Overcoming LC/MS Data Overload in Biotherapeutic Peptide Mapping

Broadcast Date: Wednesday, November 18, 2009
Time: 1:00 pm EDT
Sponsored by

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Scott Berger, Ph.D.
Market Development Manager
Waters

Adam W. Lucka, Ph.D.
Scientist II
Alexion Pharmaceuticals

Yan Wang, Ph.D.
Scientist 1
Biogen Idec

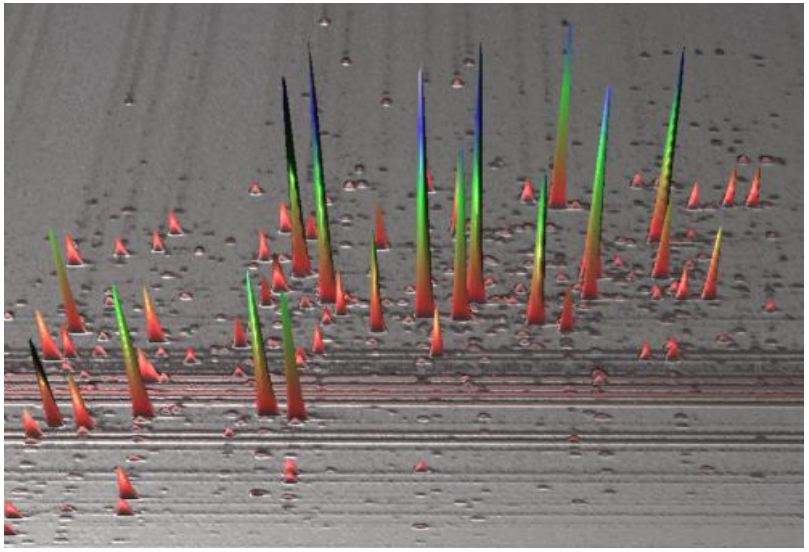
John Sterling
Editor-in-Chief,
Genetic Engineering & Biotechnology News

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Embracing the complexity of LC/MS peptide map analysis to drive better informatics.

Scott J. Berger,
Biopharmaceutical Sciences Group, Waters Corporation

Manhattan, NYC



Tryptic Enolase, Yeast

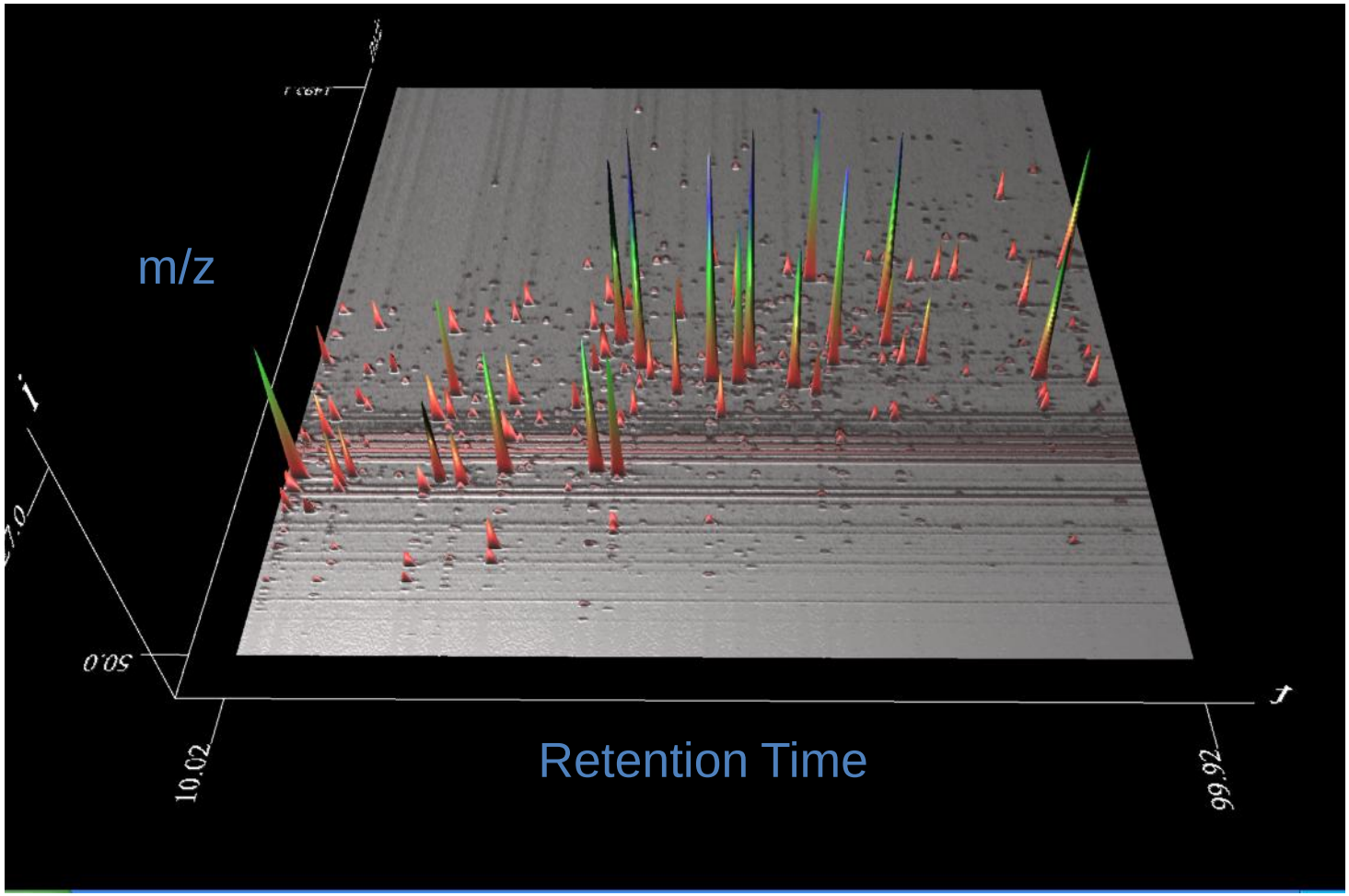
LC/(MS) Peptide Mapping of Biotherapetics

- Structural Characterization
  - 1° Do peptides conform to the predicted sequence?
  - 3° Are my disulfides in proper arrangement?
- Detecting and Monitoring Modifications
  - Co- and Post-translational modifications
    - Glycosylation, Proteolytic maturation, disulfide bonds
  - Determine product and process related impurities
    - Deamidations, oxidations, pyE, truncations, etc.
    - Host Cell Proteins

Desire for Analytical Completeness

Data analysis, not data acquisition, is the productivity limiting step

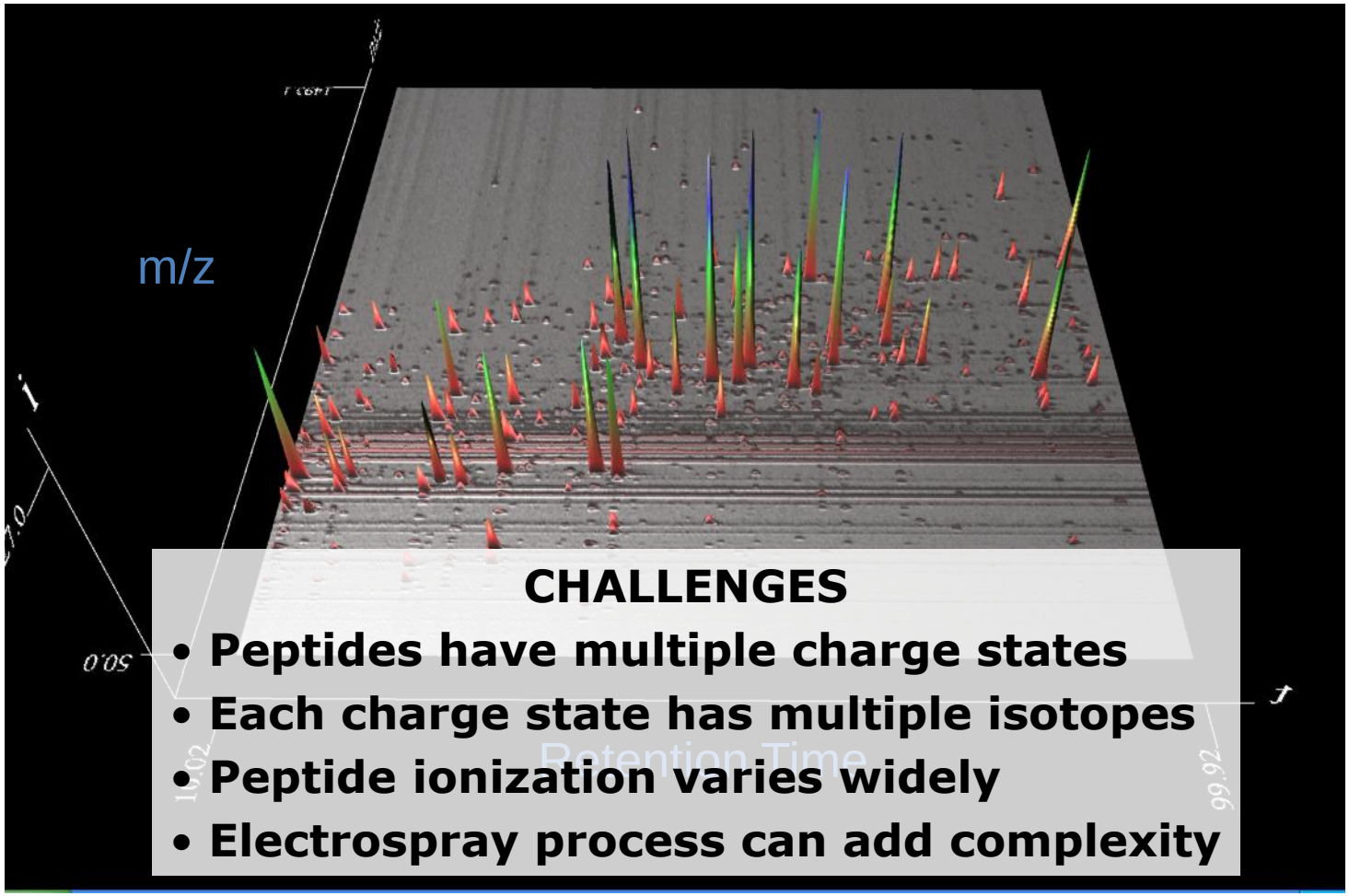
LC/ESI-TOF MS Peptide Map



Yeast Enolase (47 kDa)

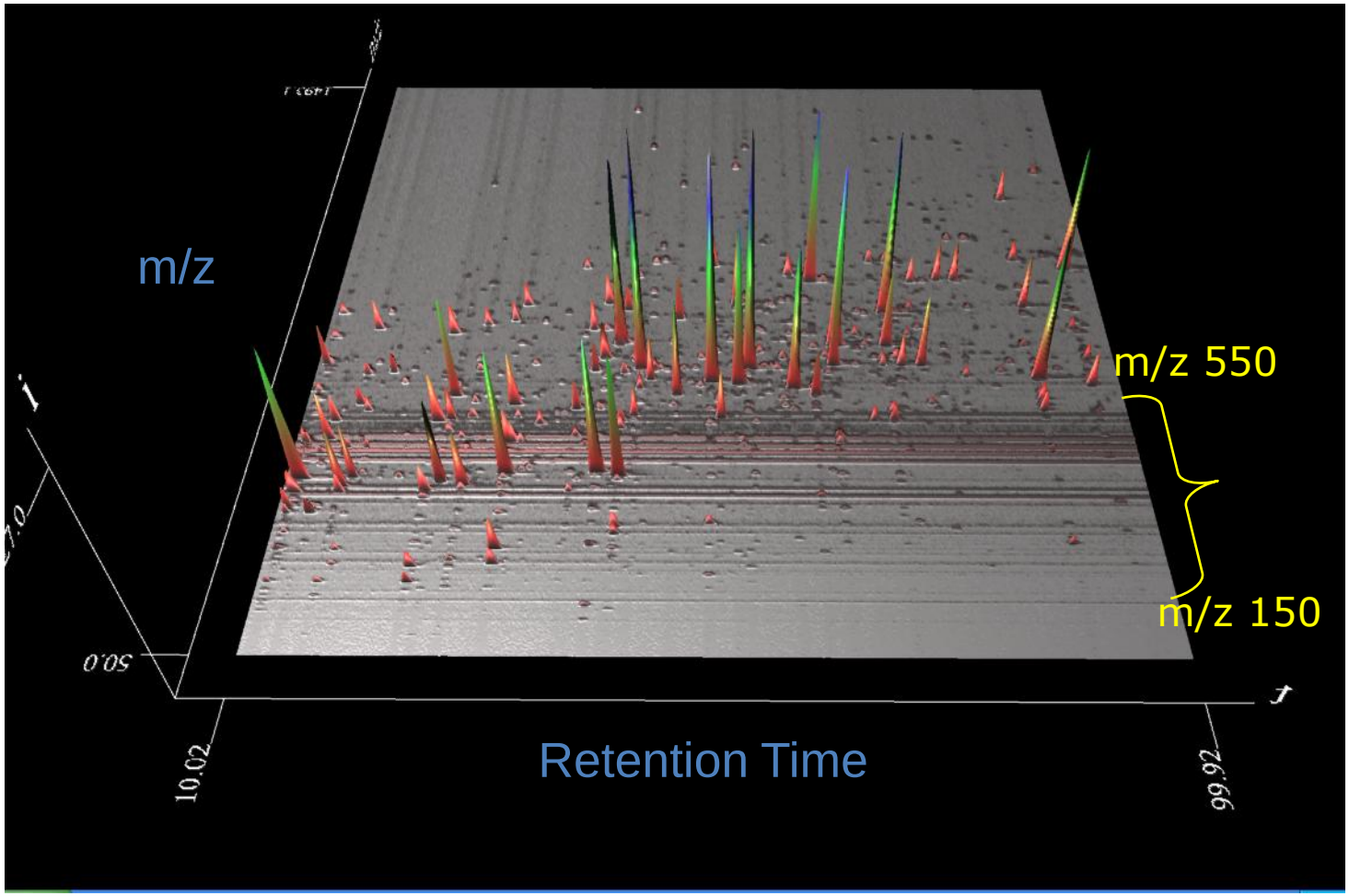
Challenges with Electrospray Ionization (ESI) Mass Detection

>850 Unique Components Detected!



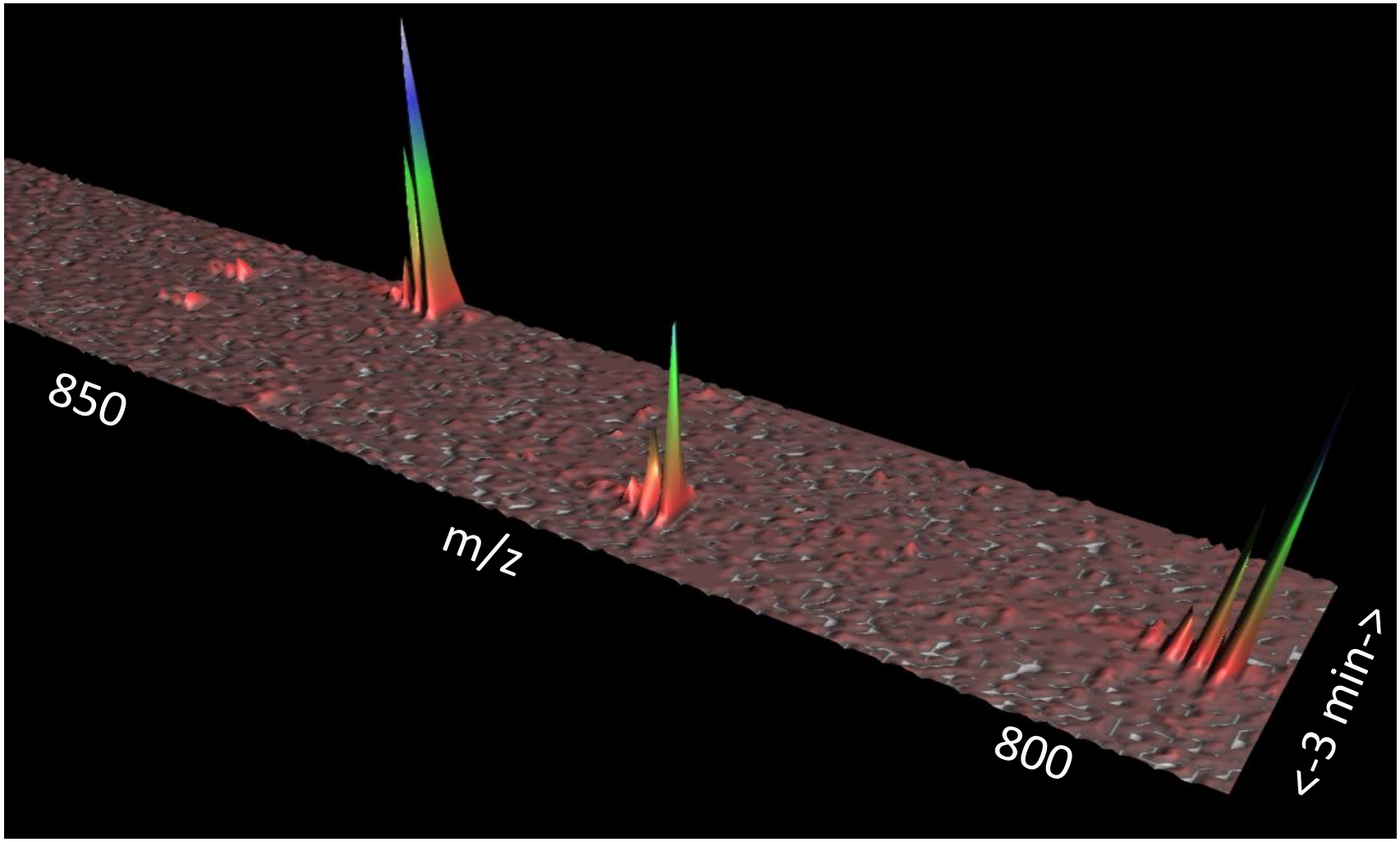
Yeast Enolase (47 kDa)

Chromatographic profiles can distinguish components from background ions

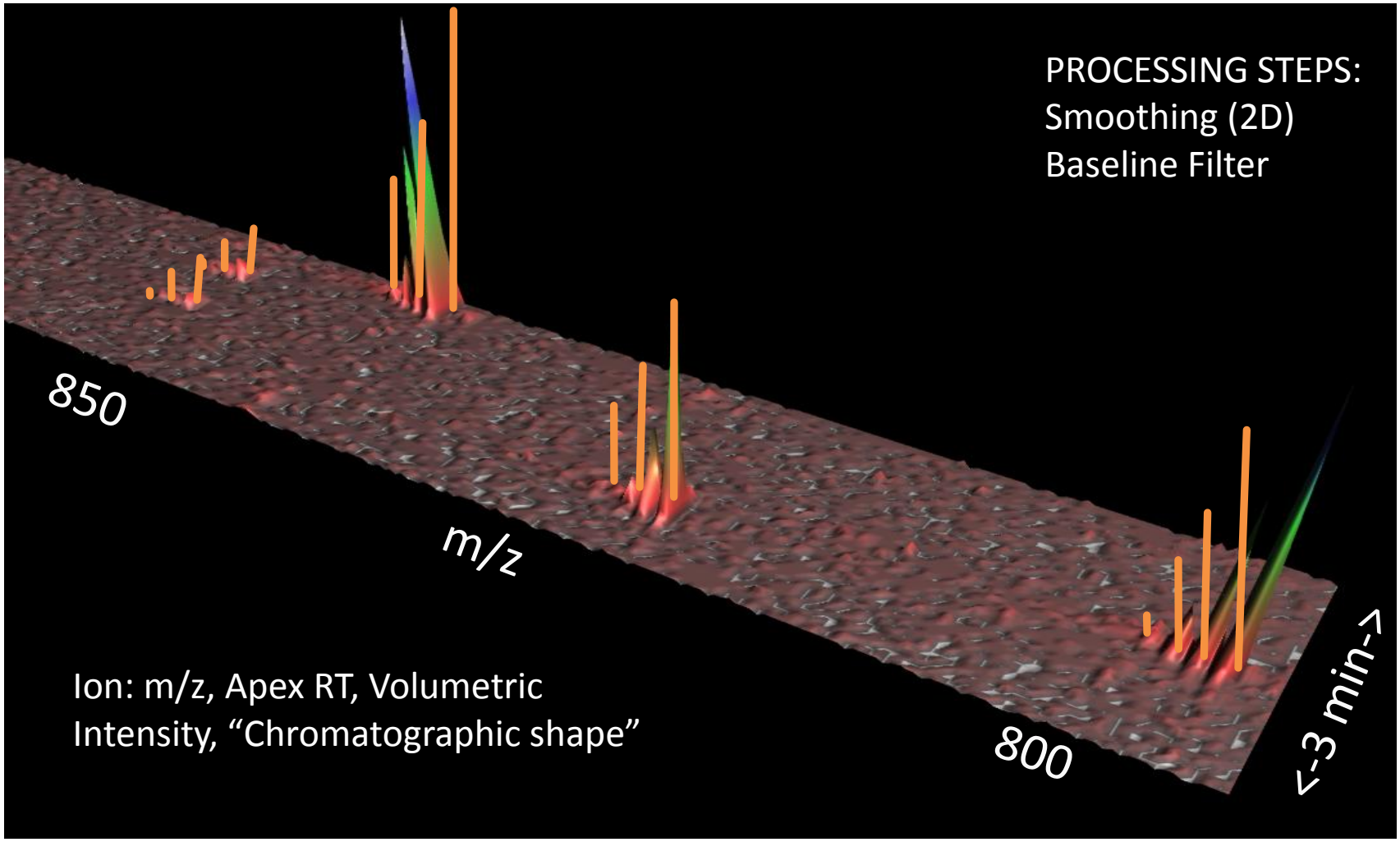


Yeast Enolase (47 kDa)

