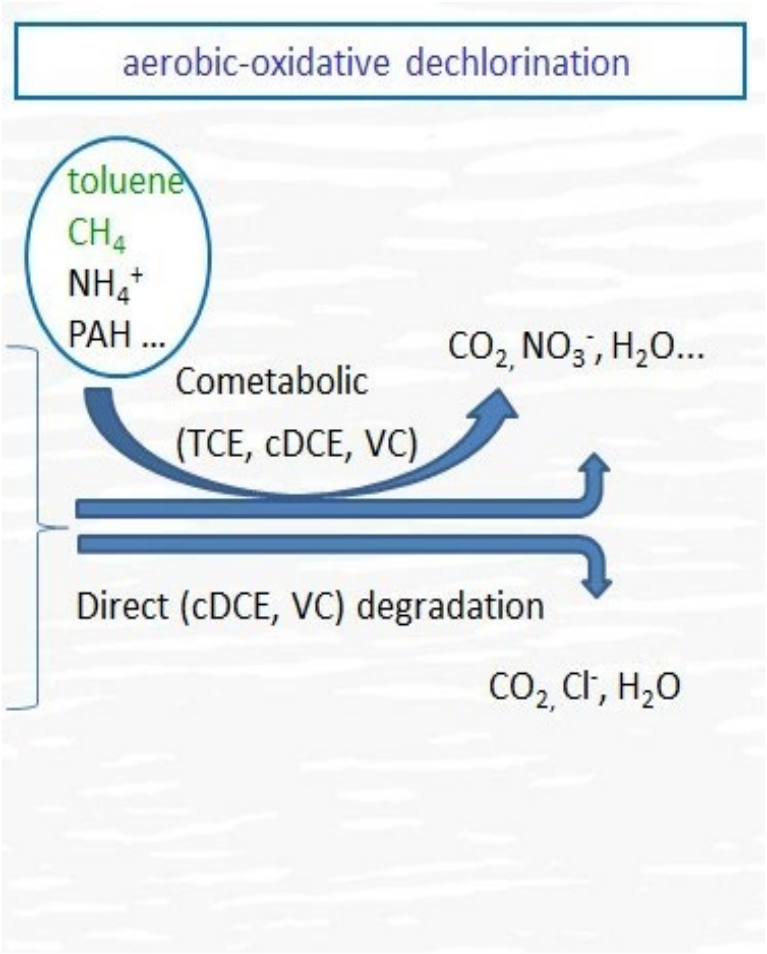
Maximal multiplexing
Rule of thumb: 4 to 6 transitions per peptide, 3 to 5 peptides per protein

High quantitative accuracy
Rule of thumb: At least 8 -10 data points over the peak

as short as possible

Microbial Dehalogenation of CVOCs





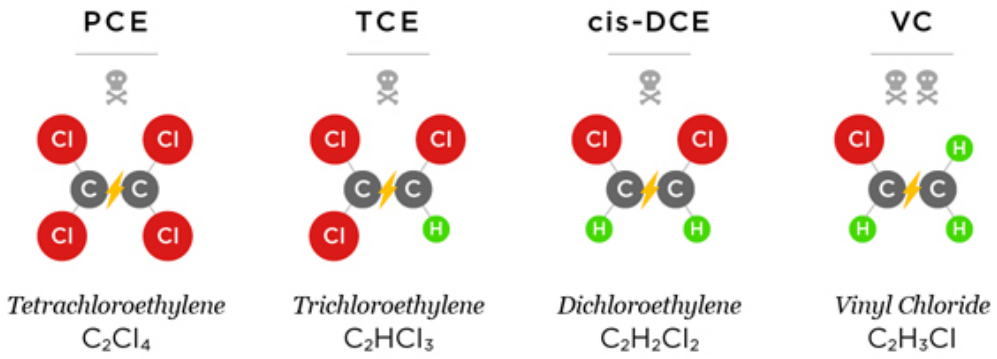
Dehalococcoides

BATTELLE

Microbial Dehalogenation of CVOCs

ESTCP

ER201726



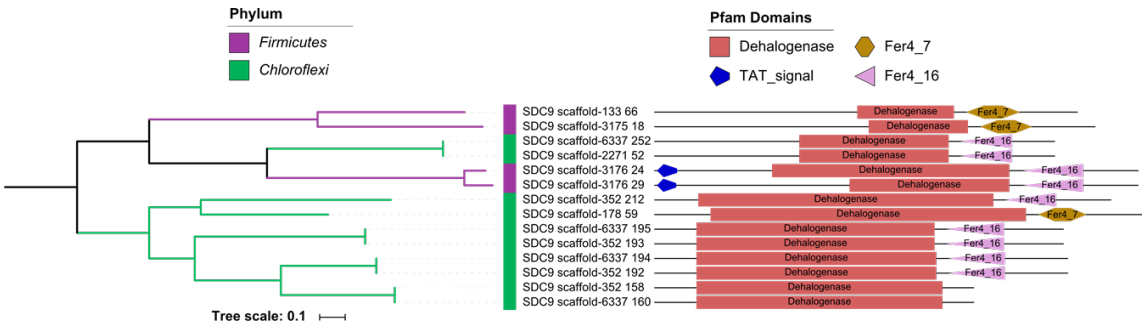
TCE → cDCE → VC → ethene

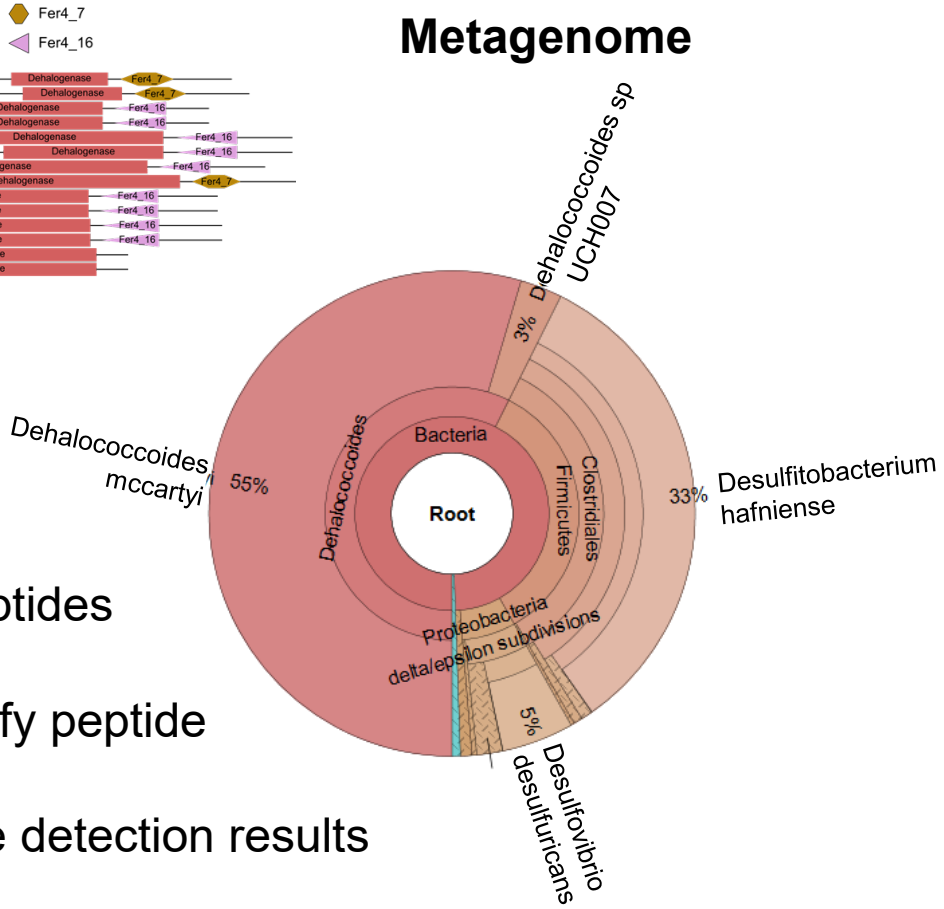
Microbe	Other	*tceA*	*tceA*	*bvcA*	*tceA*	*vcrA*	*bvcA*
Desulfitobacterium, Dehalobacter, Sulfurospirillum, Desulfuromonas, Geobacter	M	-	-	-	-	-	-
Dehalococcoides mccartyi strain 195, *Dehalococcoides* sp. strain FL2	-	M	M	-	C	-	-
Dehalococcoides sp. strain GT, VS	-	-	-	-	-	M	-
Dehalococcoides sp. BAV1	-	-	-	M	-	-	M

Key genes in reductive dechlorination of of chlorinated ethenes; M = metabolic, C = cometabolic *Source: J.Barnes (conference talk)*

Michalsen, M.M., Kucharzyk, K.H., Meisel, J.E., Hatzinger, P., Loffler, F., Wilson, J., Istok, J. Validation of Advanced Molecular Biological Tools for Monitoring Chlorinated Solvent Bioremediation and Estimating Degradation Rates. Eleventh International Conference on Remediation of Chlorinated and Recalcitrant Compounds (Palm Springs, California; April 8-12, 2018).

Microbial Dehalogenation of CVOCs





- SDC-9™ metagenome sequenced;
- RDase genes identified and RDase peptides annotated
- Correct* RDase peptides used to identify peptide targets in SDC-9 proteomics
- Approach avoids false negative peptide detection results

*metagenomic guided proteomics reduces probability of misalignment of peptides or using incorrectly annotated sequences from NCBI