Raw LC/MS Data

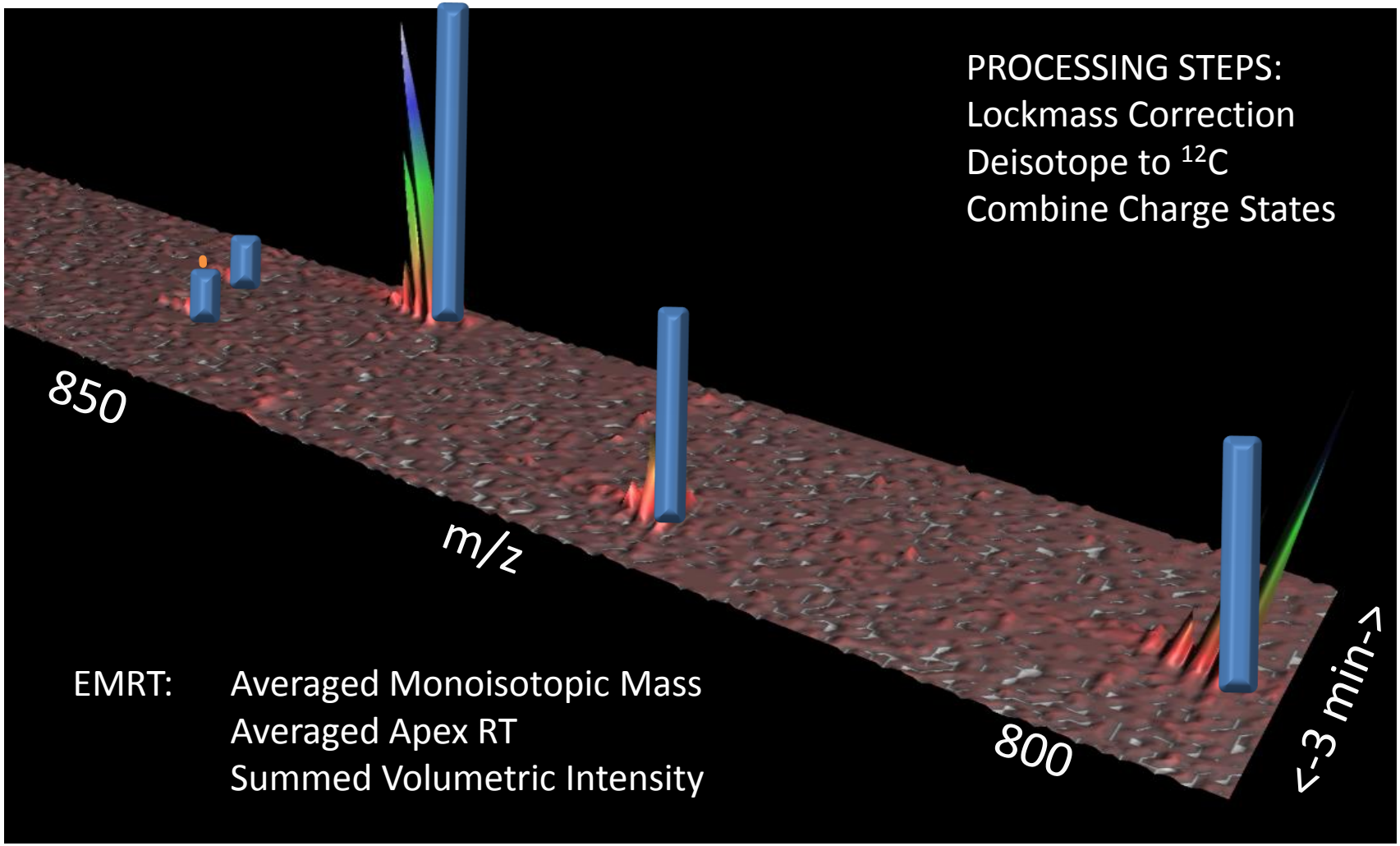


Isotope Detection



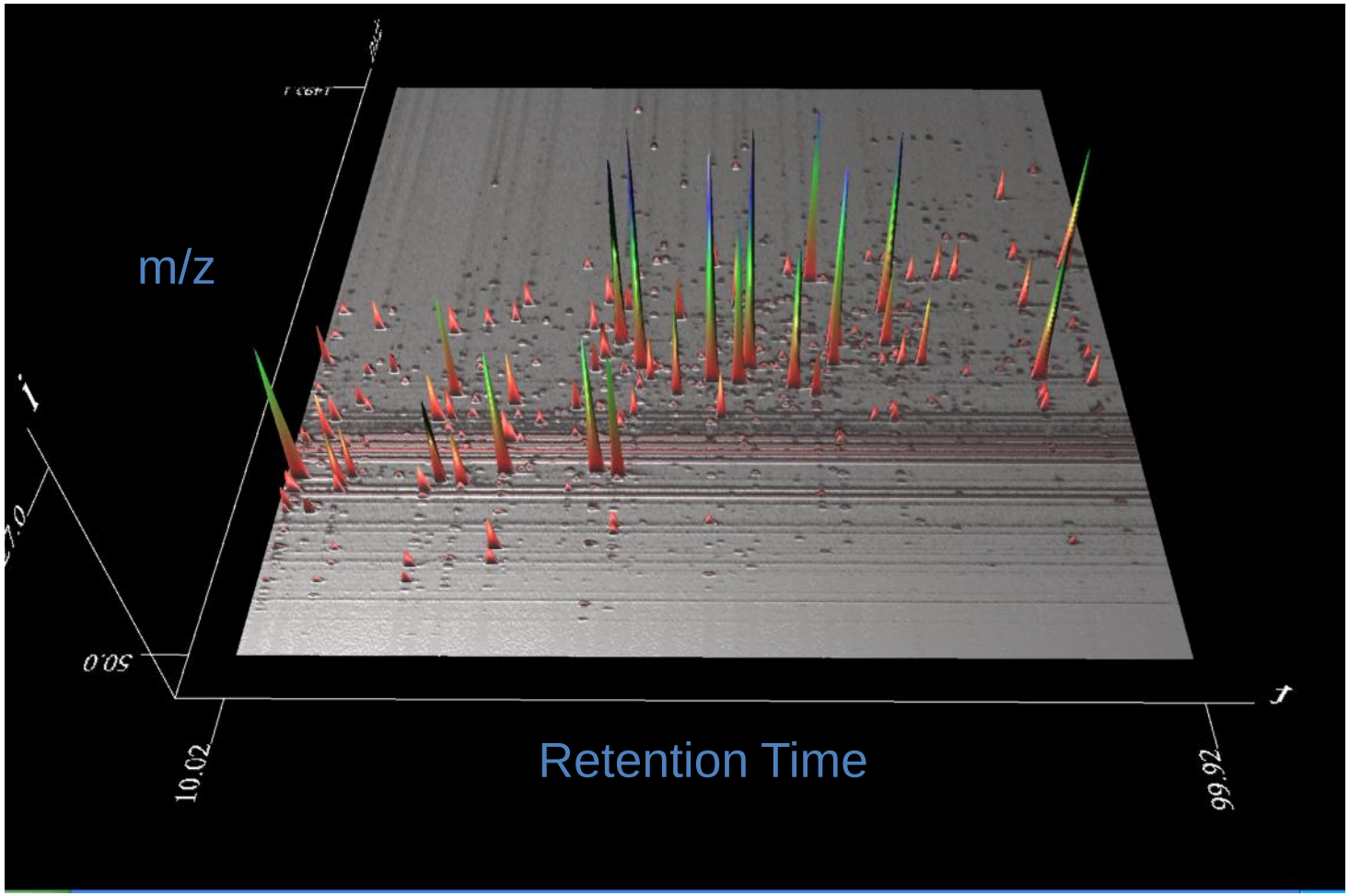
LC/MS Component (Peptide) Detection

EMRT (Exact Mass Retention Time)



LC/MS Peptide Map Complexity

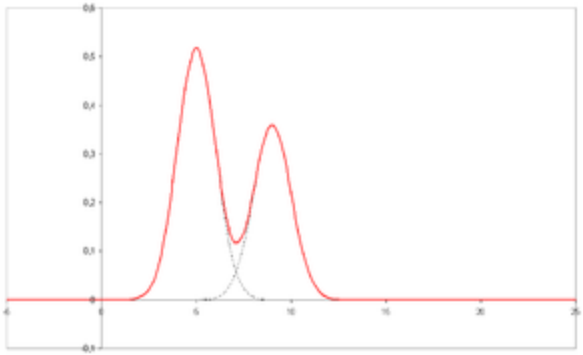
>850 Unique Components Detected!

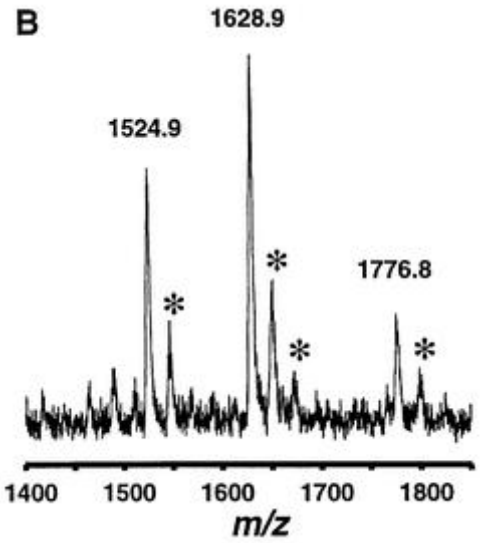


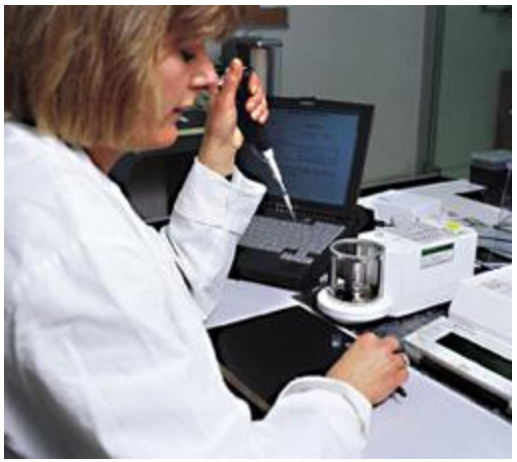
Yeast Enolase (47 kDa)

Where does the complexity of LC/MS peptide map analysis originate?

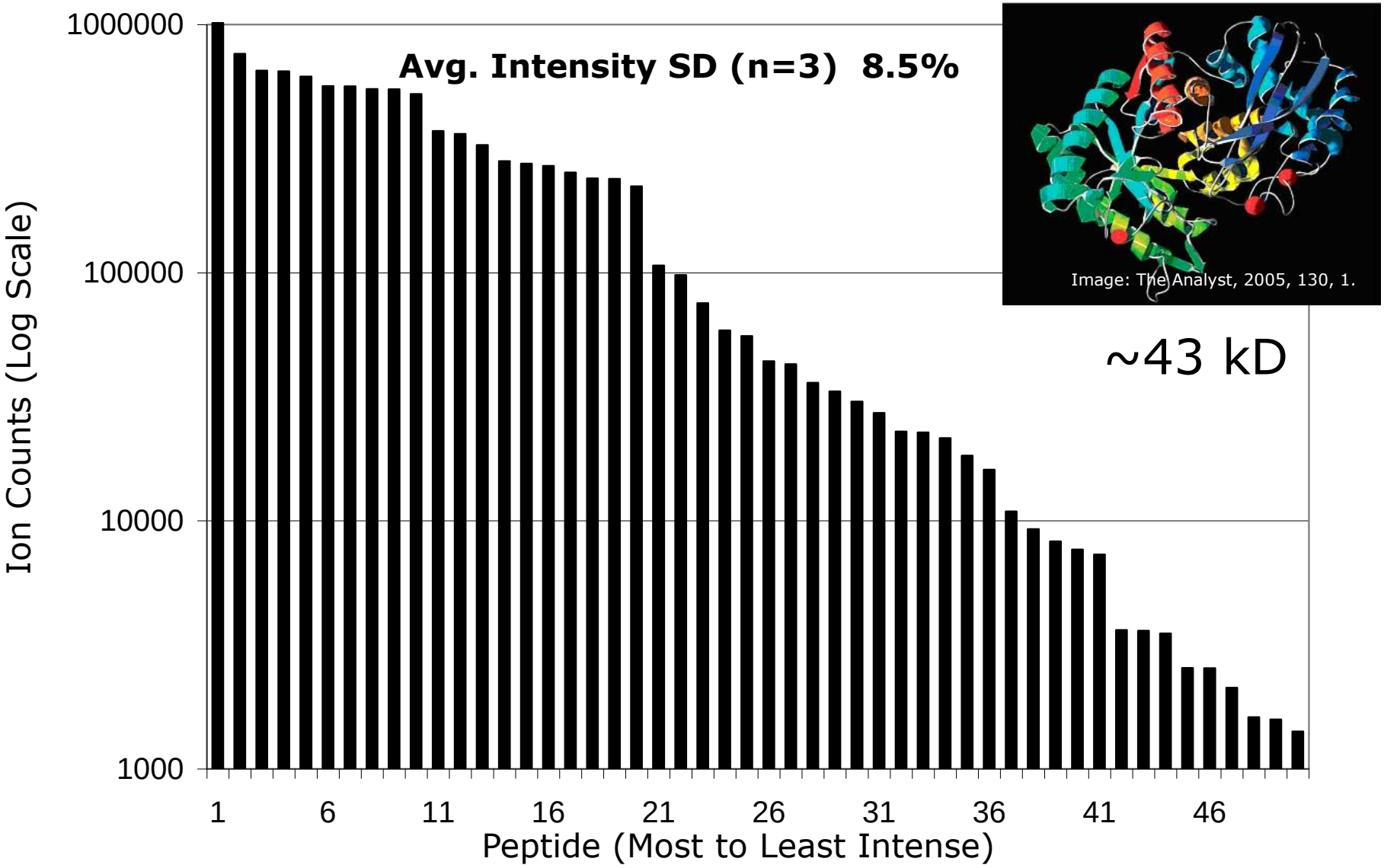
- Protein Itself
- Sample Prep Artifacts
- Chromatography Artifacts
- MS Artifacts
- Process Impurities (HCPs)



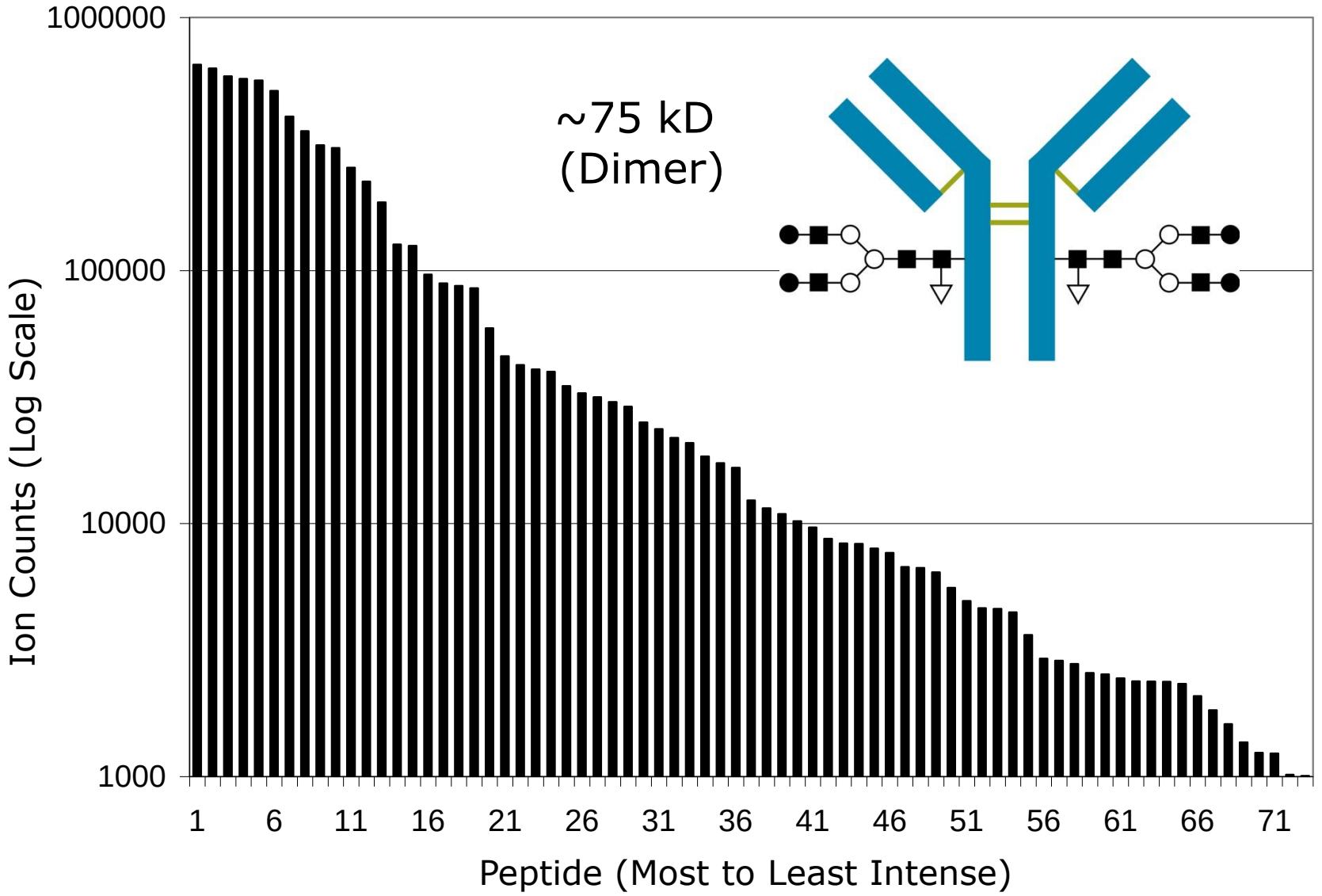




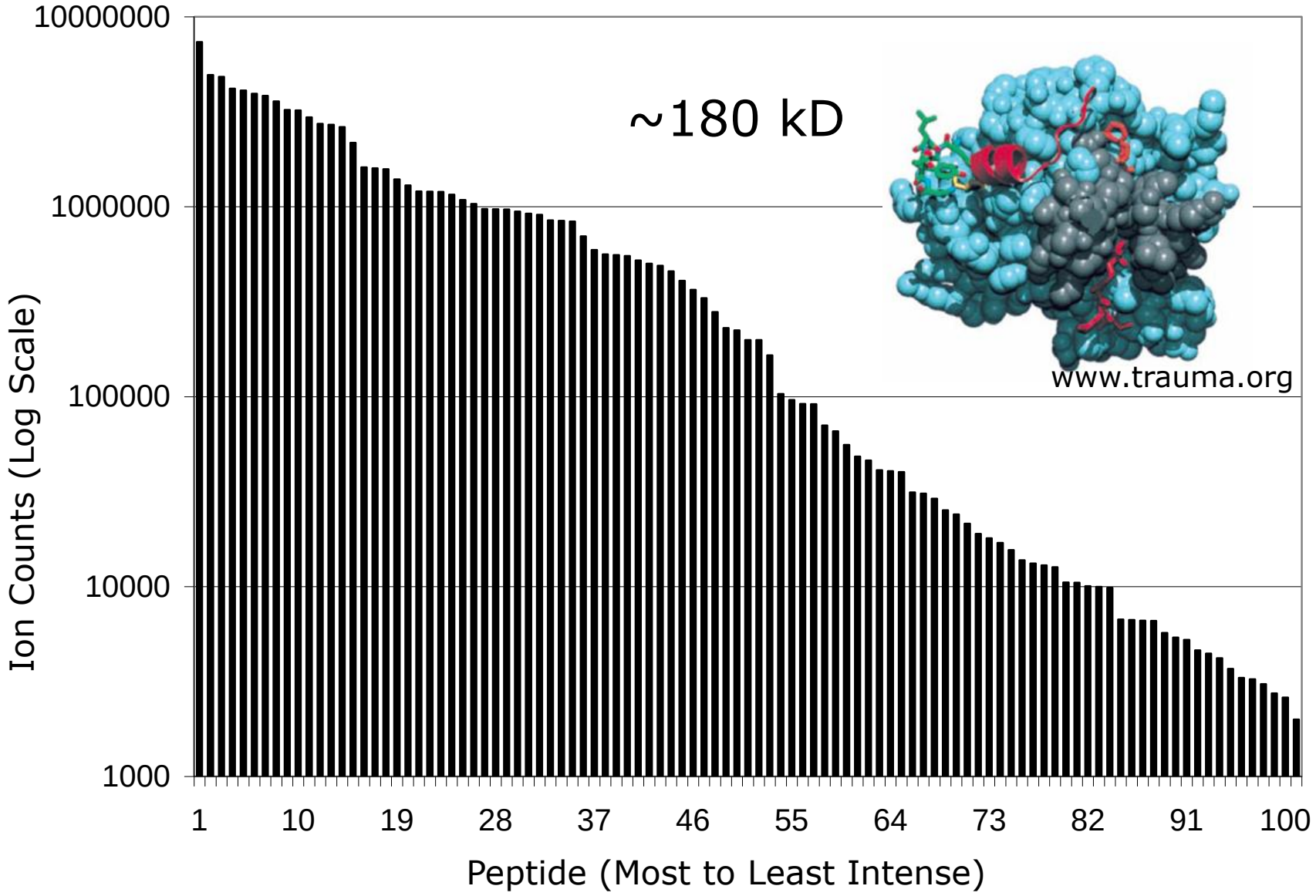
Enolase tryptic peptides are detected over ~2.5 orders of dynamic range