Metagenomics and Shotgun Proteomics

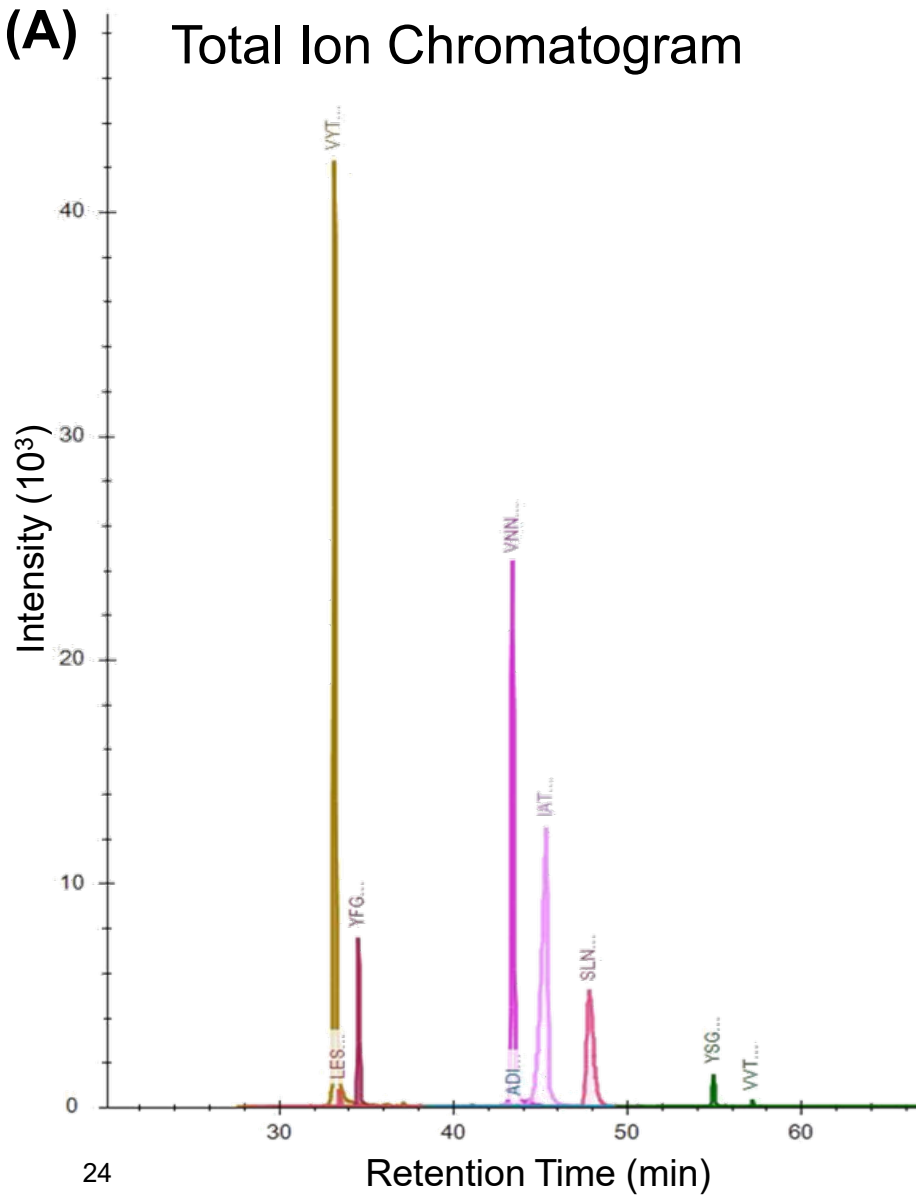
RDase #	RDase Identifier	Host	Accession # of closest RDase in Database	Query Coverage	% Amino Acid Identity	Predicted function
1	6337_195	DHC	WP_058292018.1	96%	99%	?
2	6337_194	DHC	KSV18849.1	99%	100%	?
3	352_158	DHC	KSV18948.1	98%	100%	?
4	6337_252	DHC	WP_010935983.1	99%	100%	?
5	352_212	DHC	AEI59454.1	99%	99%	VcrA
6	178_59	DHC	WP_062900263.1	99%	99%	TceA
7	6337_160	DHC	KSV18948.1	98%	100%	?
8	2271_52	DHC	WP_010935983.1	99%	100%	?
9	352_192	DHC	KSV18849.1	99%	100%	?
10	352_193	DHC	WP_058292018.1	96%	99%	?
11	3176_24	Dsf	CAD28790.2	99%	94%	PceA
12	133_66	Dsf	WP_015043198.1	98%	40%	?
13	3175_18	Dsf	CDX01551.1	99%	100%	?
14	3176_29	Dsf	WP_025206074.1	99%	82%	PceA

32 unique RDase peptides identified, 14 down-selected
Targets verified by spiking isotopically labeled peptides

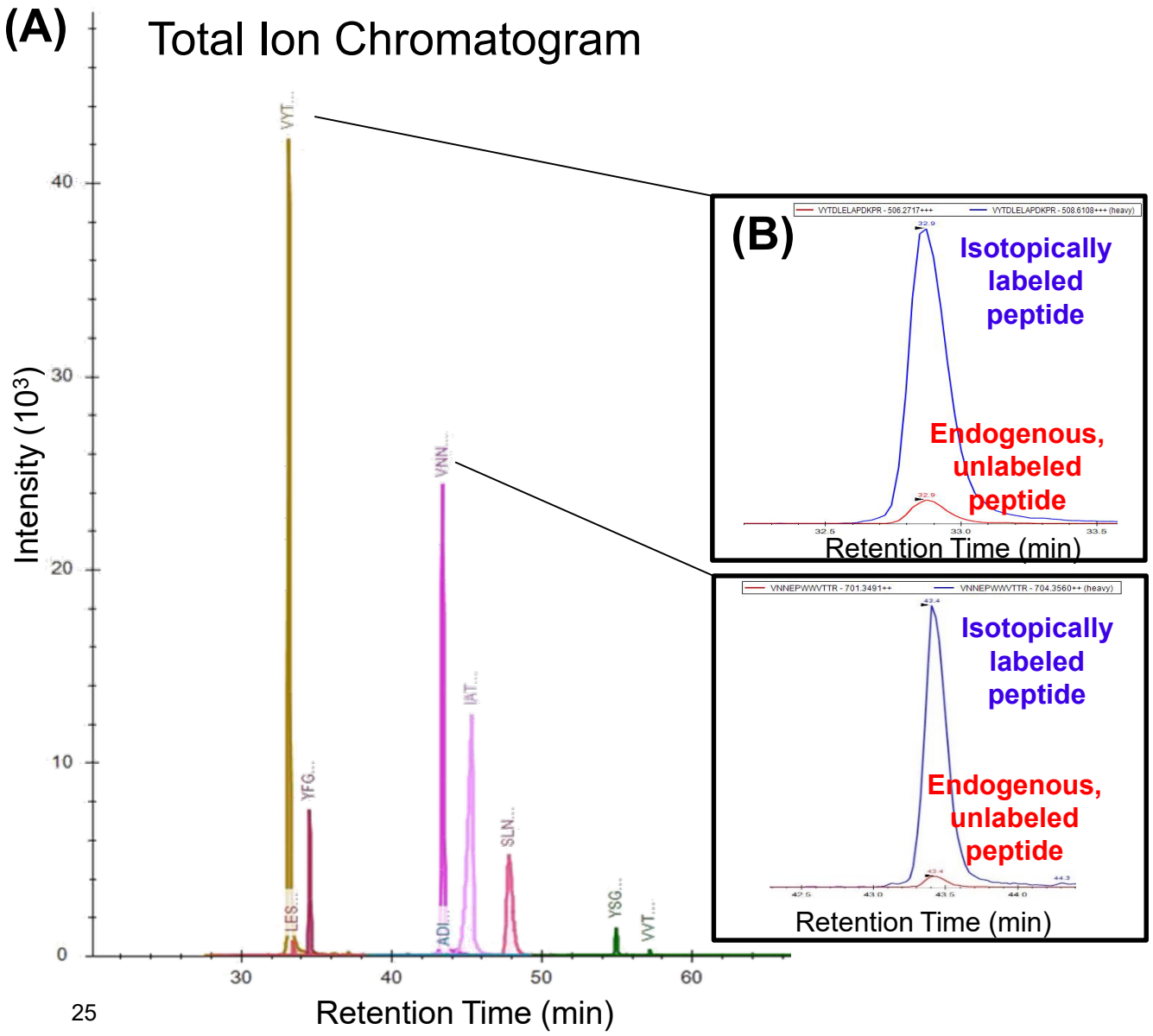
Protein and Peptide Targets

Protein	Peptide ID	Peptide Sequence
FdhA	FdhA 2	SGSEIAFTGGLIK
	FdhA 5	ALGIVYLDSQAR
PceA	PceA4	IATQIPLLQDAAR
	PceA5	LESGYVQNMVK
	PceA7	DFWNNPEPIK
	PceA8	TSPSLISSATVGK
TceA	TceA2	DVDDLLSAGK
	TceA3	VSSIIEPR
	TceA4	VNNEPWWVTTR
	TceA5	YFGASSVGAIK
VcrA	VcrA1	WGLYGPPHDSAPPDGSVPK
	VcrA2	YFGAGDVGALNLADPK
	VcrA3	VPDHAVPINFK
	VcrA4	GVYEGPPDAPFTSWGNR
	VcrA6	DQPWYYVK