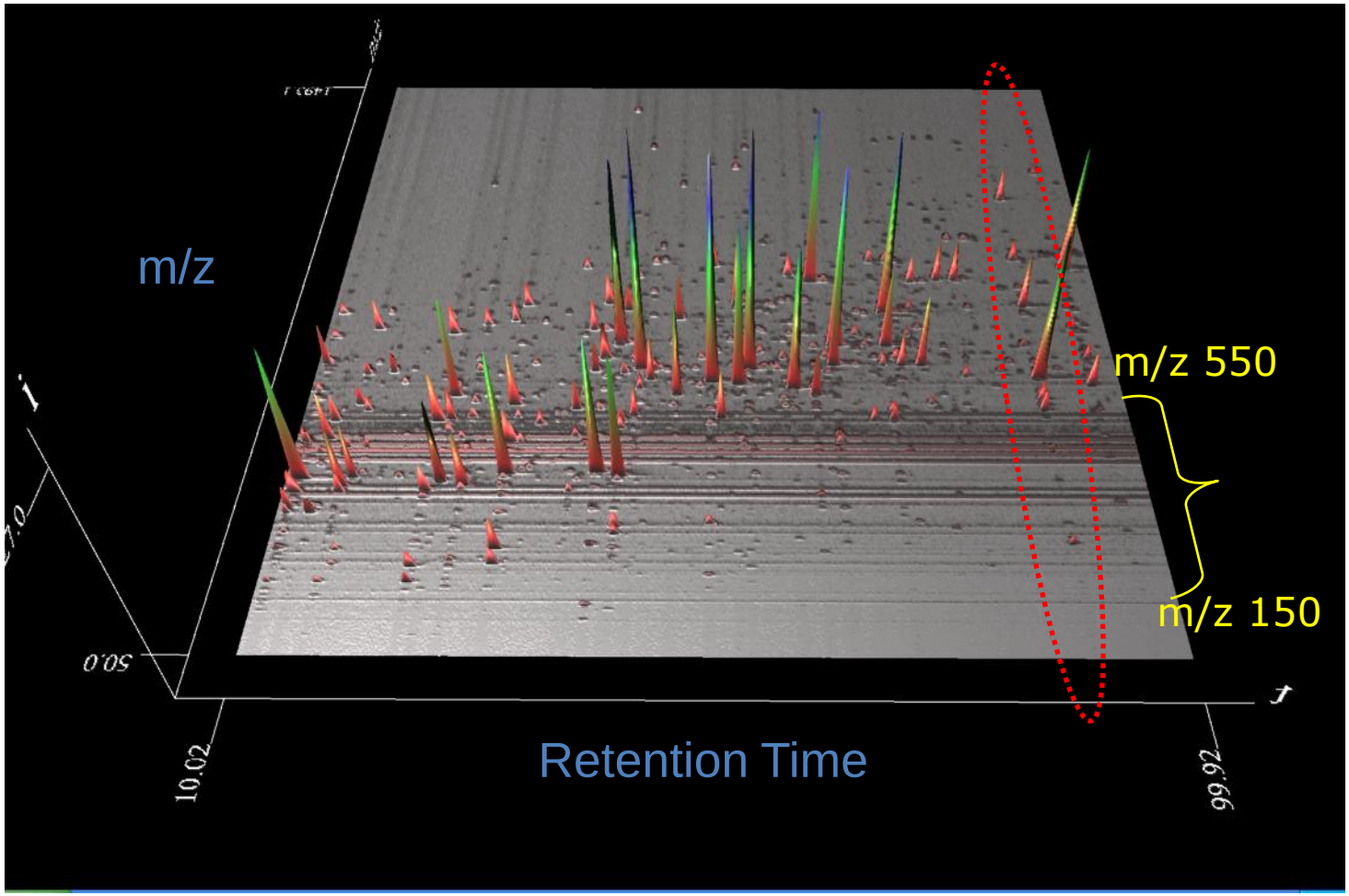
Antibody tryptic peptides are detected over ~ 3 orders of dynamic range



Recombinant Factor VIII tryptic peptides are detected over ~4 orders of dynamic range



ESI can generate many species with chromatographic alignment to the precursor peptide

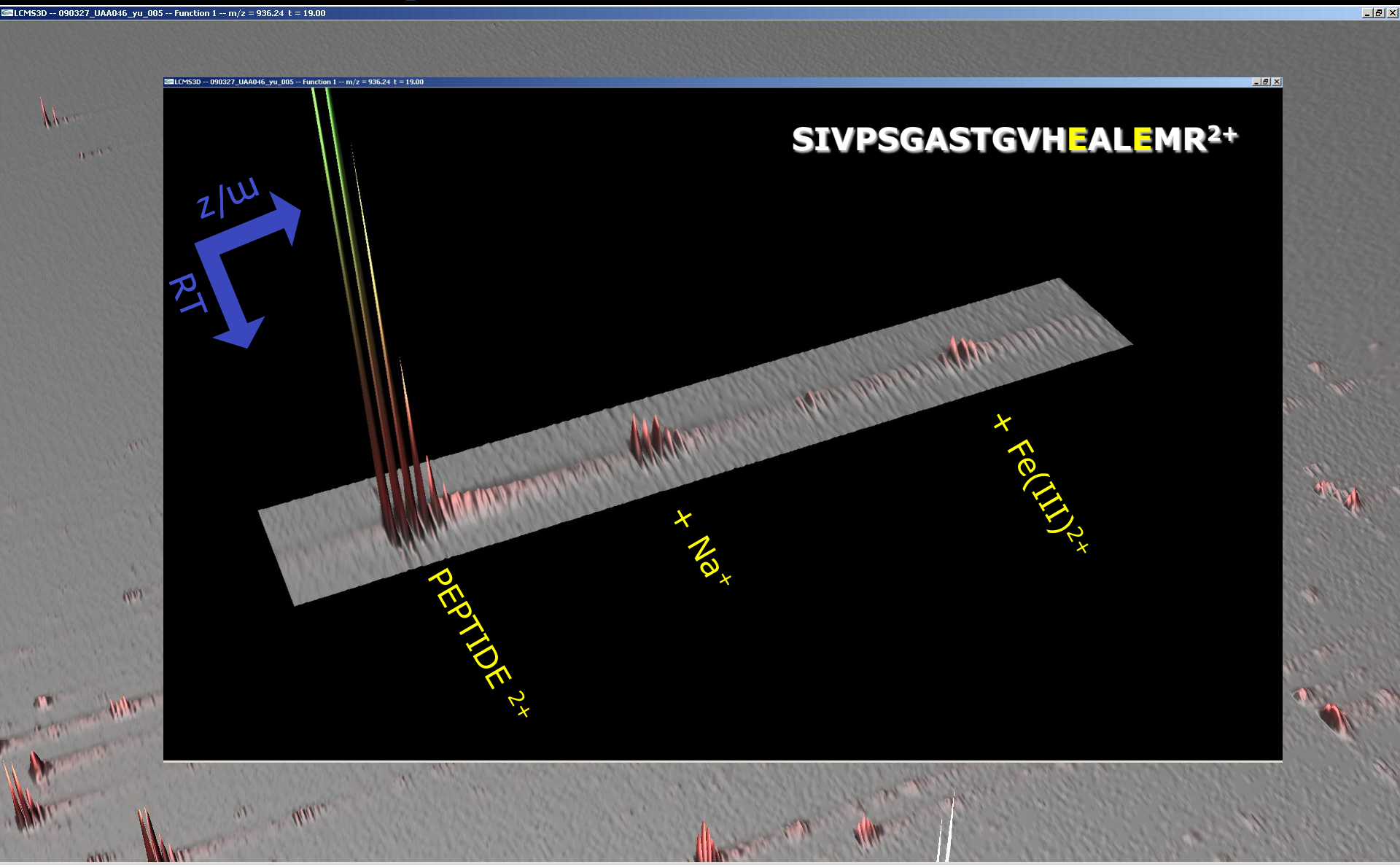


Yeast Enolase (47 kDa)

Components that demonstrate chromatographic time alignment with a detected peptide

Class of Ion	Ion Mass	Fragmentation
Adducts (e.g. Na^+)	M + Adduct	Normal
Neutral Loss (-NH_3, -H_2O)	M-17 or M-18 Da	Normal/- Loss
In-source fragment	(-) Varies	Normal + Internal
Noncovalent Dimer	M x2	Normal
In Source Oxidation	M+16 Da	Normal/+16Da

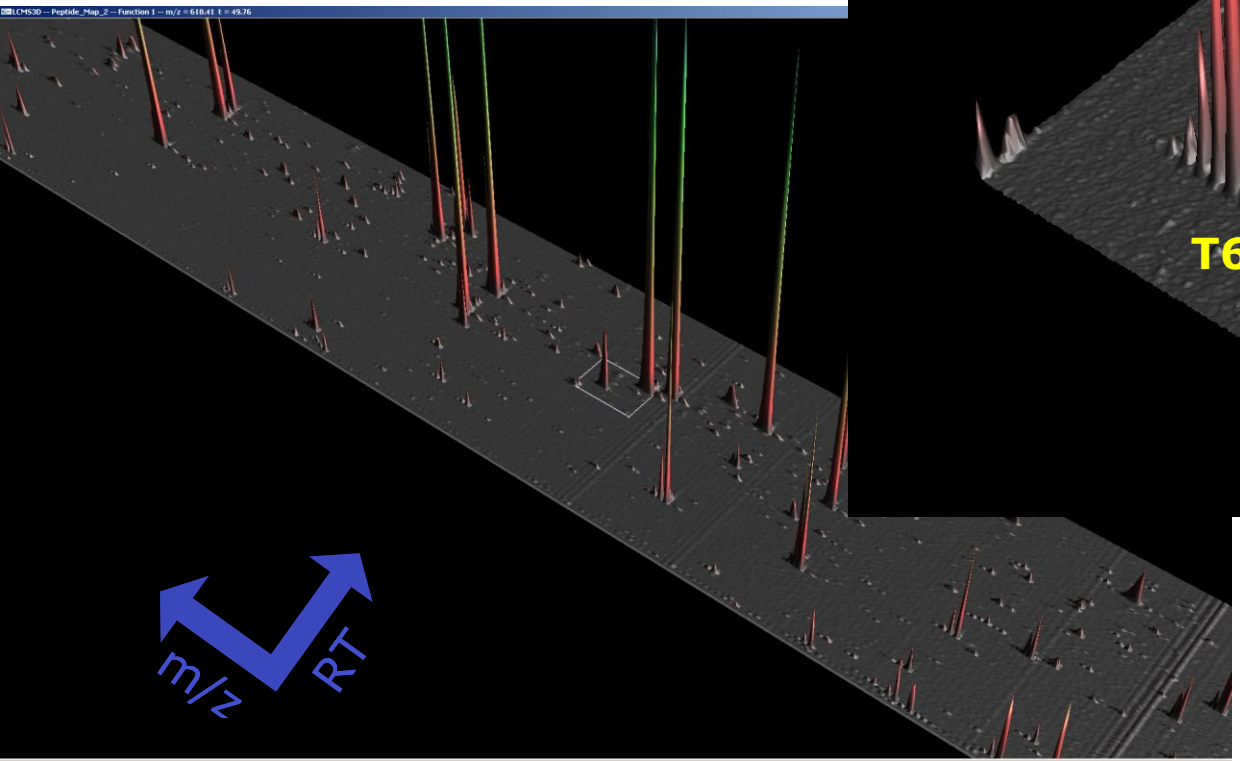
Informatics needs to be aware that these processes occur and properly assign these classes of components appropriately

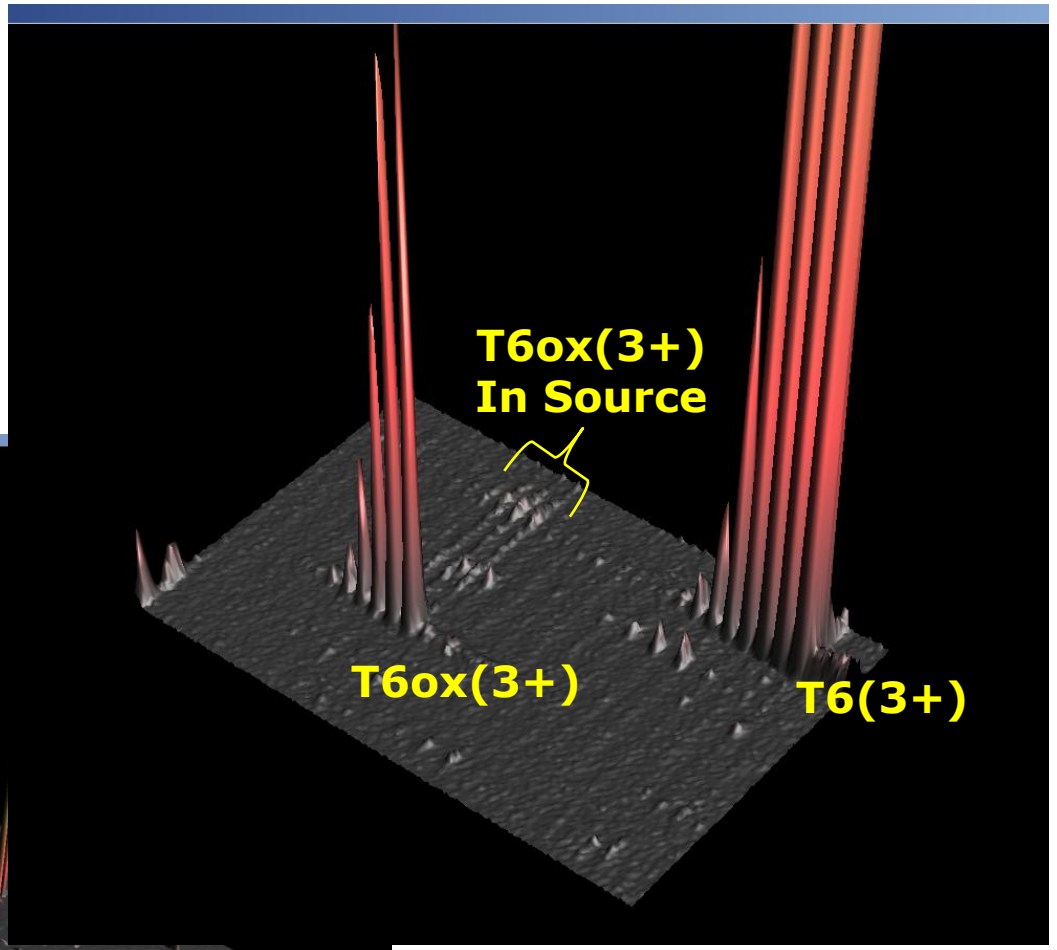

LCMS3D -- 090327_UAA046_yu_005 -- Function 1 -- m/z = 936.24 t = 19.00
SIVPSGASTGVHEALEMR2+
m/z
RT
PEPTIDE 2+
+ Na+
+ Fe(III)2+

Using chromatographic time alignment to distinguish "real" from "in-source" oxidation

Peptide	RT	m/z	Intensity
T6(3+)	54.4	614.302	1015369
T6ox(3+)	47.6	619.637	99732
T6ox(3+)	54.4	619.645	1645

SIVPSGASTGVHEALE**M**R





Map components that can be chromatographically resolved from a detected peptide

Heterogeneity of Sample:

- Modifications T1*

Complexity from Sample Preparation:

- Under-digested Peptides (Missed Cleavages, e.g. T6-7)
- Over-digested Peptides (SemiTryptic, e.g. T3n6 or T3c3)
  - Additional cleavage at non-canonical peptide bond
  - Can have the same mass as in-source fragments
- Sample Prep Artifacts (e.g. Cys Alkylation)

Reduction and Under-alkylation can result in detection of disulfide bound peptides

RT (Min)	m/z	Charge State	Mass (Da)	Intensity (Counts)
78.1	658.313	4	2629.2202	289578

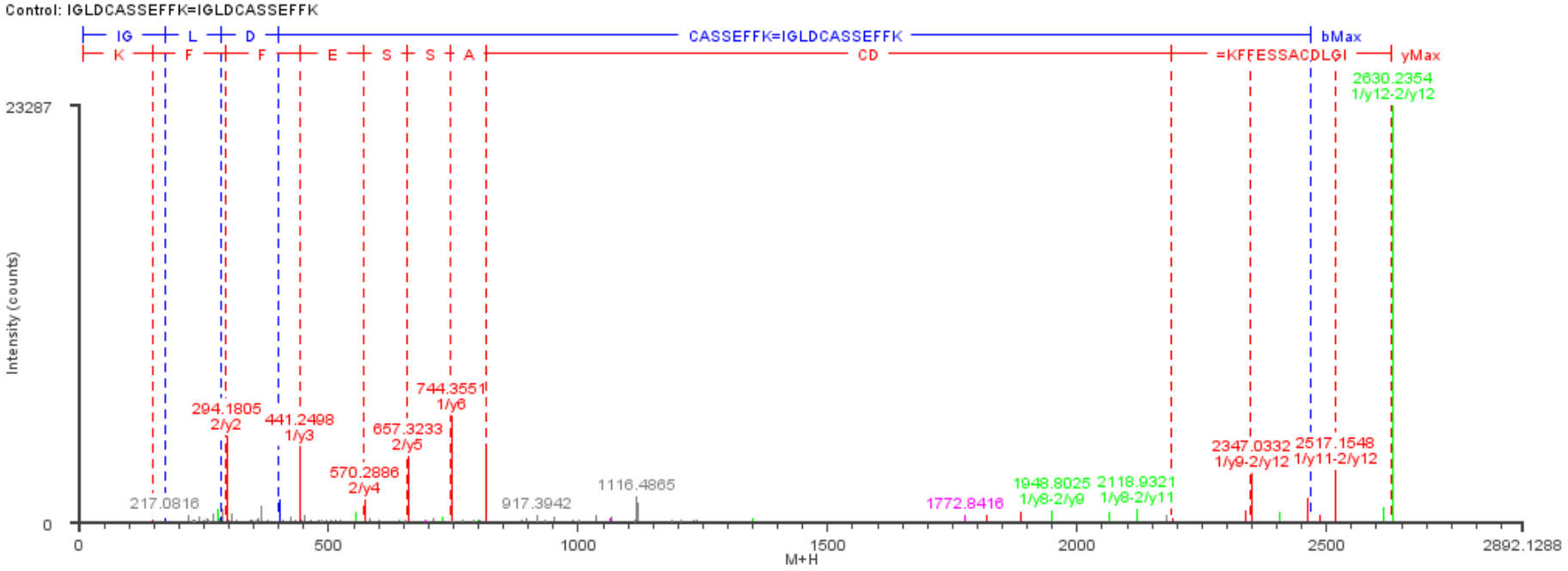
	Protein	Peptide	Fragment Number	Start	End	Modifiers	Calculated Peptide Mass (Da)	b/y Possible	Control RT (Min)	Control m/z	Control Charge State	Control Mass (Da)	Control Intensity (Cou...
1	Enolase	SIVPSGASTGVHEALEMR	1:T6	32	49		1839.9149	34	54.4	614.3133	3	1839.9161	1015170.0
2	Enolase	IEEELGDNAVFAGENFHHGDKL	1:T51-52	415	436		2440.1294	42	57.3	611.0422	4	2440.1372	917319.0
3	Enolase	TAGIQIVADDLTVTNPK	1:T38	312	328		1754.9414	32	66.0	878.4774	2	1754.9388	761587.0
4	Enolase	AVDDFLISLDGTANK	1:T14	88	102		1577.7937	28	74.4	789.9035	2	1577.7911	651682.0
5	Enolase	NWNDVIAPAFVK	1:T11	67	78		1285.7030	22	58.6	643.8572	2	1285.6985	647772.0
6	Enolase	WLTGPQLADLYHSLMK	1:T35	272	287		1871.9603	30	90.9	624.9948	3	1871.9604	617700.0
7	Enolase	AAQDSFAAGWGVMVSHR	1:T44	358	374		1788.8365	32	63.5	597.2865	3	1788.8357	565335.0
8	Enolase	VNQIGTLSESIK	1:T43	346	357		1287.7034	22	46.5	644.8584	2	1287.7009	563821.0
9	Enolase	LGANAILGVSLAASR	1:T16	105	119		1411.8147	28	68.4	706.9136	2	1411.8113	550375.0
10	Enolase	SGETEDTFIADLVVGLR	1:T45	375	391		1820.9155	32	96.8	911.4658	2	1820.9156	549213.0
11	Enolase	GNPTVEVELTTEK	1:T4	15	27		1415.7144	24	44.0	708.8654	2	1415.7148	525508.0
12	Enolase	IEEELGDNAVFAGENFHHGDK	1:T51	415	435		2327.0454	40	50.7	582.7725	4	2327.0583	372181.0
13	Enolase	YGASAGNVGDEGGVAPNIQT...	1:T27	201	233		3256.6099	64	108.4	1086.5535	3	3256.6365	362762.0
14	Enolase	IGSEVYHNLK	1:T23	185	194		1158.6033	18	26.2	580.3069	2	1158.5980	326725.0
15									78.1	658.3130	4	2629.2202	289578.0
16	Enolase	IGLDCASSEFFK	1:T30	243	254		1315.6118	22	68.1	658.8119	2	1315.6079	281786.0
17	Enolase	AADALLLK	1:T42	338	345		813.4960	14	42.4	407.7532	2	813.4906	275217.0
18	Enolase	YPIVSIEDPFAEDDWEAWSH...	1:T37	289	311		2827.2805	44	107.1	943.4408	3	2827.2986	269466.0
19	Enolase	HLADLSK	1:T19	132	138		782.4286	12	11.4	392.2197	2	782.4236	253431.0
20	Enolase	TFAEALR	1:T22	178	184		806.4286	12	32.8	404.2184	2	806.4209	239676.0
21	Enolase	TSPYVLPVPFLNVLNGGSHAG...	1:T21	141	177		3736.9651	72	108.3	1246.6765	3	3737.0056	239031.0
22	Enolase	YDLDFK	1:T32	258	263		799.3752	10	44.7	400.6925	2	799.3691	222965.0
23	Enolase	WLTGPQLADLYH	1:T35c4	272	283		1412.7089	22	74.3	707.3617	2	1412.7075	181008.0

Reduction and Under-alkylation can result in detection of disulfide bound peptides

RT (Min)	m/z	Charge State	Mass (Da)	Intensity (Counts)
78.1	658.313	4	2629.2202	289578