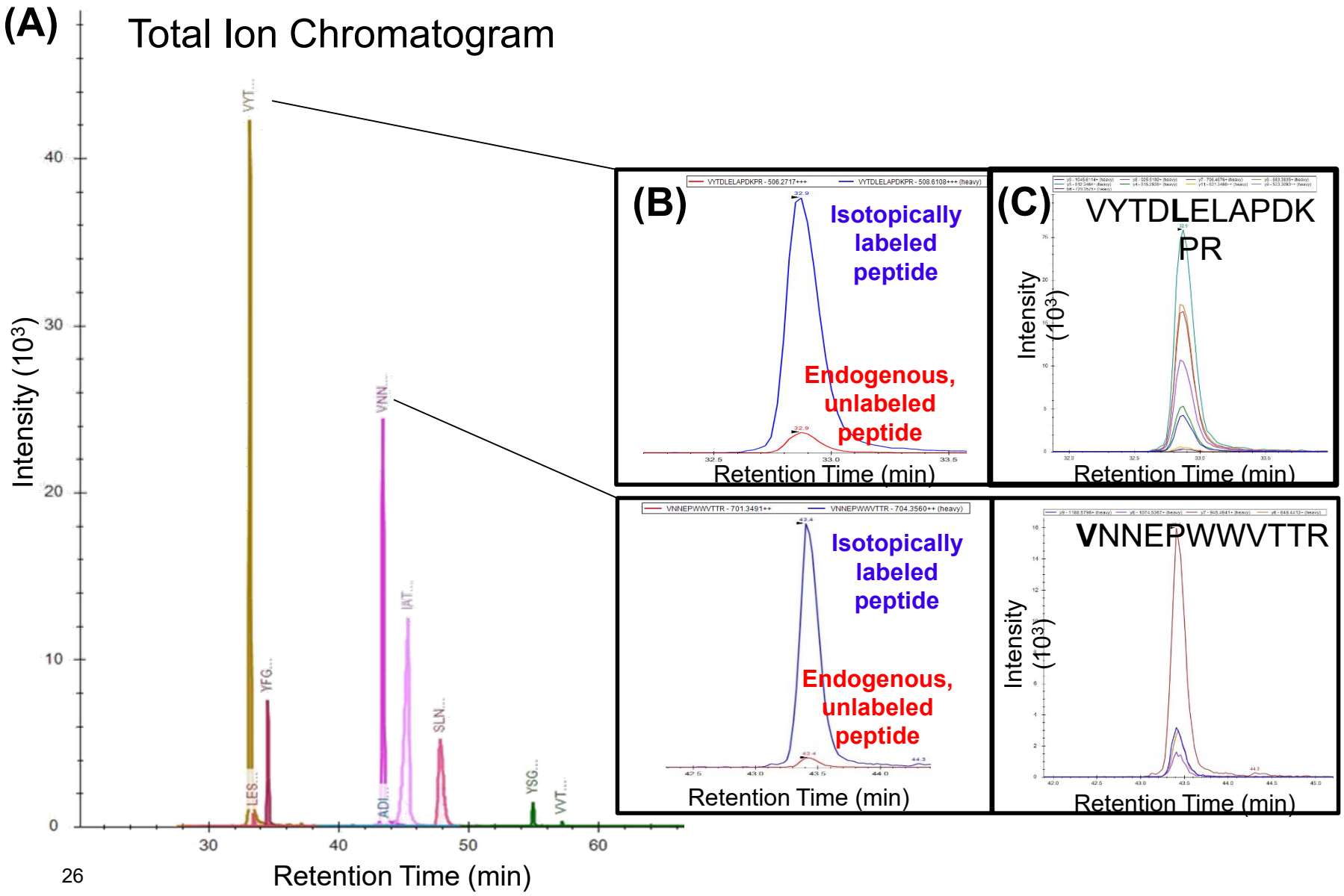
Real life SRM



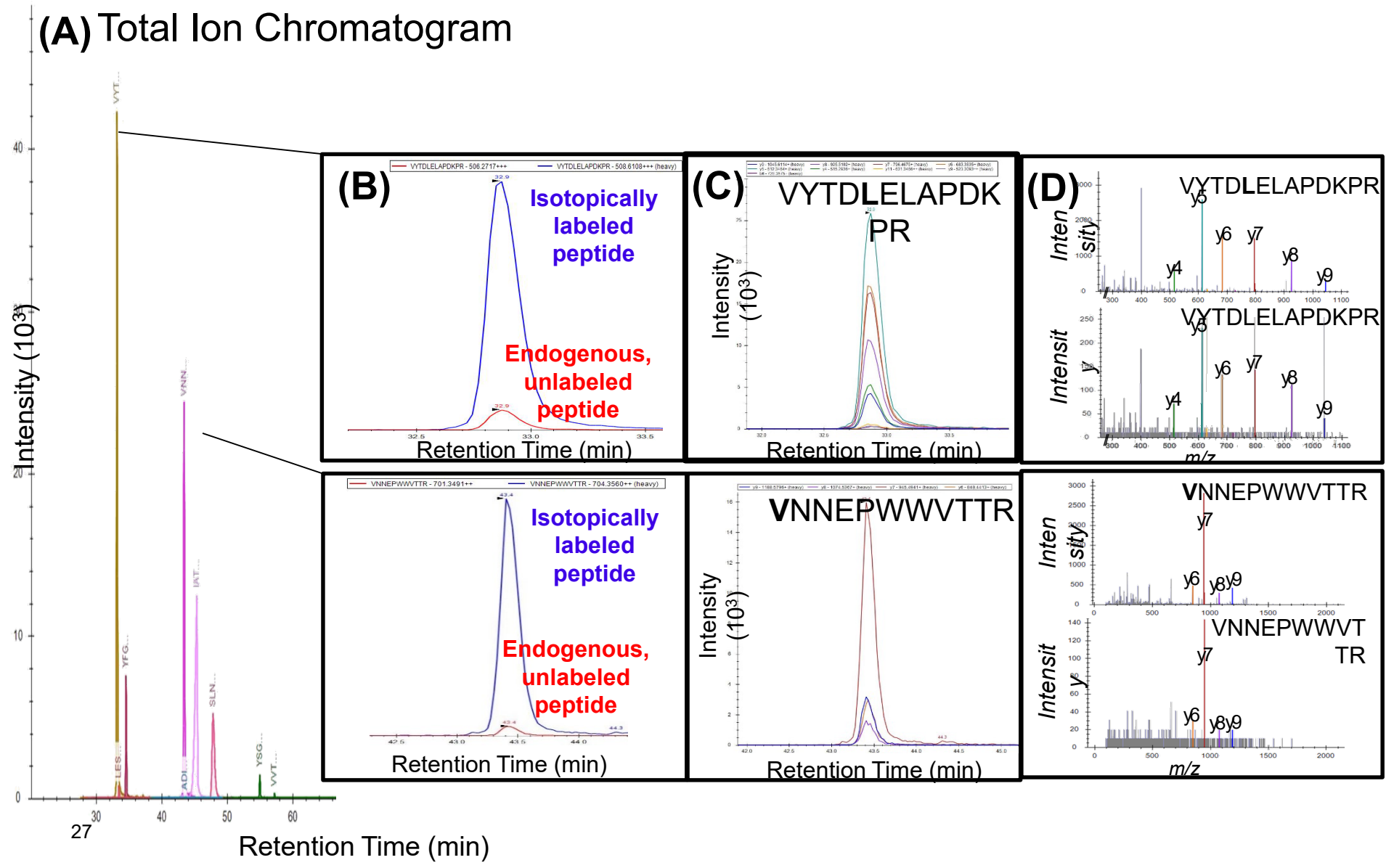
Real life SRM



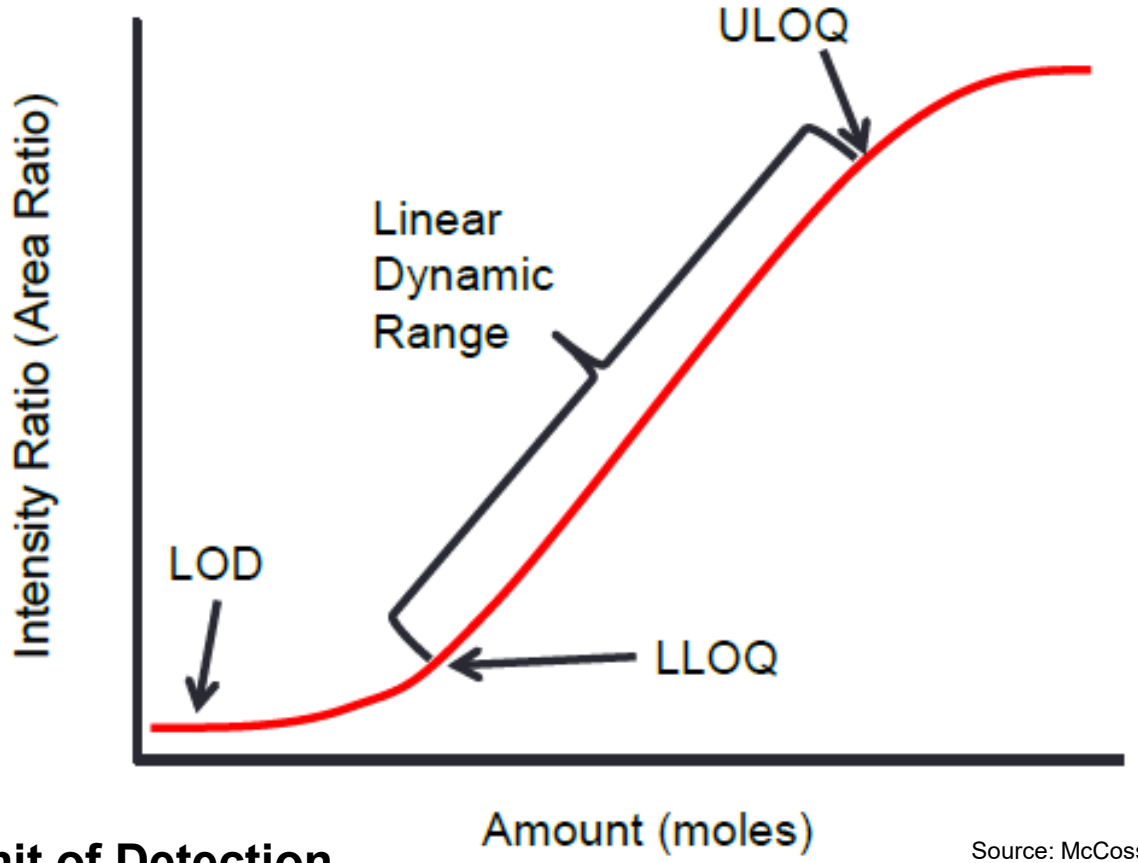
Real life SRM



Real life SRM



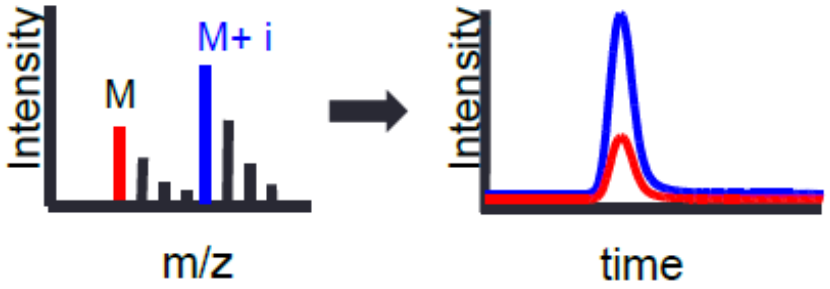
Quantitative Analysis



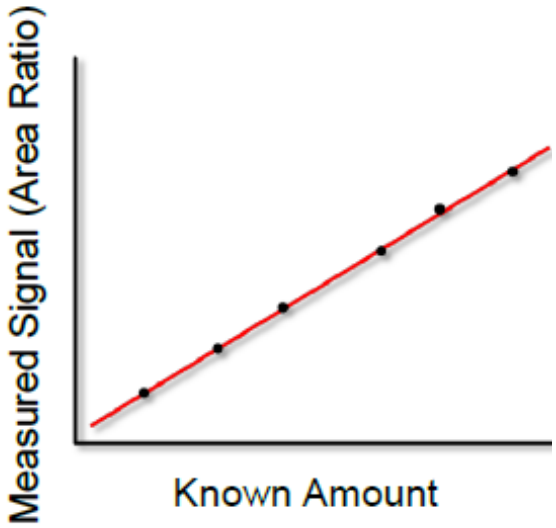
LOD: Limit of Detection
LLOQ: Lower Limit of Quantitation
ULOQ: Upper Limit of Quantitation