IGLDCASSEFFK=IGLDCASSEFFK

MS^E Fragmentation



Cys alkylation artifacts can increase peptide map complexity

Carbamidomethylation on **Tyr (Y)**	Peptide + 57.0215 Da	*Anal. Biochem*. **54**, 170-177 (1973)
Carbamidomethylation on **Met (M)** +Decomposition (In Source Fragment)	Peptide + 57.0215 Da Peptide - 48.0034 Da	*J. Mass Spectrom*. **35**, 572–575 (2000)

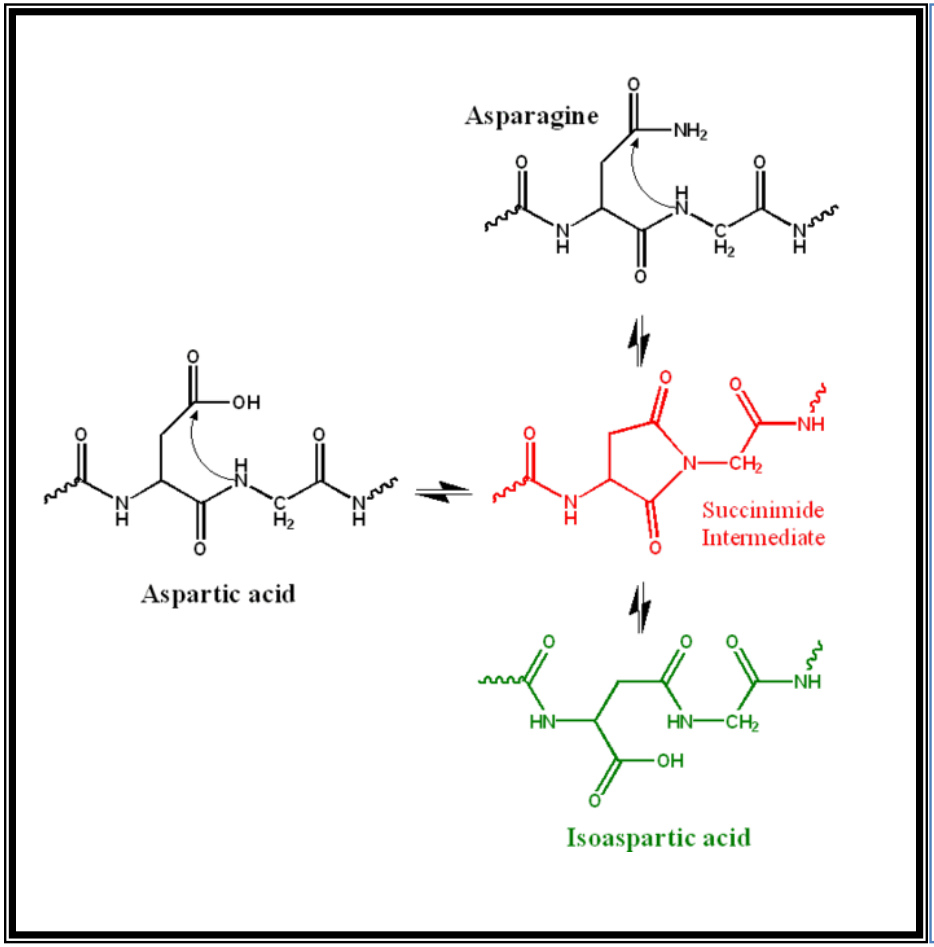
Reduction(DTT)/Alkylation(IA) + DTT Quench (7 Artifact Peptides)
Reduction(DTT)/Alkylation(IA) – No DTT Quench (30 Artifact Peptides)

Chromatogram | Spectrum | Coverage Map | Protein Digests | Peak Match Data

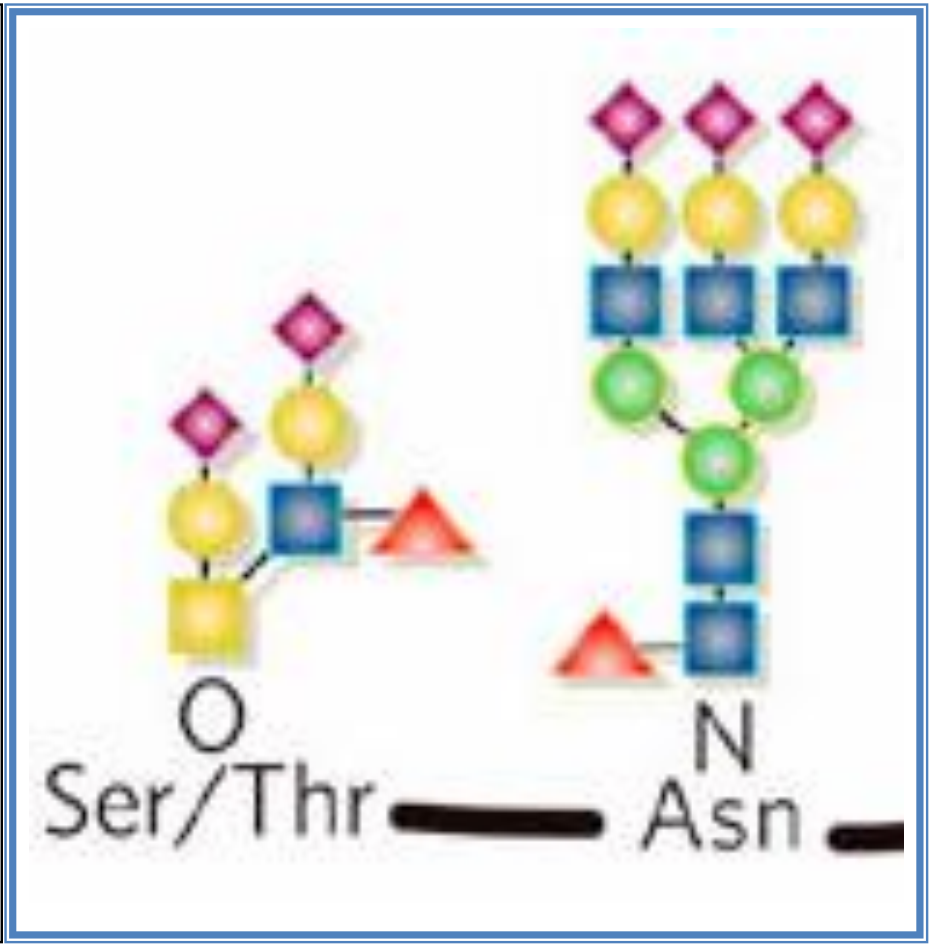
	Protein	Peptide	Fragment Number	Modifiers
1	IgG1_humanized	LLIYSASFLYSGVPSR	1:T5*	carbamidomethyl Y(1)
2	IgG1_humanized	DIQMTQSPSSLSASVGDR	1:T1*	Carbamidomethyl M(1)
3	IgG1_humanized	WGGDGFYAMDYWGQGTLVTVSSASTK	2:T12*	carbamidomethyl Y(1),Carbamidomethyl M(1)
4	IgG1_humanized	DTLMISR	2:T21*	Carbamidomethyl M(1)
5	IgG1_humanized	YADSVK	2:T7*	carbamidomethyl Y(1)
6	IgG1_humanized	TTPPVLDSDGSFFLYSK	2:T38*	carbamidomethyl Y(1)
7	IgG1_humanized	NTAYLQMNSLR	2:T10*	carbamidomethyl Y(1),Carbamidomethyl M(1)
8	IgG1_humanized	WGGDGFYAMDYWGQGTLVTVSSASTK	2:T12*	decom of carbamidomethyl M(1)
9	IgG1_humanized	WGGDGFYAMDYWGQGTLVTVSSASTK	2:T12*	Carbamidomethyl M(1)
10	IgG1_humanized	DSTYSLSSTLTLSK	1:T15*	carbamidomethyl Y(1)
11	IgG1_humanized	DIQMTQSPSSLSASVGDR	1:T1*	Carbamidomethyl M(1)
12	IgG1_humanized	DIQMTQSPSSLSASVGDR	1:T1*	decom of carbamidomethyl M(1)
13	IgG1_humanized	WGGDGFYAMDYWGQGTLVTVSSASTK	2:T12*	decom of carbamidomethyl M(1)
14	IgG1_humanized	FNWYVDGVEVHNAK	2:T23*	carbamidomethyl Y(1)
15	IgG1_humanized	WGGDGFYAMDYWGQGTLVTVSSASTK	2:T12*	Carbamidomethyl M(1)
16	IgG1_humanized	DSTYSLSSTLTLSK	1:T15*	carbamidomethyl Y(1)
17	IgG1_humanized	DIQMTQSPSSLSASVGDR	1:T1*	Carbamidomethyl M(1)
18	IgG1_humanized	TTPPVLDSDGSFFLYSK	2:T38*	carbamidomethyl Y(1)
19	IgG1_humanized	DTLMISR	2:T21*	decom of carbamidomethyl M(1)
20	IgG1_humanized	ASQDVNTAVAWYQQKPGK	1:T3*	carbamidomethyl Y(1)
21	IgG1_humanized	DTLMISR	2:T21*	Carbamidomethyl M(1)
22	IgG1_humanized	DTYIHWVR	2:T3*	carbamidomethyl Y(1)
23	IgG1_humanized	ASQDVNTAVAWYQQKPGK	1:T3*	carbamidomethyl Y(1)

Two modification types where separation is key to producing proper detection and assignment

Deamidation

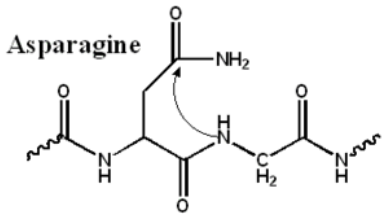


Glycopeptides

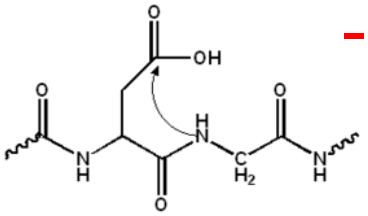


Deamidation: At neutral or alkaline pH Asn can transform to Asp, isoAsp, Asn_{suc}

N

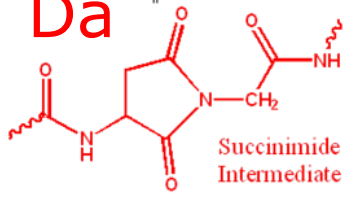


D



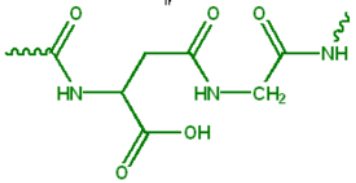
Aspartic acid

-17 Da



Succinimide Intermediate

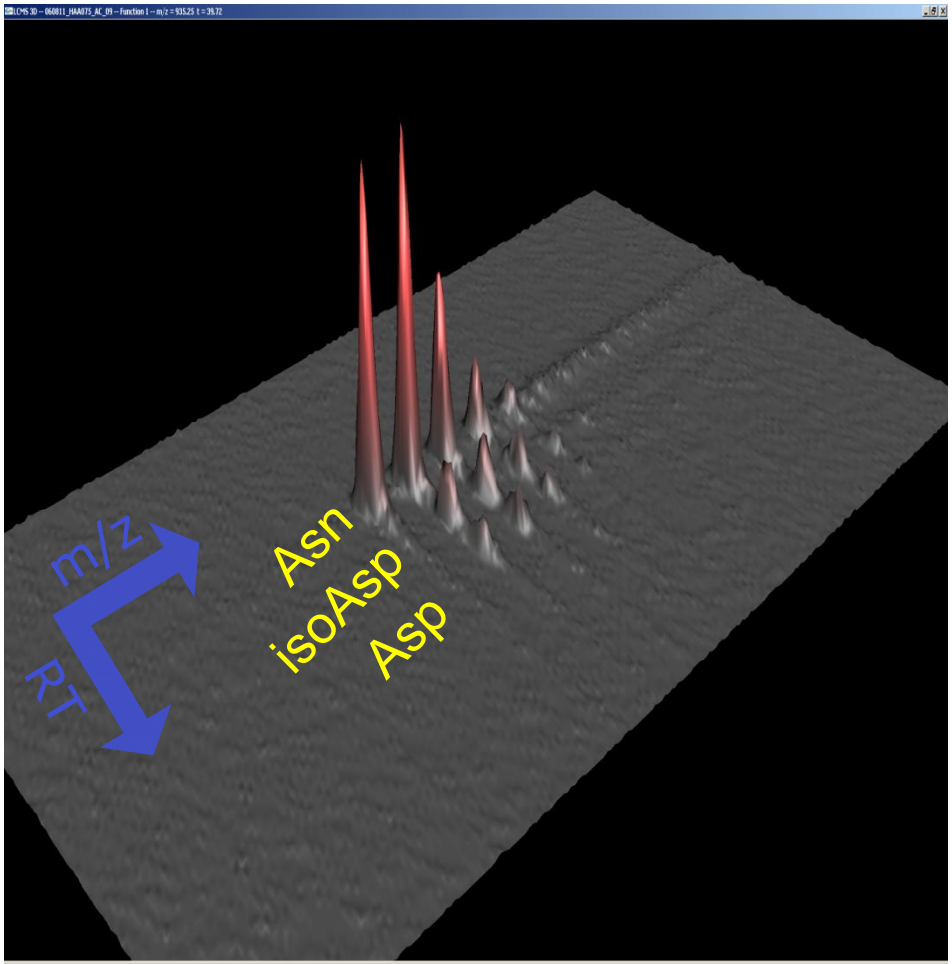
+ 1 Da



Isoaspartic acid

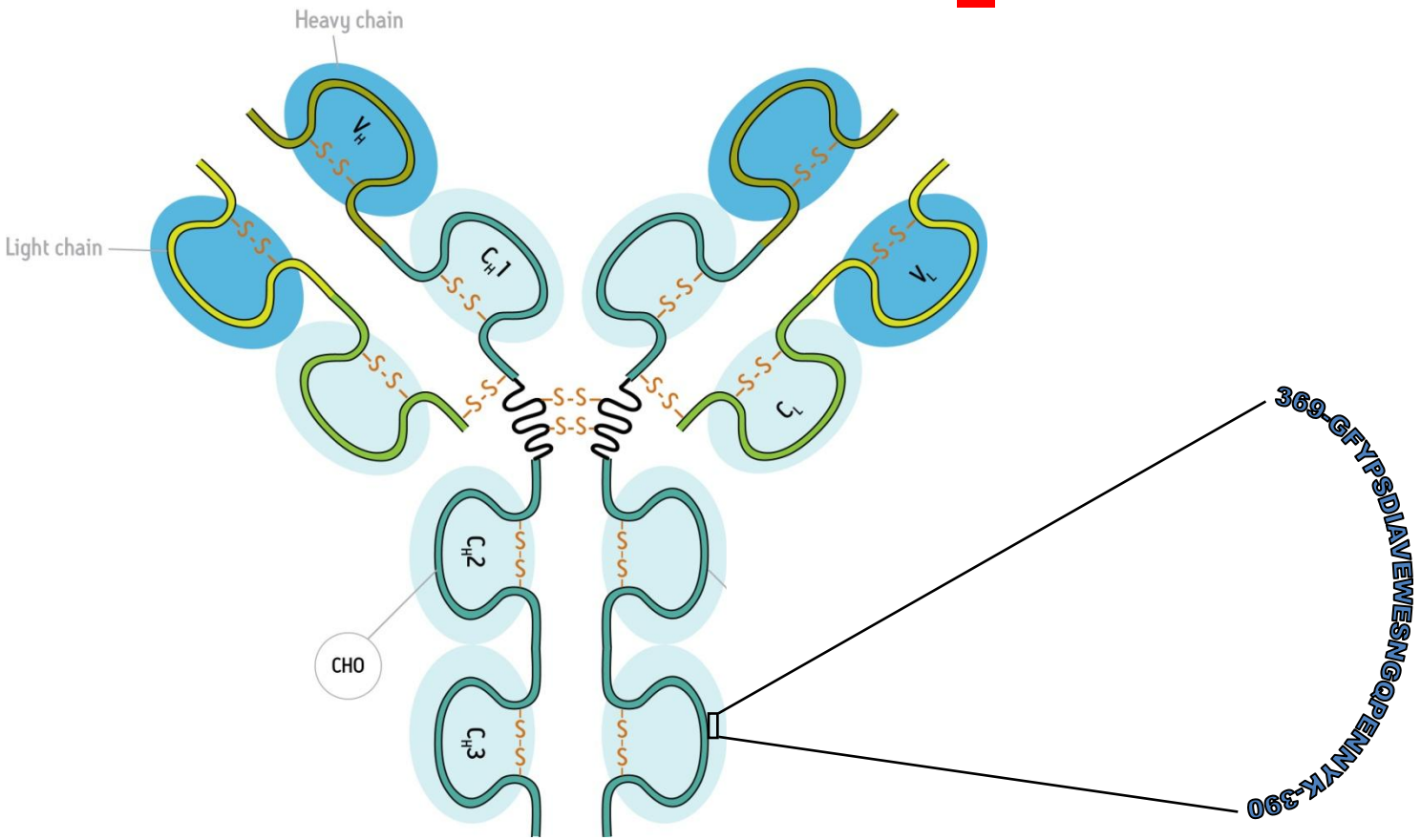
+ 1 Da

isoD



Deamidation of the PENNY Peptide

369-GFYPSDIAVEWESNGQPENNYK-390



Separation of PENNY Peptide and its deamidated forms

1 - GFYPSDIAVEWESNGQPENNYK

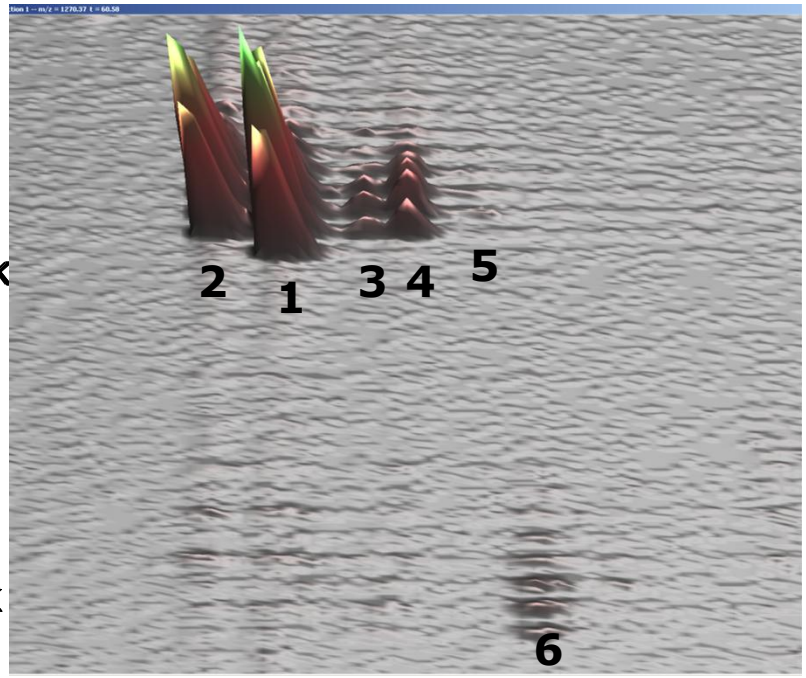
2 - GFYPSDIAVEWESisoDGQPENNYK

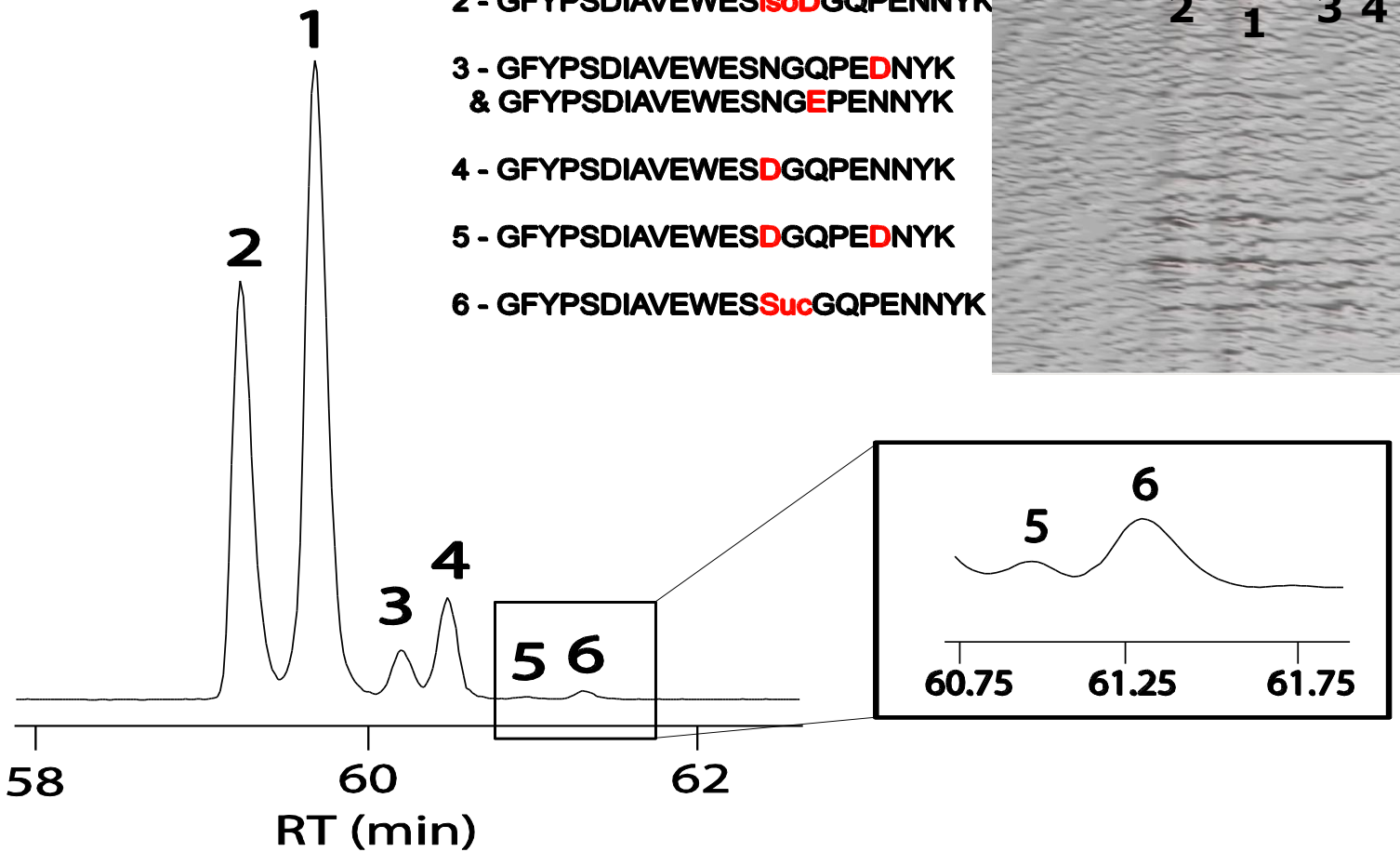
3 - GFYPSDIAVEWESNGQPEDNYK
& GFYPSDIAVEWESNGEPENNYK