Source: McCoss Lab, Skyline Workshop 2017, University of Washington

Quantitative Analysis



$$R_0 = \frac{A_0}{A_i} \propto \frac{n_0}{n_i}$$



$$R_0 = k_{0/i} \cdot n_0 + R_i \qquad n_0 = \left(\frac{R_0 - R_i}{k_{0/i}} \right)$$

Where: $k_{0/i}$ = is the slope of the standard curve $R_i \approx 0$ Area ratio from a blank … only internal std

Source: McCoss Lab, Skyline Workshop 2017, University of Washington

BATTELLE

Instrument Detection Limit (IDL) vs Method Detection Limit (MDL)

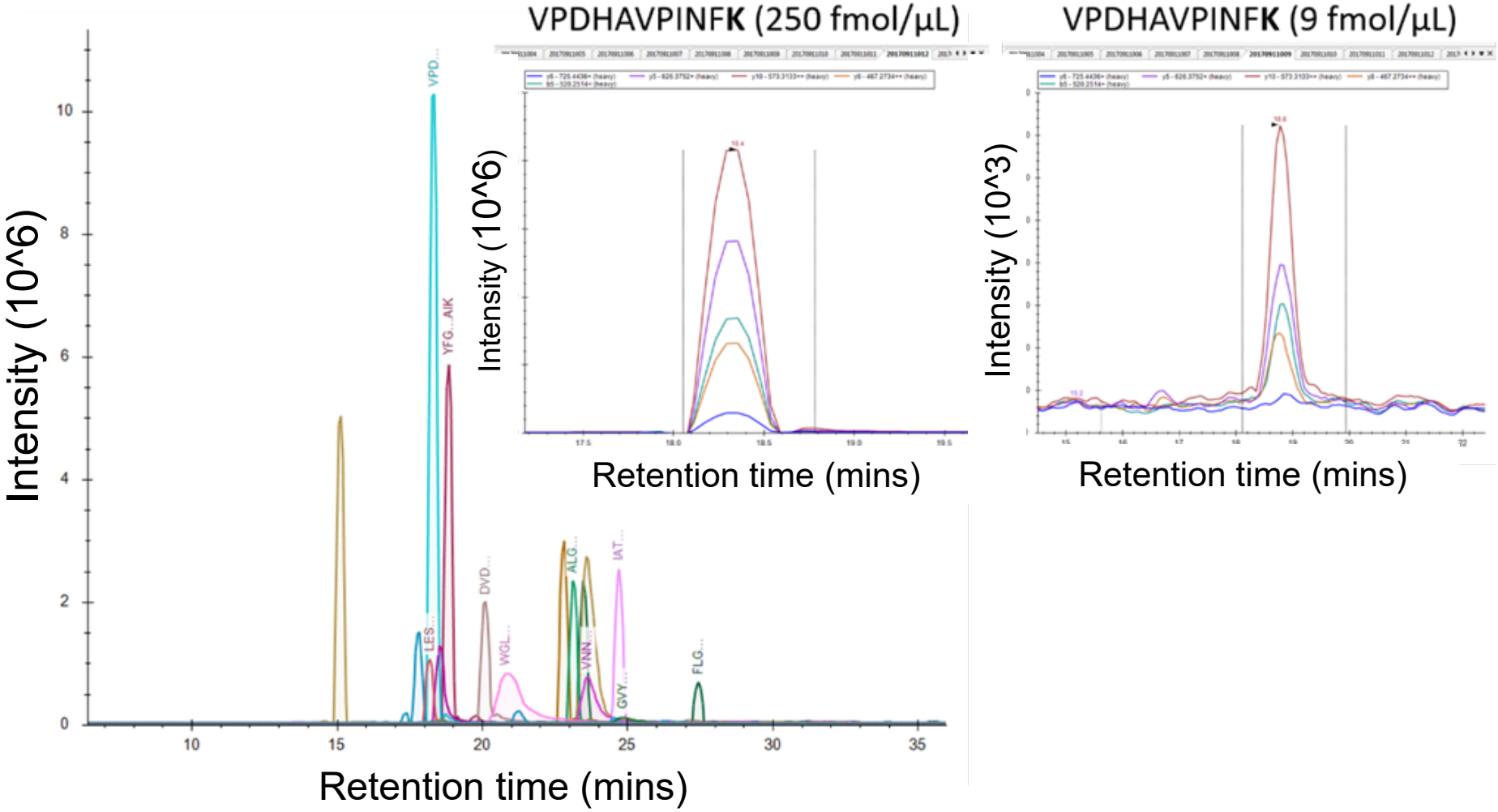
IDL = Instrument Detection Limit

- Peptides injected at a range of concentrations (0.1 fmol/ul to 250 fmol/ul)
- Instrument variability check
- Lower level detection limit established per peptide