4 - GFYPSDIAVEWESDGQPENNYK

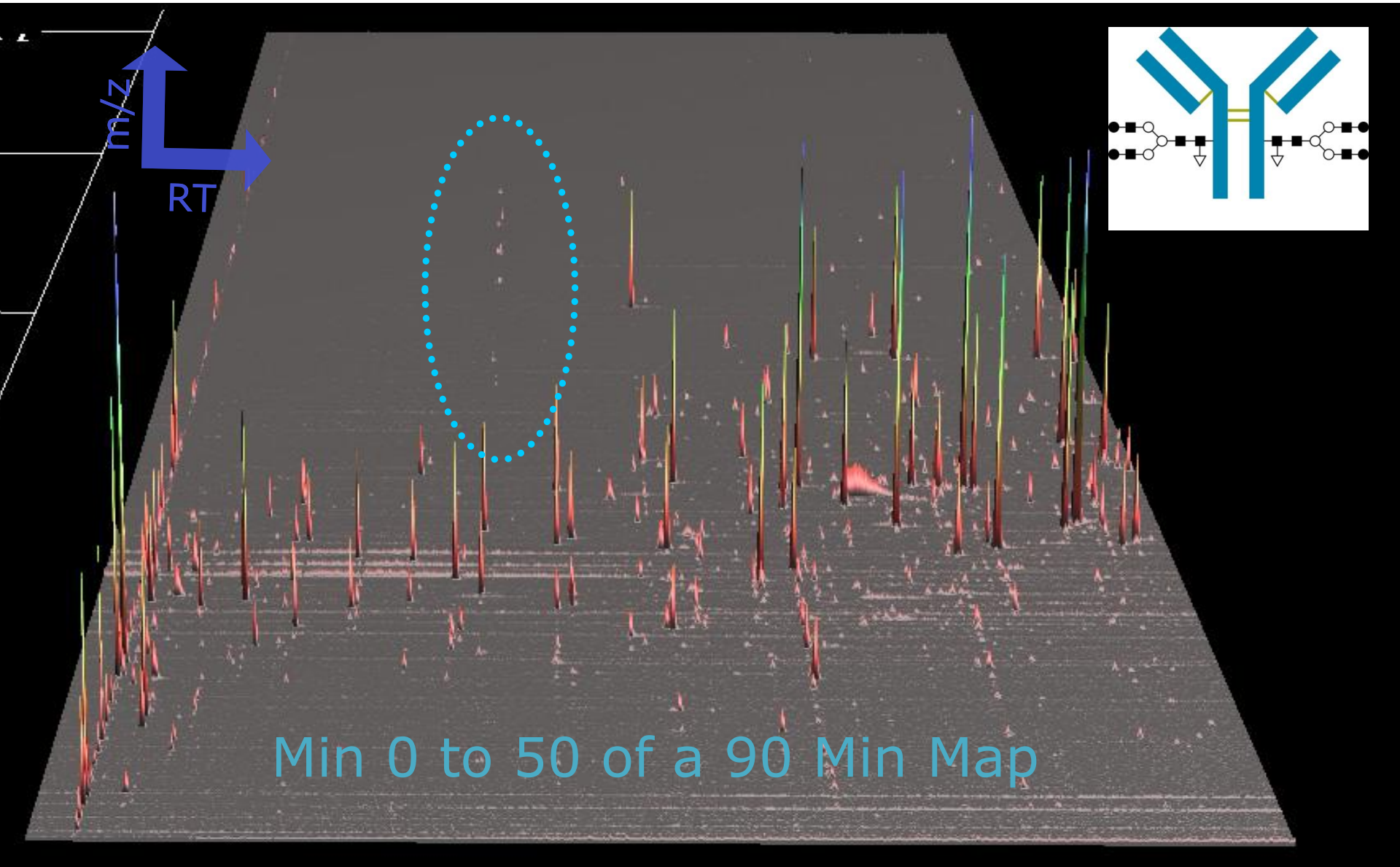
5 - GFYPSDIAVEWESDGQPEDNYK

6 - GFYPSDIAVEWESSucGQPENNYK

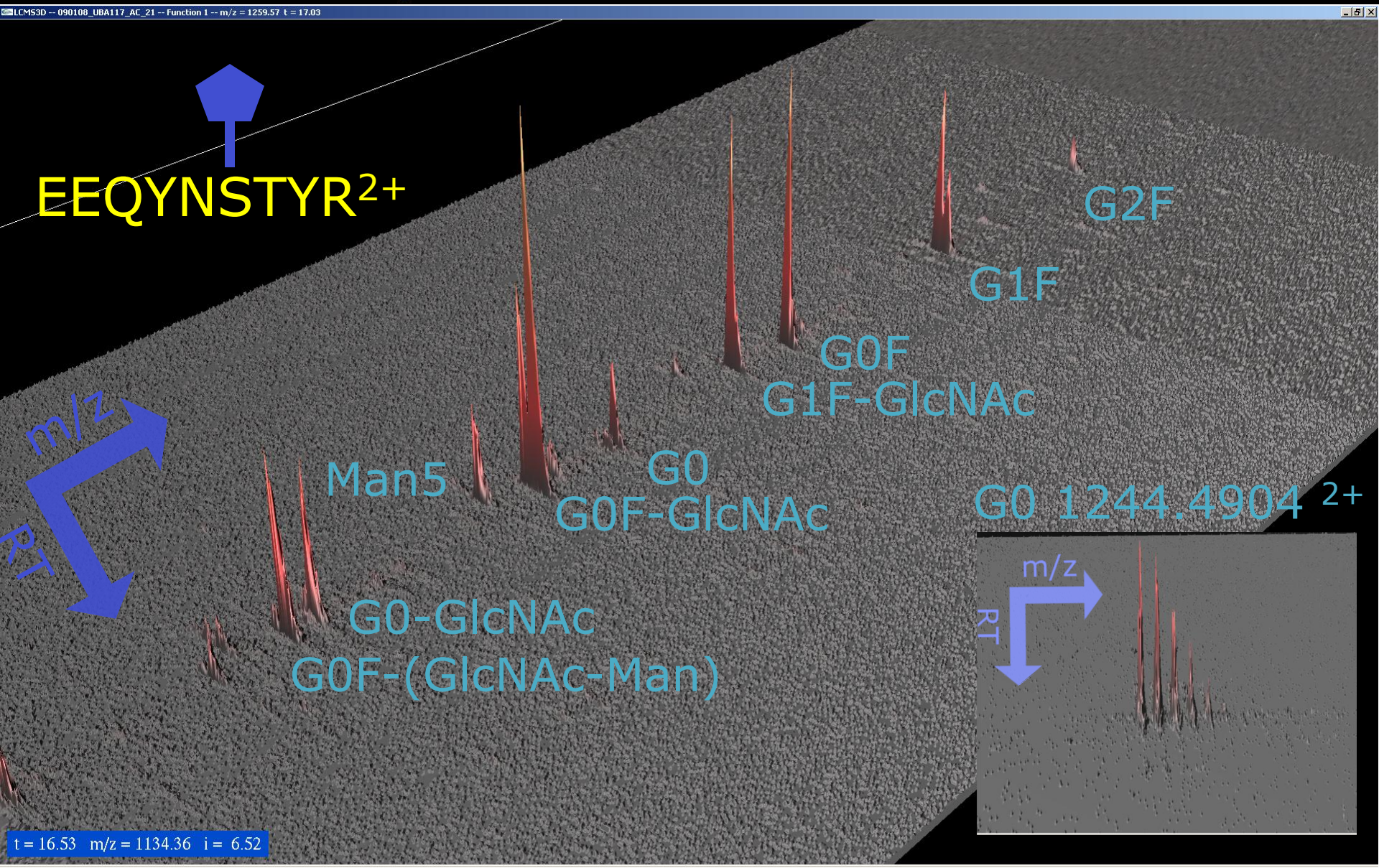




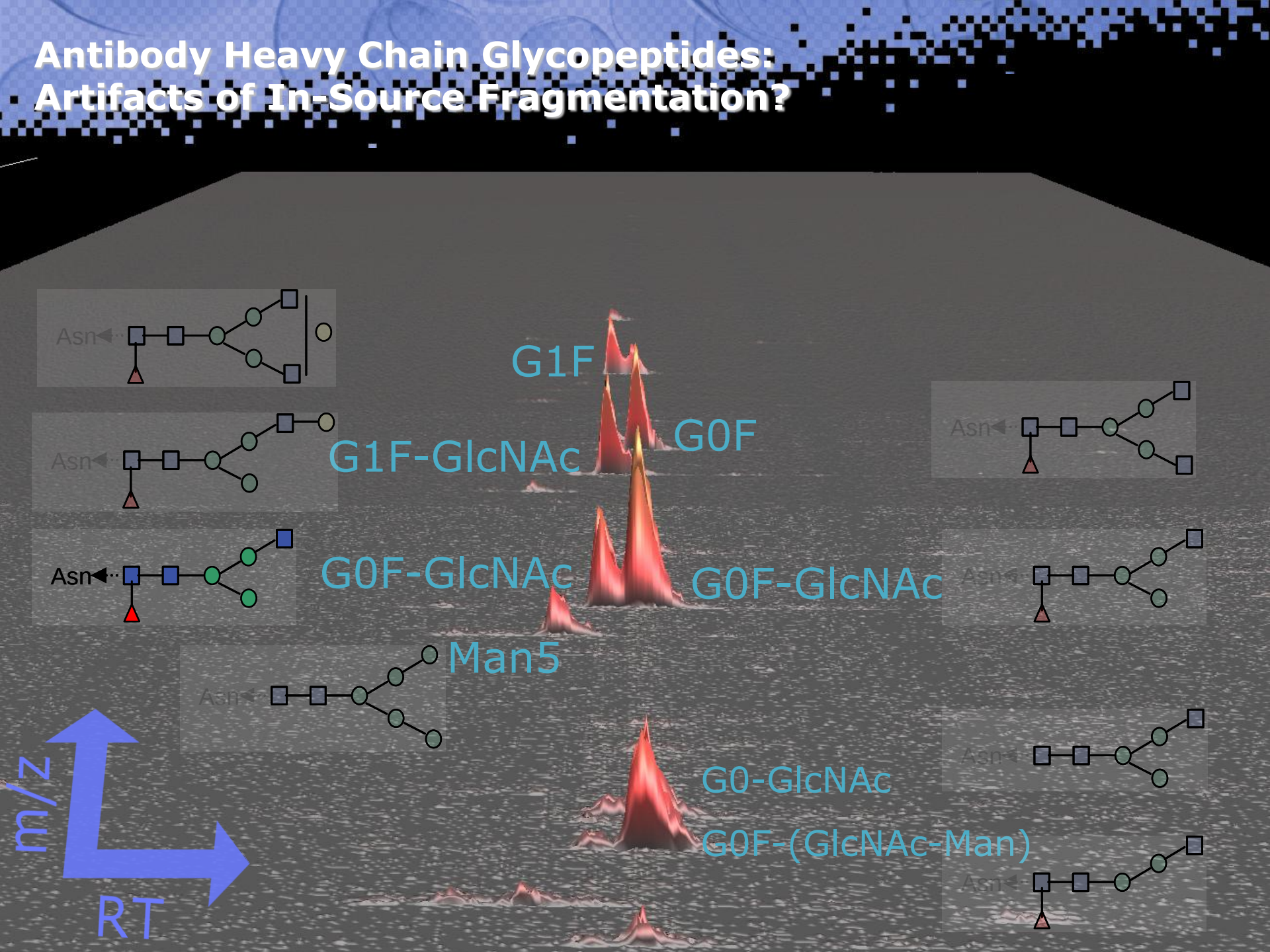
Antibody heavy chain glycopeptides demonstrate poor ionization



Antibody Heavy Chain Glycopeptides



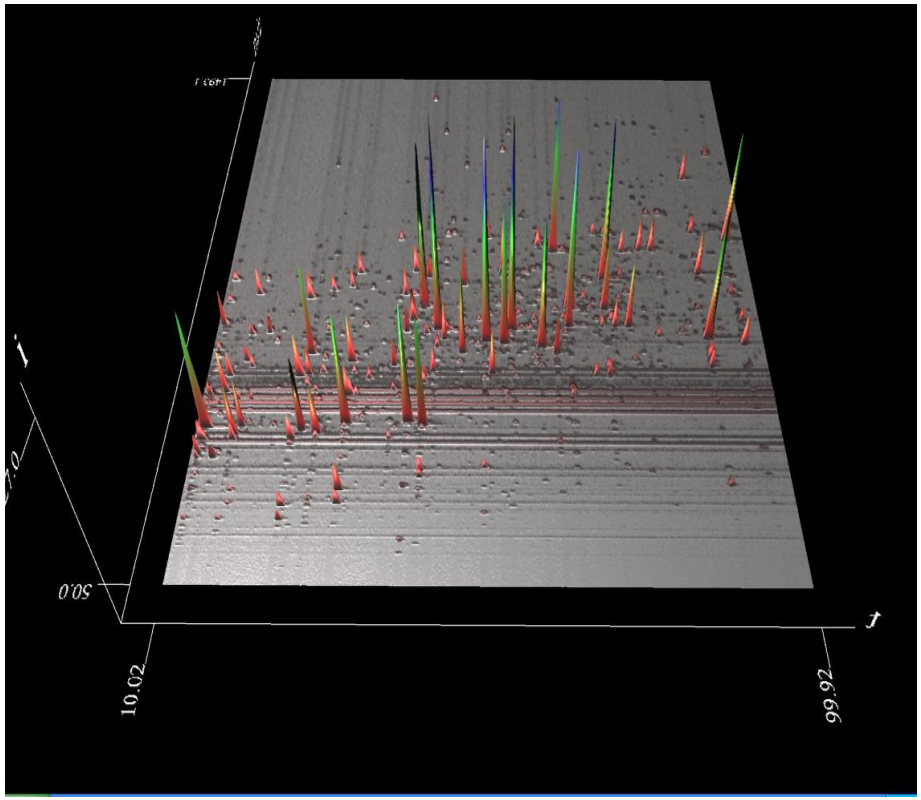
Antibody Heavy Chain Glycopeptides: Artifacts of In-Source Fragmentation?



Peptide Map Component Distribution

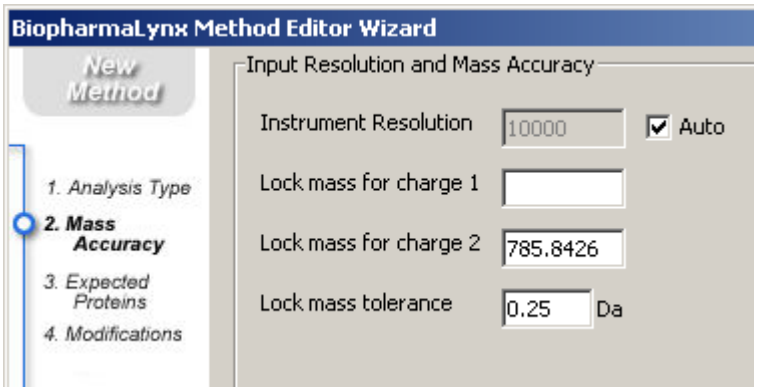
	Top100	All820
Identified	86%	63%
Tryptic	44%	9%
Semi-Tryptic	33%	49%
Source Fragment Source Ox Neutral Loss ESI Dimer	15%	29%
Missed Cleavage	8%	12%

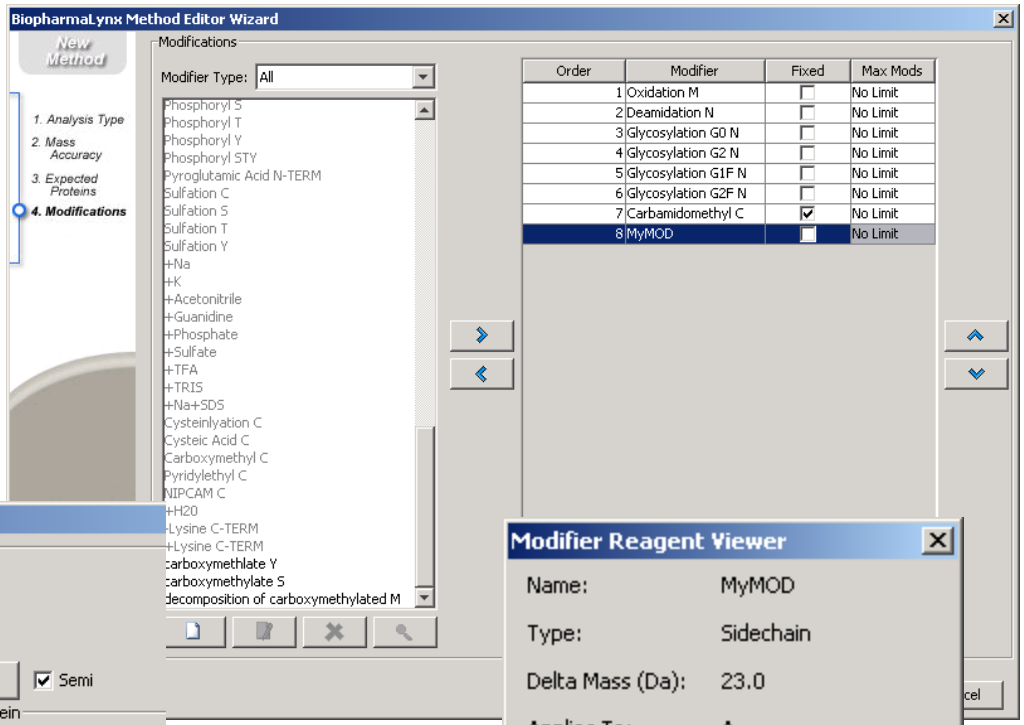
Enolase Peptide Map

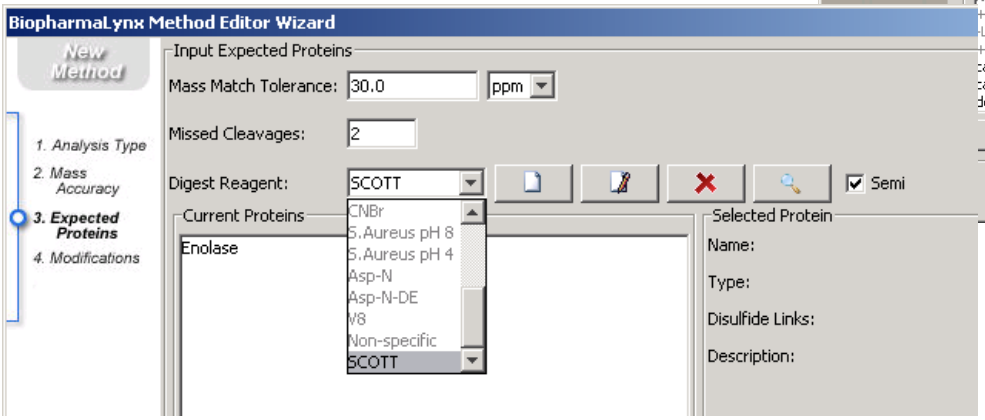


Informatics that only looks for Tryptic fragments does not comprehensively analyze a tryptic peptide map

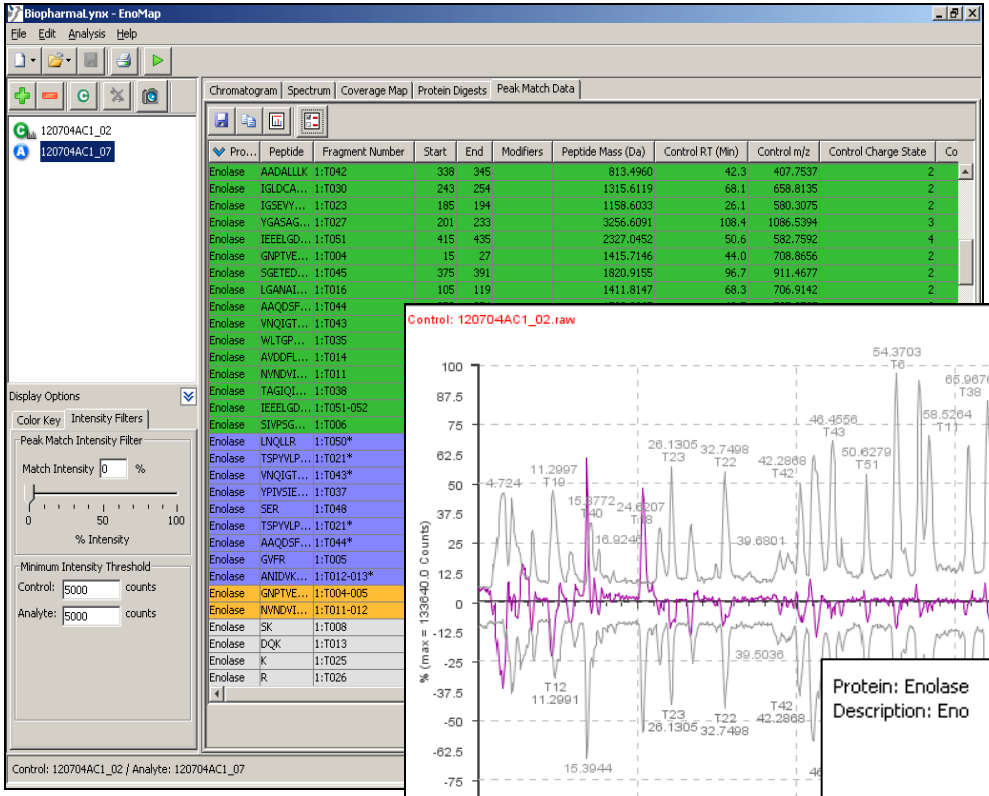
Peptide mapping does not require database searching

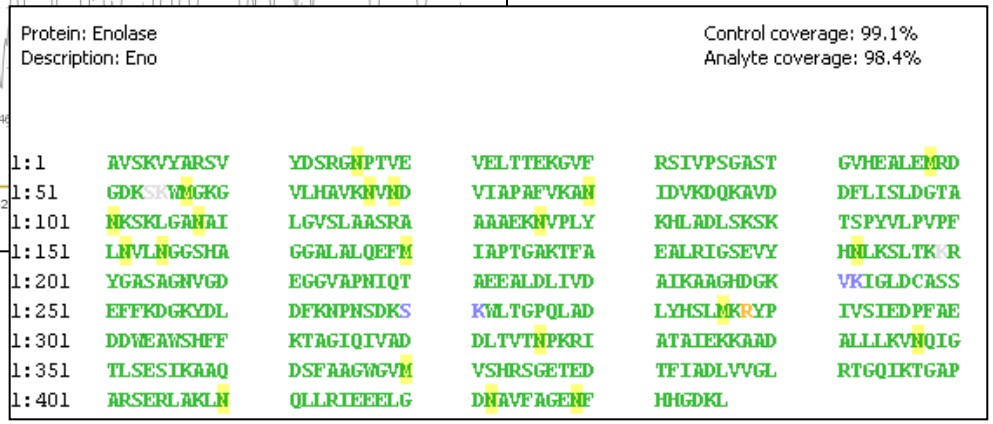






Automating data processing, display, and annotatation in BiopharmaLynx

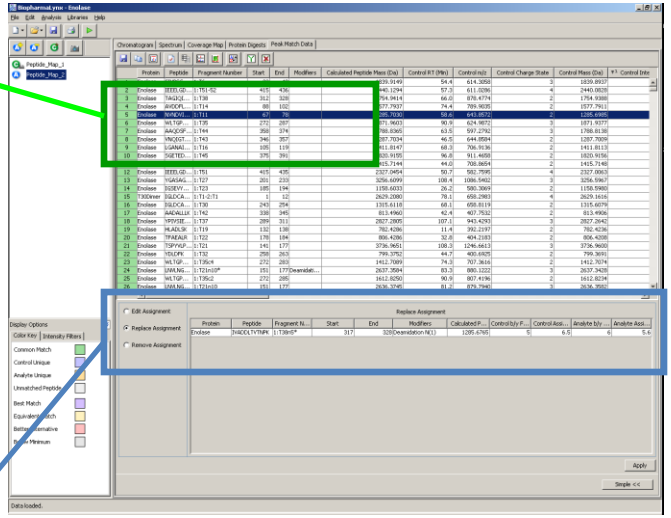




BiopharmaLynx™

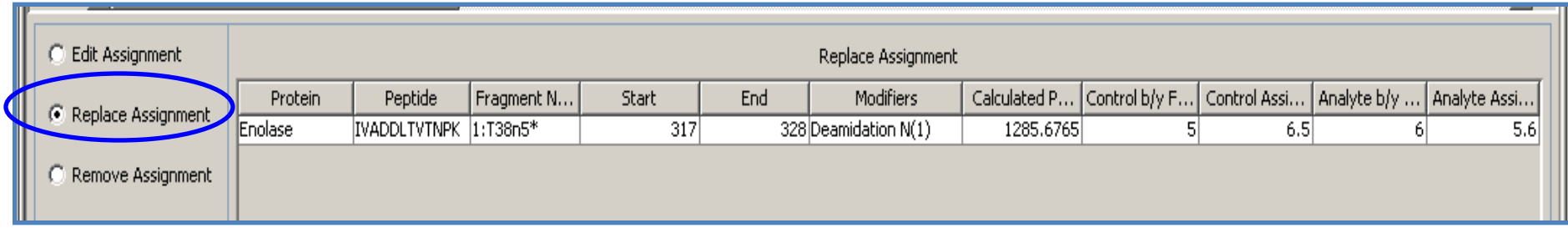
Overriding Automated Assignments

4	Enolase	AVDDFL...	1:T14	88	102		1577.7937
5	Enolase	NVNDVI...	1:T11	67	78		1285.7030
6	Enolase	WLTGP...	1:T35	272	287		1871.9603

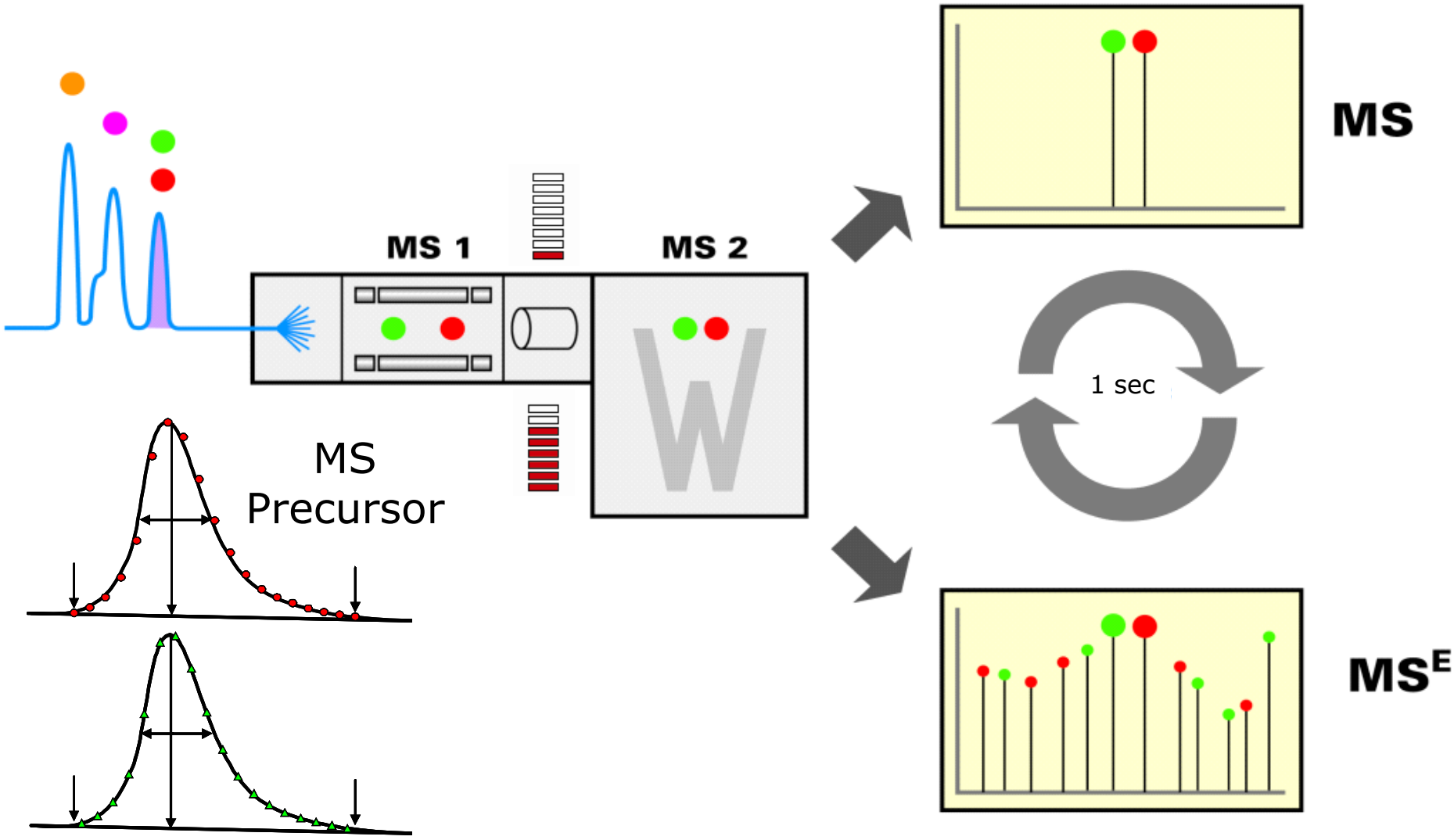


User can:

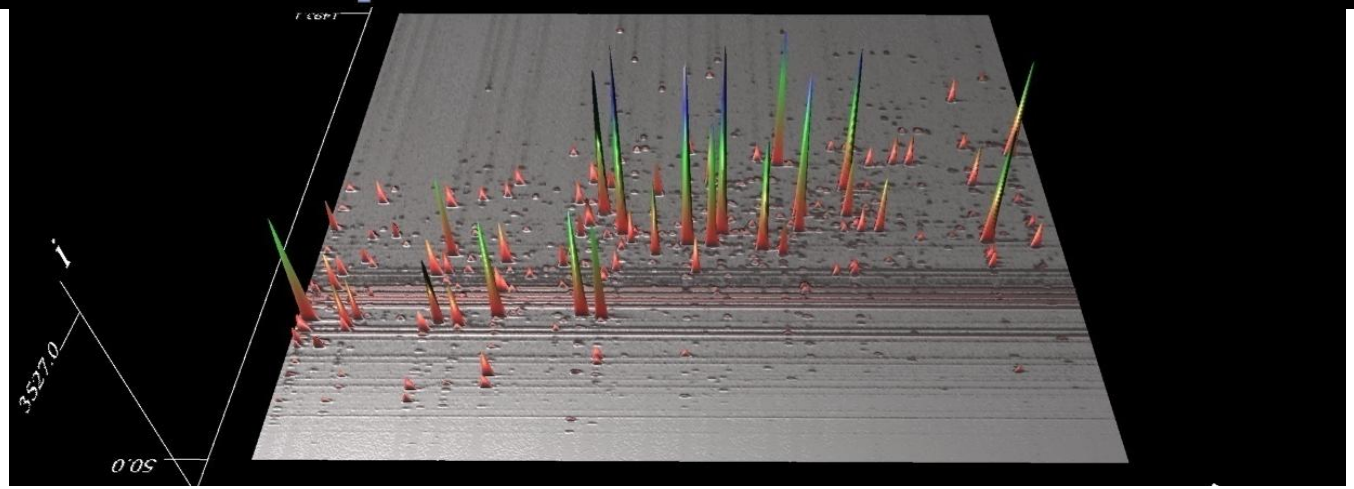
- Replace an assignment (Alternative)
- Manually edit an assignment
- Remove an assignment

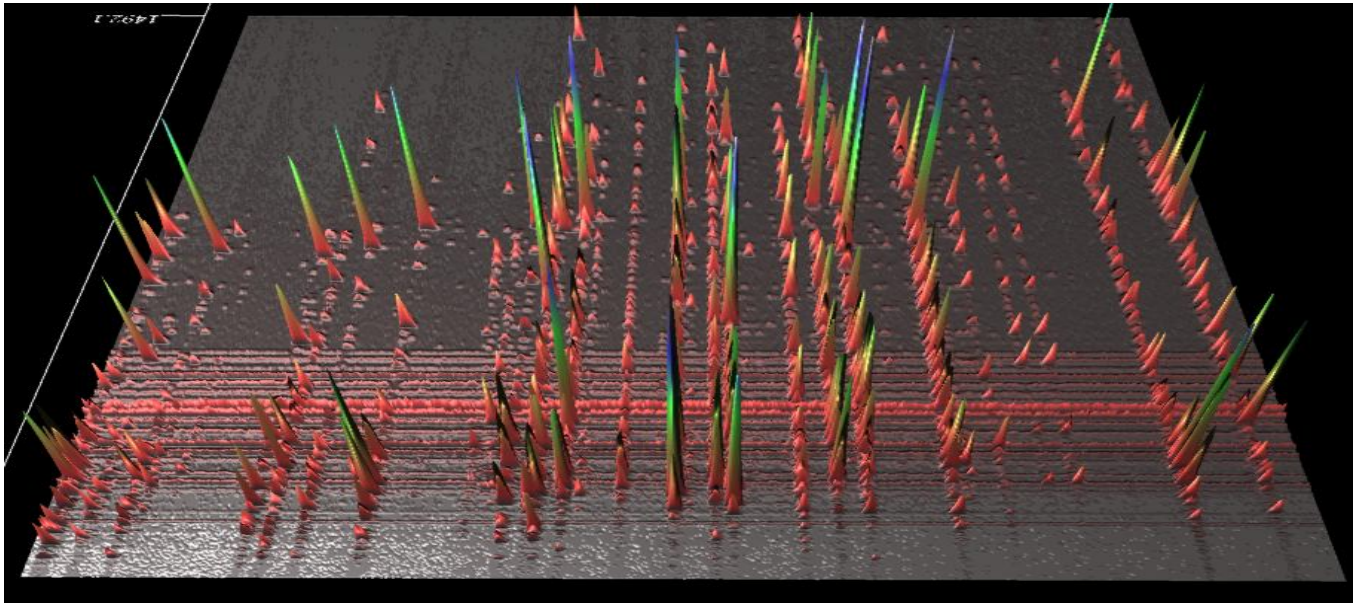


LC/MS^E Data Acquisition



Time Alignment of Low Energy (MS) and Elevated (MS^E) Energy Data



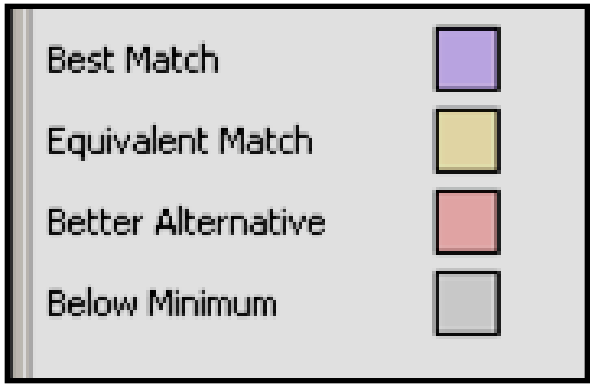


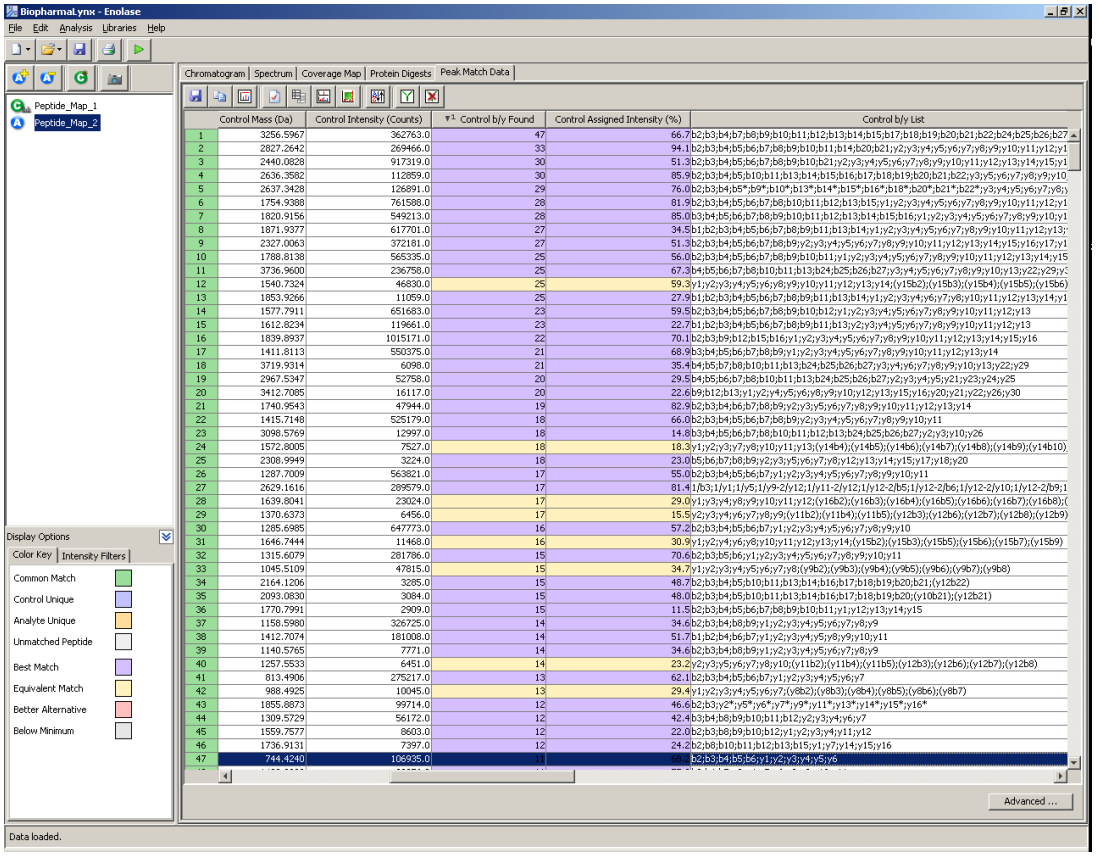
Retention Time (80 min)

Yeast Enolase Digest

BiopharmaLynx 1.2 : Validate Peptide Map Assignments with MS^E Fragmentation Data

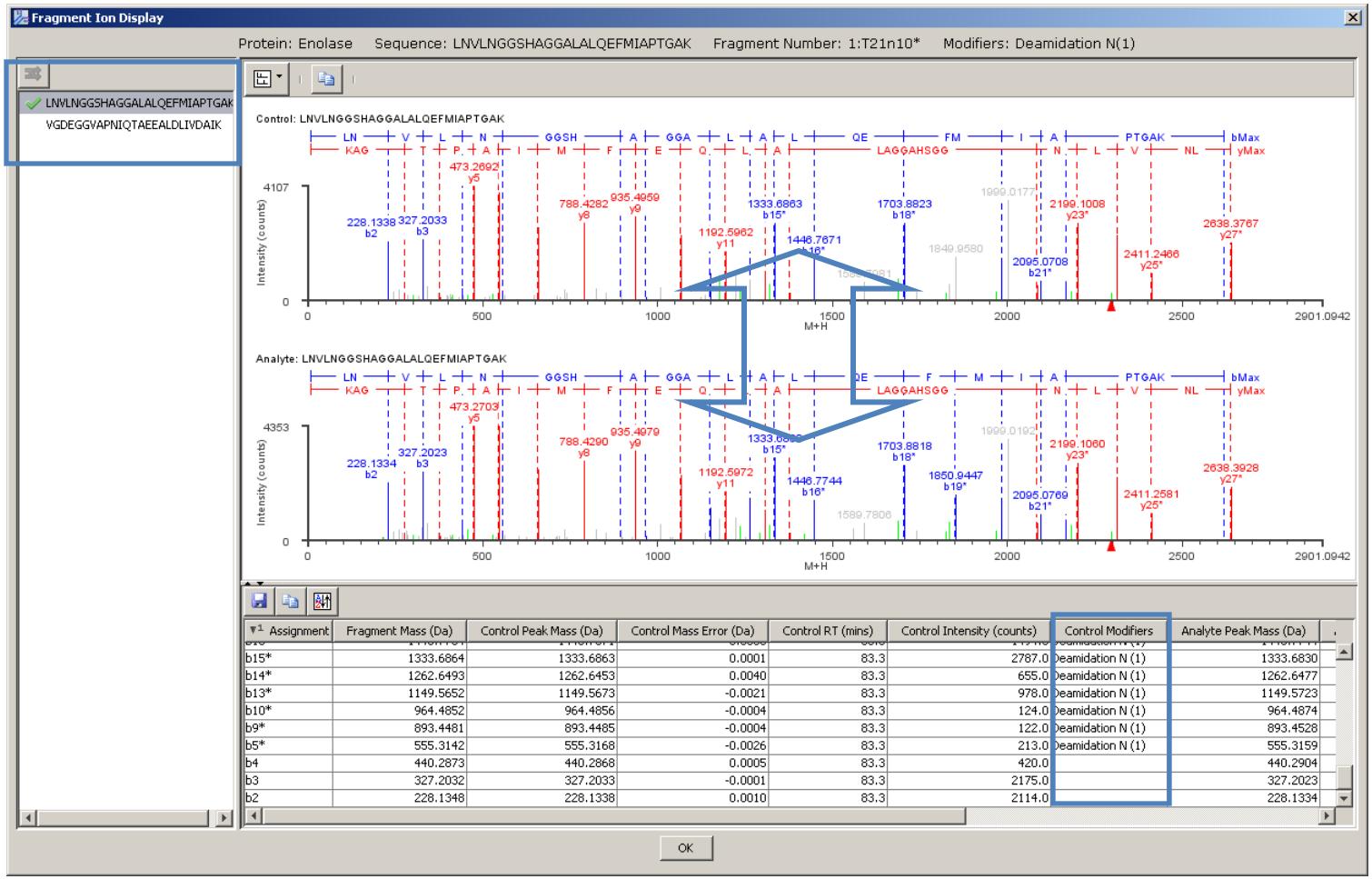
Control b/y Found	Control Assigned Intensity (%)	Control b/y List
47	66.7	b2;b3;b4;b7;b8;b9;b10;b11;b12;b13;b14;b15;b17;b18;b19;b20;b21;b22;b24;b25;b26;b27
33	94.1	b2;b3;b4;b5;b6;b7;b8;b9;b10;b11;b14;b20;b21;y2;y3;y4;y5;y6;y7;y8;y9;y10;y11;y12;y1





BiopharmaLynx 1.2 : Validate Peptide Map Assignments with MS^E Fragmentation Data

1. Fragmentation patterns can be compared between analyses
2. Potential alternative peptide assignments tested and selected
3. Modifications can be assigned to specific residues



Assignment	Fragment Mass (Da)	Control Peak Mass (Da)	Control Mass Error (Da)	Control RT (mins)	Control Intensity (counts)	Control Modifiers	Analyte Peak Mass (Da)
b15*	1333.6864	1333.6863	0.0001	83.3	2787.0	Deamidation N (1)	1333.6830
b14*	1262.6493	1262.6453	0.0040	83.3	655.0	Deamidation N (1)	1262.6477
b13*	1149.5652	1149.5673	-0.0021	83.3	978.0	Deamidation N (1)	1149.5723
b10*	964.4852	964.4856	-0.0004	83.3	124.0	Deamidation N (1)	964.4874
b9*	893.4481	893.4485	-0.0004	83.3	122.0	Deamidation N (1)	893.4528
b5*	555.3142	555.3168	-0.0026	83.3	213.0	Deamidation N (1)	555.3159
b4	440.2873	440.2868	0.0005	83.3	420.0		440.2904
b3	327.2032	327.2033	-0.0001	83.3	2175.0		327.2023
b2	228.1348	228.1338	0.0010	83.3	2114.0		228.1334

Identifying Unknown Process Impurity Proteins (HCP's)

LC/MS^E Data

IDENTITY^E high definition proteomics

Match to: **Protein_X** Score: **89**

```
  1 MTQFTDIDKL AVSTIRILAV DTVSKANSGH PGAPLGMAPA AHVLWSQMRM
 51 NPTNPDWINR DRFVLSNGHA VALLYSMLHL TGYDLSIEDL KQFRQLGSRT
101 PGHPEFELPG VEVTTGPLGQ GISNAVGMAM AQANLAATYN KPGFTLSDNY
151 TYVFLGDGCL QEGISSEASS LAGHLKLGNL IAIYDDNKIT IDGATSISFD
201 EDVAKRYEAY GWEVLYVENG NEDLAGIAKA IAQAKLSKDK PTLIKMTTTI
251 GYGSLHAGSH SVHGAPLKAD DVKQLKSKFG FNPDKSFVVP QEVYDHYQKT
301 ILKPGVEANN KWNKLFSEYQ KKFPELGAEL ARRLSGQLPA NWESKLPTYT
351 AKDSAVATRK LSETVLEDVY NQLPELIGGS ADLTPSNLTR WKEALDFQPP
401 SSGSGNYSGR YIRYGIREHA MGAIMNGISA FGANYKPYGG TFLNFVSYAA
451 GAVRLSALSG HPVIWVATHD SIGVGEDGPT HQPIETLAHF RSLPNIQVWR
501 PADGNEVSAA YKNSLESKHT PSIIALSRQN LPQLEGSSIE SASKGGYVLQ
551 DVANPDIILV ATGSEVSLSV EAAKTLAAKN IKARVVSLPD FFTFDKQPLE
601 YRLSVLPDNV PIMSVEVLAT TCWGKYAHQS FGIDRFGASG KAPEVFKFFG
651 FTPEGVAERA QKTIAFYKGD KLISPLKKAF
```

BiopharmaLynx™

Unassigned Components

.PKL Export File

ProteinLynx Global Server
MASCOT
Other Bioinformatics Tools

Conclusions

- LC/ESI-Tof MS a powerful tool for protein characterization and peptide map development
- The resulting data sets are complex
  - Sample Heterogeneity
  - Sample Processing Artifacts
  - LC and MS Analysis Effects
- Our understanding of these processes has enhanced our ability to build better informatics tools.
  - Analysis can be automated (Days -> Minutes)
  - Increase uniformity and confidence of results
  - Eases sample-sample comparisons

Examination of a Biotherapeutic Characterization Strategy.