MDL = Method Detection Limit

- Peptides injected at a range of concentrations (0.1 fmol/ul to 250 fmol/ul)
- Peptides prepared as experimental samples (extraction, digestion, cleanup)
- Samples run with the optimized SRM method
- Lower level of quantification (LLOQ) and level of detection (LOD) established for each peptide

Method Detection Limit (MDL)



Method Detection Limit (MDL)

Protein	ID	Peptide[1]	MDL 1	MDL2	MDL3	Established MDL
			fmol/mL			
FdhA	FdhA2	SGSEIAFTGGLIK	3	3	3	**3**
	FdhA5	ALGIVYLDSQA**R**	3	3	1	**3**
	FdhA8	NQAVSAPGEA**K**	3	3	3	**3**
PceA	PceA4	IATQIPLLQDAA**R**	9	9	9	**9**
	PceA5	LESGYVQNMV**K**	3	3	3	**3**
	PceA7	DFWNNPEPI**K**	1	1	1	**1**
	PceA8	TSPSLISSATVG**K**	0.3	0.3	1	**1**
TceA	TceA2	DVDDLLSAG**K**	0.3	3	3	**3**
	TceA3	VSSIIEP**R**	0.3	0.3	1	**1**
	TceA4	VNNEPWWVTT**R**	9	9	9	**9**
	TceA5	YFGASSVGAI**K**	0.3	0.3	1	**1**
VcrA	VcrA1	WGLYGPPHDSAPPDGSVP**K**	9	9	3	**9**
	VcrA2	YFGAGDVGALNLADP**K**	27	27	27	**27**
	VcrA3	VPDHAVPINF**K**	0.3	0.3	1	**1**
	VcrA4	GVYEGPPDAPFTSWGN**R**	83	27	27	**83**
	VcrA6	DQPWYV**K**	1	1	1	**1**

[1] Bolded letters denote heavy ^{13}C and ^{15}N labeled amino acid; the maximum of three MDL test replicates was established as the MDL

RDase Concentration Detection in Pure Cultures

Peptide ID	Peptide Sequence	Previously reported concentrations*	Approximate DHC cell concentrations		
			10^7	5.25E+07	1.71E+07
			SDC-9	DHC 195	DHC FL-2
FdhA 2	SGSEIAFTGGLIK	KB1 culture (TCE): 3100 - 3500 fmol/mL; D2 culture: 2300-3500 fmol/mL	3,085.0	115.26	#N/A
FdhA 5	ALGIVYLDSQAR		4,162.5	207.70	6.37
FdhA 8	NQAVSAPGEAK		#N/A	#N/A	#N/A
PceA4	IATQIPLLQDAAR	D2 culture: 45 fmol/mL;	4,617.5	#N/A	#N/A
PceA5	LESGYVQNMVK		1,572.5	#N/A	#N/A
PceA7	DFWNNPEPIK		11,670.8	#N/A	#N/A
PceA8	TSPSLISSATVGK		#N/A	#N/A	#N/A
TceA2	DVDDLLSAGK	KB1 culture (TCE): 300 fmol/mL; D2 culture (PCE): 850-2300 fmol/mL	13,948.8	50.51	5.51
TceA3	VSSIIEPR		2,074.3	367.01	#N/A
TceA4	VNNEPWWVTTR		#N/A	#N/A	#N/A
TceA5	YFGASSVGAIK		#N/A	235.77	21.34
VcrA1	WGLYGPPHDSAPPDGSVPK	Difficult to quantify due to low peptide sensitivity	#N/A	#N/A	#N/A
VcrA2	YFGAGDVGALNLADPK		#N/A	#N/A	#N/A
VcrA3	VPDHAVPINFK		397.5	#N/A	#N/A
VcrA4	GVYEGPPDAPFTSWGNR		#N/A	#N/A	#N/A
VcrA6	DQPWYVK		#N/A	#N/A	#N/A

Conclusions

- Sequenced SDC-9 genome
- Annotated RDase genes and 32 RDase peptides
- Found 14 conservative RDases – PRM candidates
- Identified all of the isotopically labelled RDases in culture samples
- Quantified RDases in culture samples

Battelle Team

Dr. Jayda Meisel

Dr. Craig Bartling

Larry Mullins

Steve Rosansky, PPE

Dr. Kate Kucharzyk

NAVFAC NW
Dr. Malcolm Gander

USACE ERDC
Dr. Mandy Michalsen
Dr. Fiona Crocker

Aptim
Dr. Paul Hatzinger
Dr. Mark Fuller

Scissortail Solutions
Dr. John Wilson

University of Tennessee-Knoxville
Dr. Frank Loeffler
Dr. Fadime Kara Murdoch
Dr. Steve Higgins

University of Oregon
Dr. Jack Istook

NAVFAC EXWC
Ms. Kenda Neil

Conoco-Philips
Dr. Jennifer Harris-Bush

Thank you!

Questions:

kucharzyk@battelle.org

BATTELLE

It can be done