GEN Webinar Series 2009

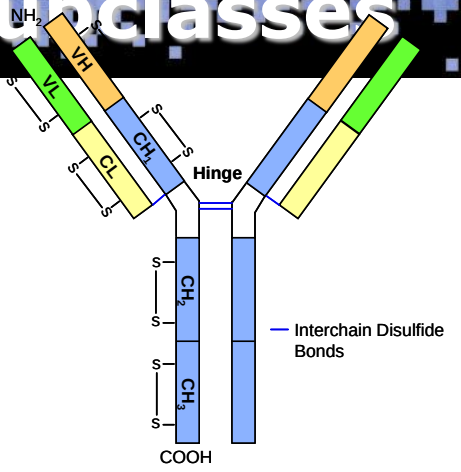
Adam Lucka

Alexion Pharmaceuticals
Protein Characterization Group

Human Antibody Structure and IgG Subclasses

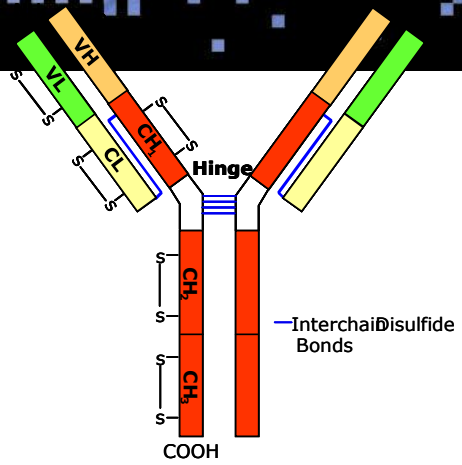
Rise to Greatness™

have three constant regions with more than 95% homology



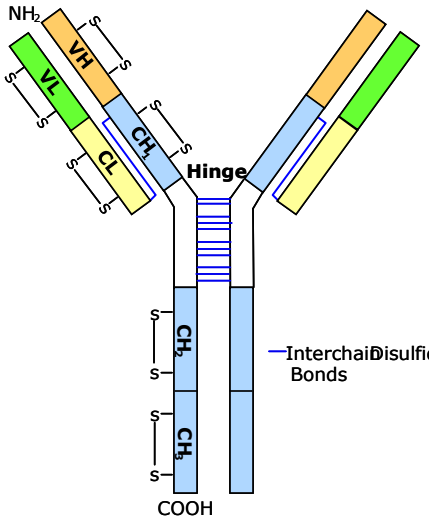
IgG_1

Two disulfide bonds
MW~148,480 Da



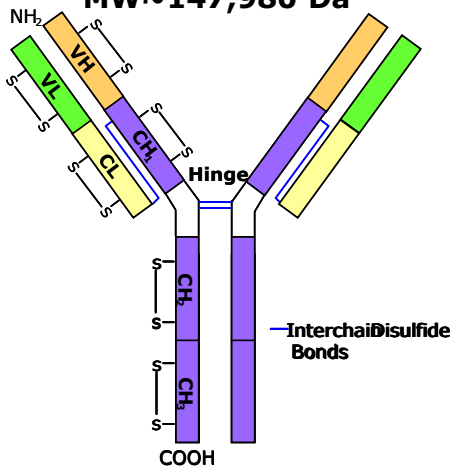
IgG_2

Four disulfide bonds
MW~147,980 Da



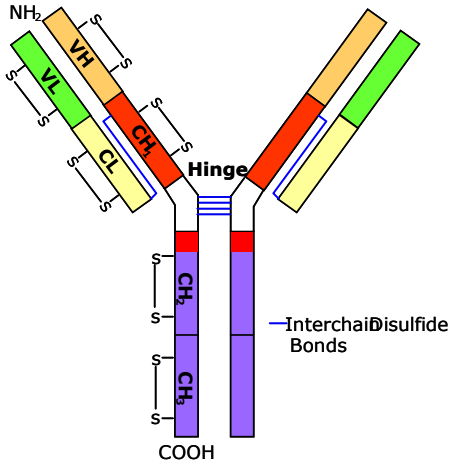
IgG_3

Eleven disulfide bonds



IgG_4

Two disulfide bonds
MW~148,030 Da



$IgG_{2/4}$

Four disulfide bonds

MW~147,870 Da

Characterizing Biotheraputics and Impurities

- Bulk Drug Substance: Formulated drug material that containing all molecular variants of the drug at the time of manufacture. (glycoforms, PTM, CAM, minor clips)

- Product Related Impurities: Molecular variants of the drug that may be present at the time of manufacture new lots/processes or formed during storage. (Truncated forms, aggregation, isomerized, mismatched S-S, deamidated, oxidized, etc.)

Protein Characterization Core

Mass Spectrometry

- MALDI-ToF: Intact MW, Free Glycan, PMF
- ESI-ToF: Intact MW
- ESI-Q-IM-ToF: Peptide Mapping, Disulfide Analysis, Small Molecule

HPLC

- Oligosaccharide
- Monosaccharide
- Sialic Acid
- SEC, RP, WAX

N-terminal Sequencing
AAA
Gel Based Separations
Chip Based Separations
AFFF-MALS
Ion Mobility
Calorimetry
Circular Dichroism
BIACORE SPR

Alexion's Approach to Drug Substance Characterization

Top-Down, Bottom-Up, Inside-Out

- Intact MW: Top Down

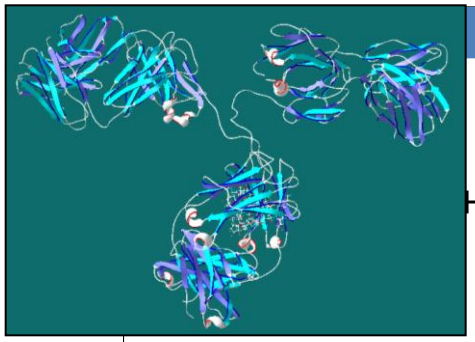
- Proteolytic Digests Peptide Maps: Bottom Up
- MS/MS sequencing: Bottom Up

- Non-reduced Digests for Disulfide Confirmation: Inside Out
- Glycopeptide and Free Glycan analysis: Inside Out

Intact MW: Top Down

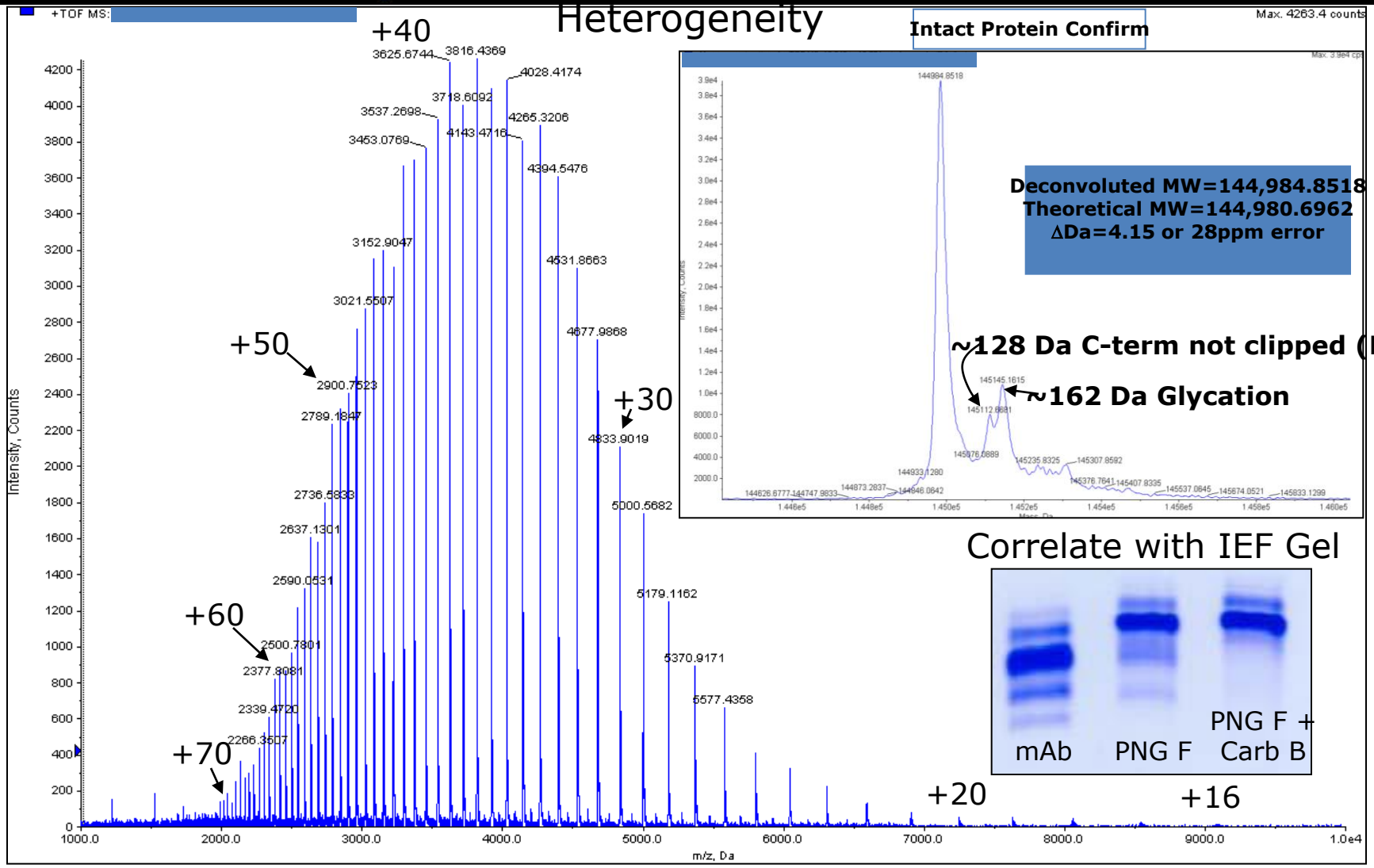
Every Day Matters™

Product Related Substance



De-Glycosylated mAb

Rise to Greatness™



Drug Characterization LC-MS Peptide Mapping: Bottom Up

- **Trypsin Digestion**
- **RP-HPLC-MS**
- **MS Data:**
  - Accurate Retention time
  - Accurate Mass
  - Peptide Confirmation via Fragmentation (MS/MS or MS^E)

Use and correlate data from multiple techniques:

Work for 100% primary sequence coverage

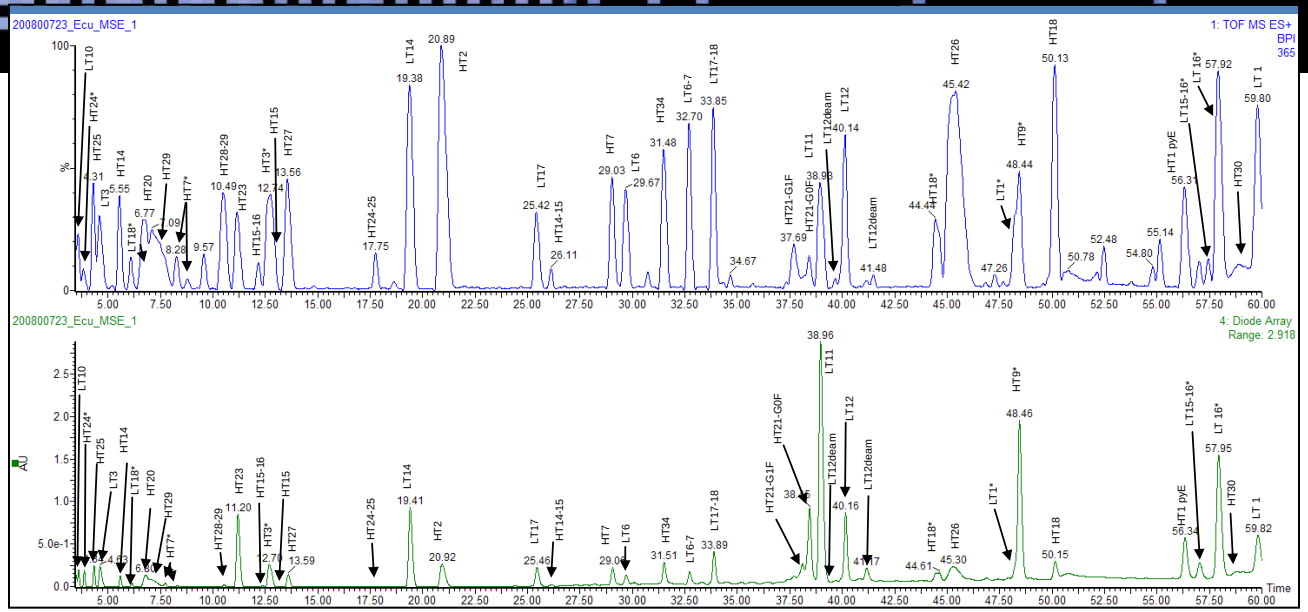
mAb Drug Substance Peptide Map

Rise to Greatness™

BPC

UV

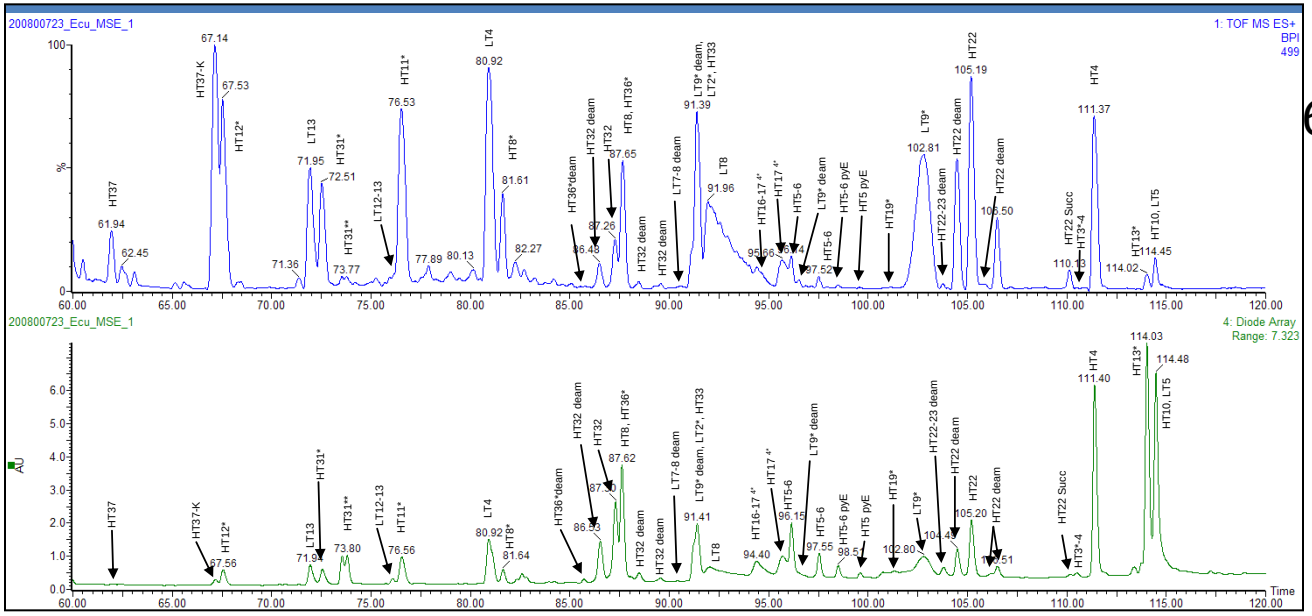
0-60 min



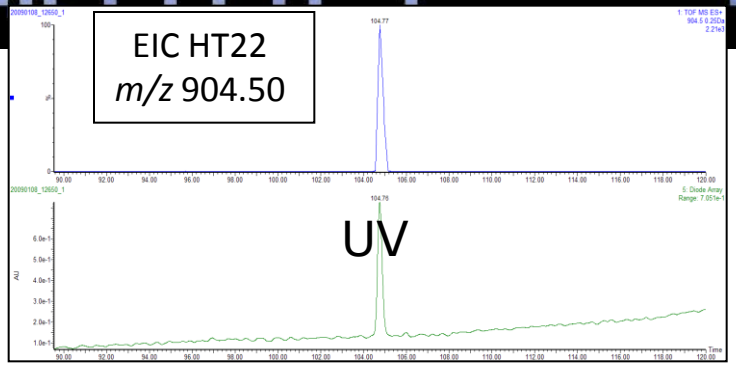
BPC

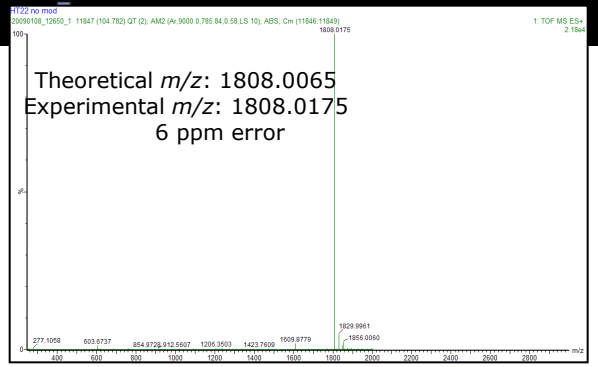
UV

60-120 min

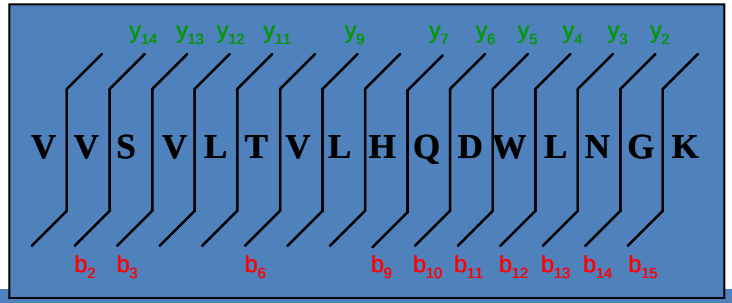


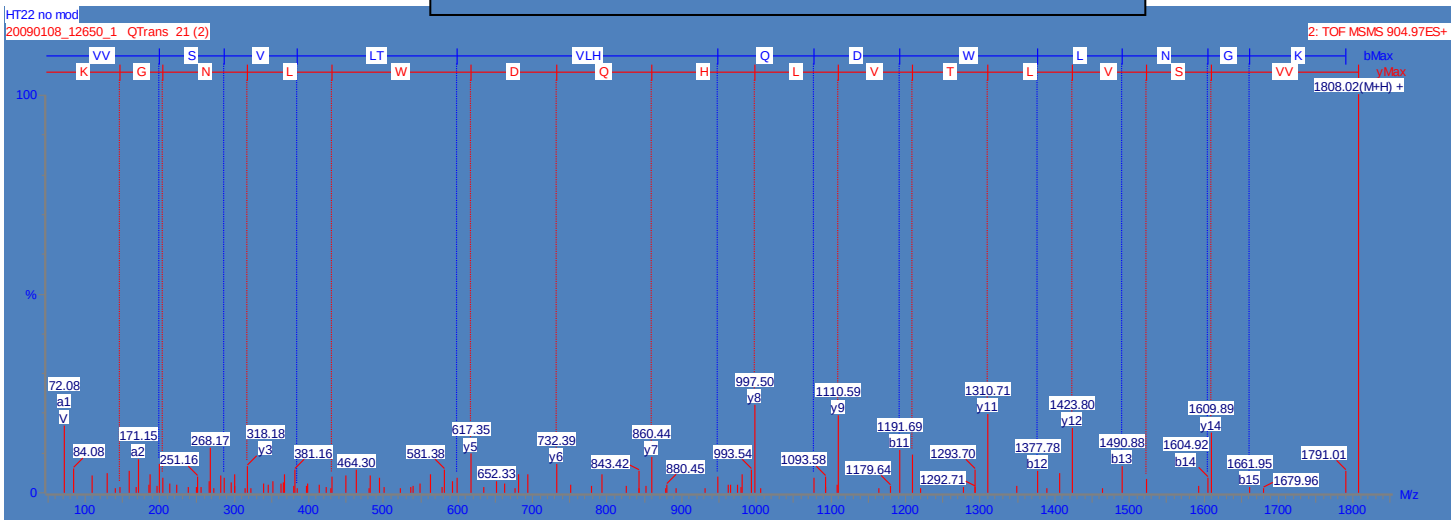
eptides Identified by RT, Accurate Mass





Accurate RT **Accurate Mass**

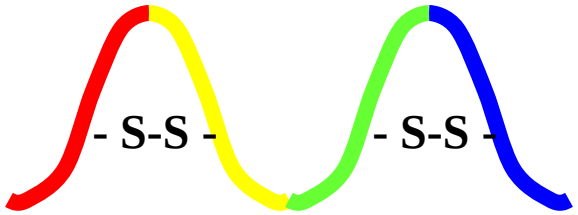




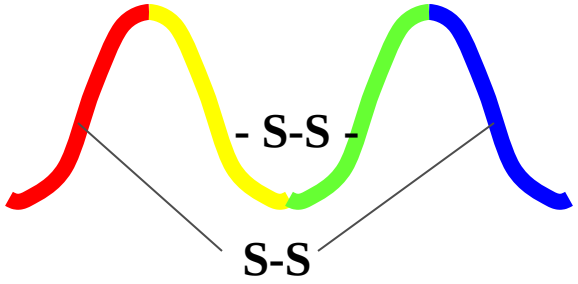
Disulfide Confirmation Via Non-Reduced Peptide Mapping

Every Day Matters™

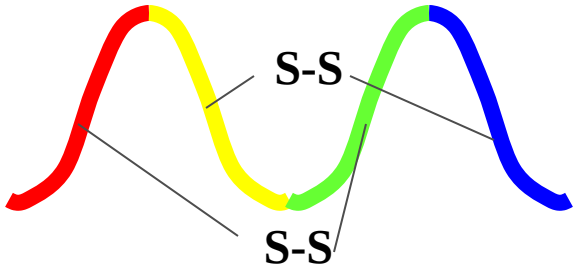
scFv 2-6, 11-17

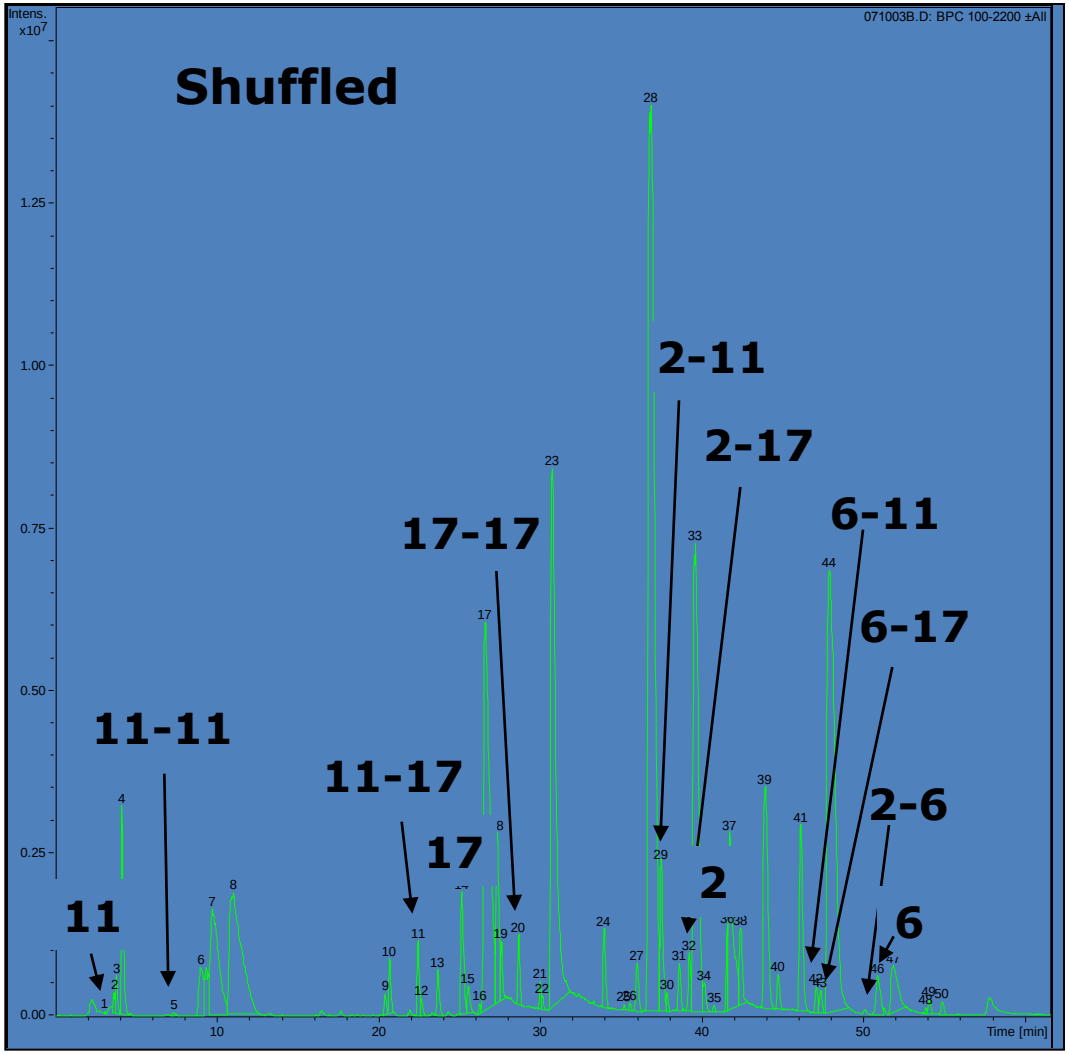


scFv 2-17, 6-11



scFv 2-11, 6-17



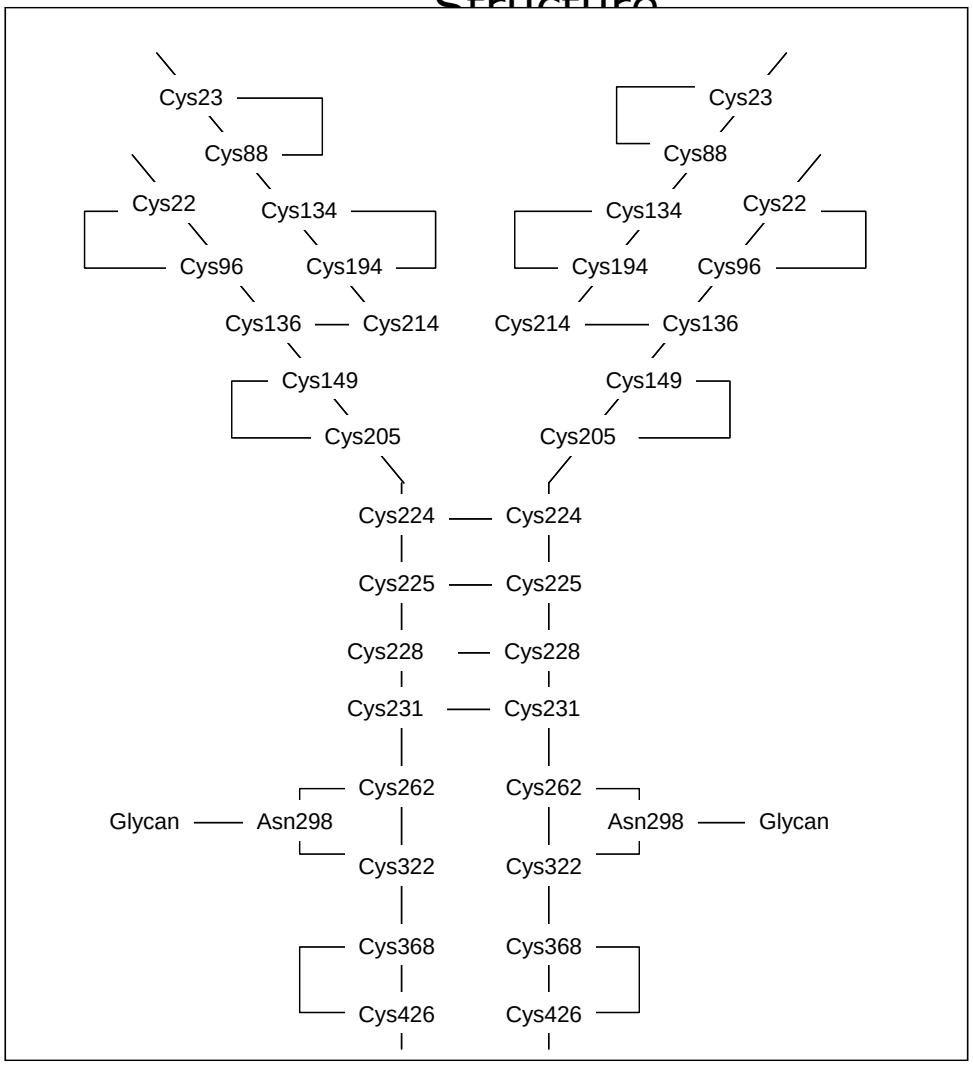


Intra and Four Additional Inter Disulfide Cross-Links 2-2, 6-6, 11-11, 17-17

Disulfide Confirmation more complex for IgG

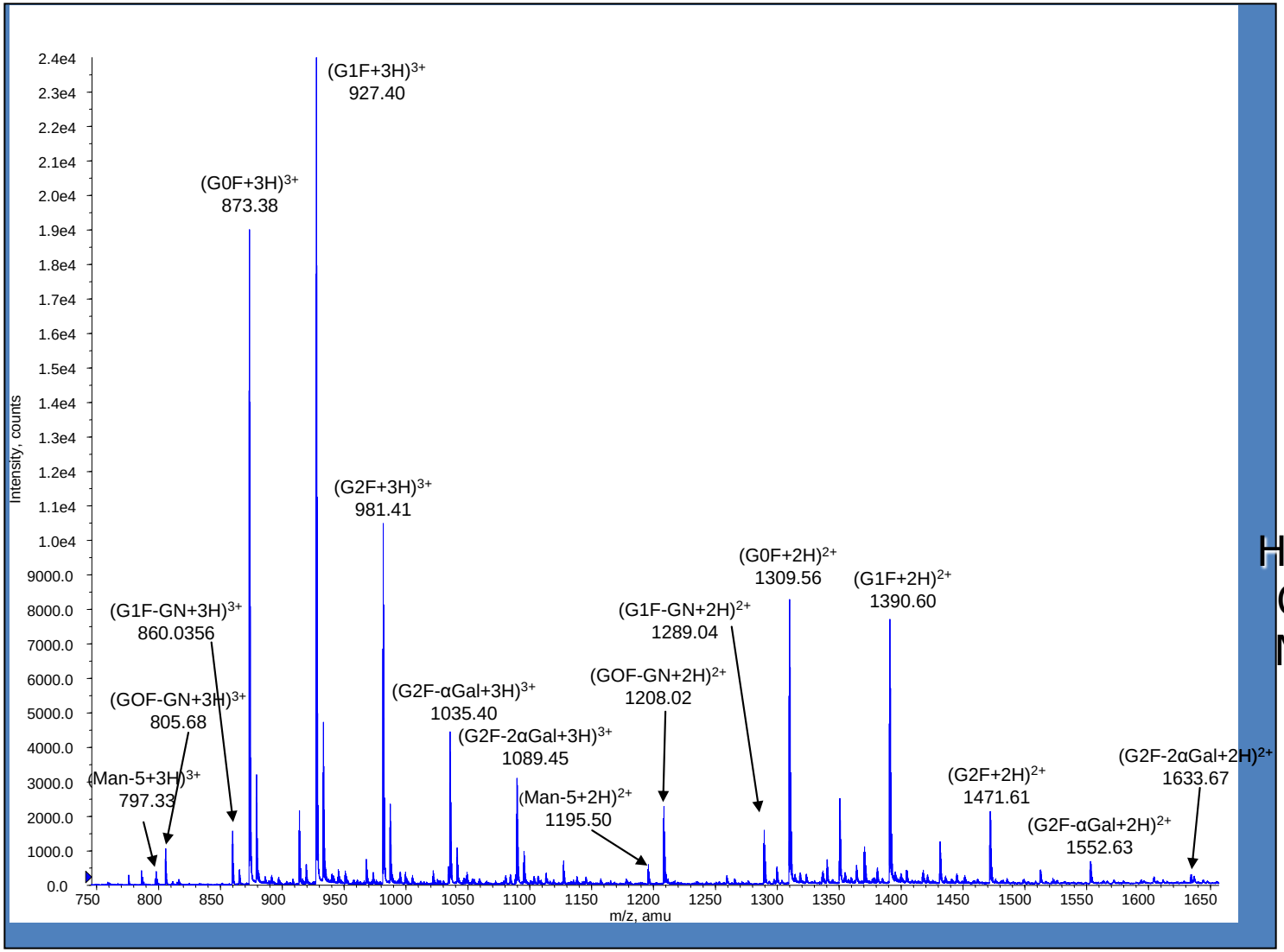
Rise to Greatness™

Disulfide Bonding Pattern for Human IgG2/4 Structure



18 total disulfides
11 unique disulfides

N-Linked Glycopeptide ESI-nLC-TOF



Correlated with MALDI-TOF Free Glycan

HPLC fluorescence Oligosaccharide, Monosaccharide, Sialic Acid

Analytical Summary of mAb LC-MS Peptide Mapping

Rise to Greatness™

% Expected Heavy Chain Sequence Identified	100%
% Expected Light Chain Sequence Identified	100%

Terminal Analysis	Sequences
Heavy Chain N-terminal	As expected
Light Chain N-terminal	As expected
Full Heavy Chain C-terminal (Relative %)	14.0% intact peptide
Clipped Heavy Chain C-terminal (Relative %)	86.0 % clipped peptide

Glycosylation Analysis	
Glycosylation Site (Asn_{298})	100% Occupancy of Asn_{298} observed. No evidence for O-glycosylation

Disulfide Analysis
HT3s-sHT9 (Cys_{22} – Cys_{96})
HT11s-sLT18 ($^{H}Cys_{136}$ – $^{L}cys_{214}$)
HT12s-sHT13 (Cys_{149} – Cys_{205})
HT17ssss-ssssHT17 (Cys_{224} – Cys_{224}), (Cys225 – Cys225), (Cys228 – Cys228), (Cys231 – Cys231)
HT19s-sHT24 (Cys_{262} – Cys_{322})
HT31s-sHT36 (Cys_{368} – Cys_{426})
LT2s-sLT5 (Cys_{23} – Cys_{88})
LT9s-sLT16 (Cys_{134} – Cys_{194})

All Disulfides confirmed with minor evidence of alternative isoforms

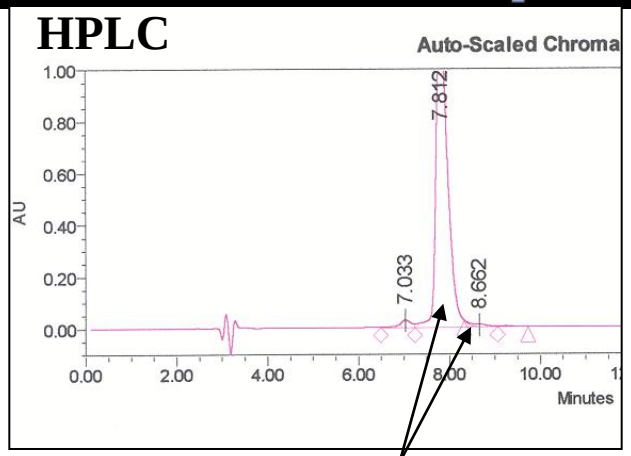
Detected Glycopeptides	Relative %
G0F-ManGN	5.6
Man5	7.2
G0F-GN	28.2
G1F-GN	5.3
G0F	100.0
G2F-GN	1.0
G1F	46.5
G2F	9.7
G1F+NGNA	3.3
G2F+NGNA	1.5

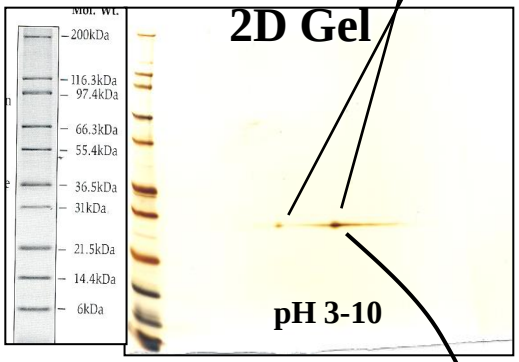
Deamidation Analysis	% Unmodified	% Deamidated
Peptide HT22 (303-318) Asn_{316}→Asp_{316}	43.7	56.3
Peptide HT32 (372-393) Asn_{385}→Asp_{385}	59	41
Peptide HT36*(418-440) Asn_{435}→Asp_{435}	78.8	21.2
Peptide LT9* (127-142) Asn_{137}→Asp_{137}	80.6	19.4
Peptide LT12* (150-169) Asn_{158}→Asp_{158}	94.1	5.9

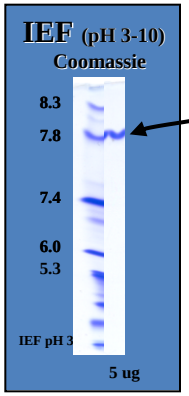
Oxidation Analysis	% Unmodified	% Oxidized
Peptide HT8 (73-87) M_{81}	5.2	94.8
…	…	…

Product Related Impurities

Rise to Greatness™







and Spots

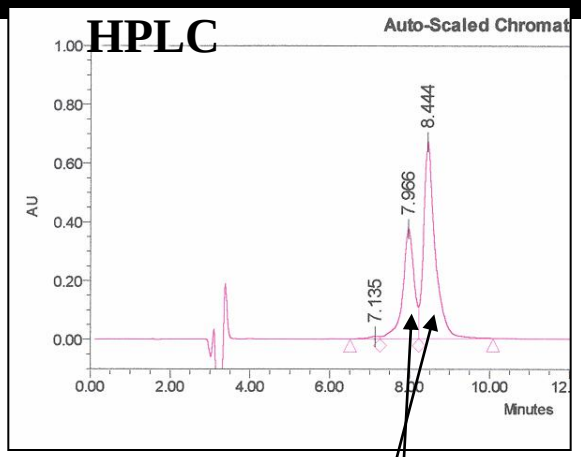
Accelerated Stability, Temperature Stressed, Oxidation, Deamidation, Acetylation, Carbamylation, …etc.

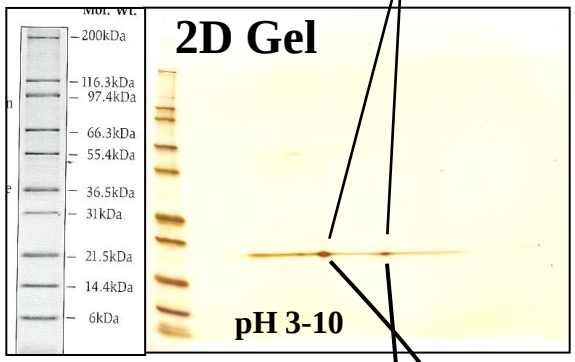
- **Identify susceptible residues**
- **Identify Accurate RT location**

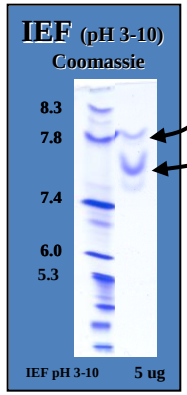
How Characterization Challenges Resembles Proteomics Challenges

100:1 to 1000:1 Purity
100:1 ionization efficiency
peptides/spectrum

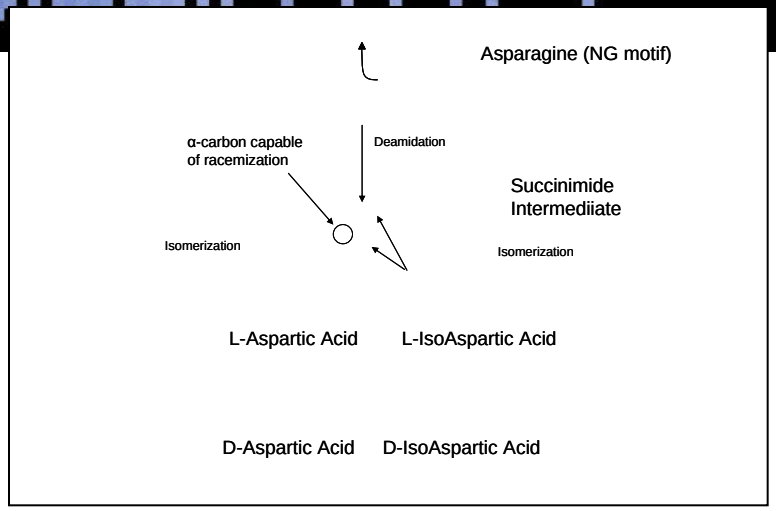
10,000:1 to 100,000:1



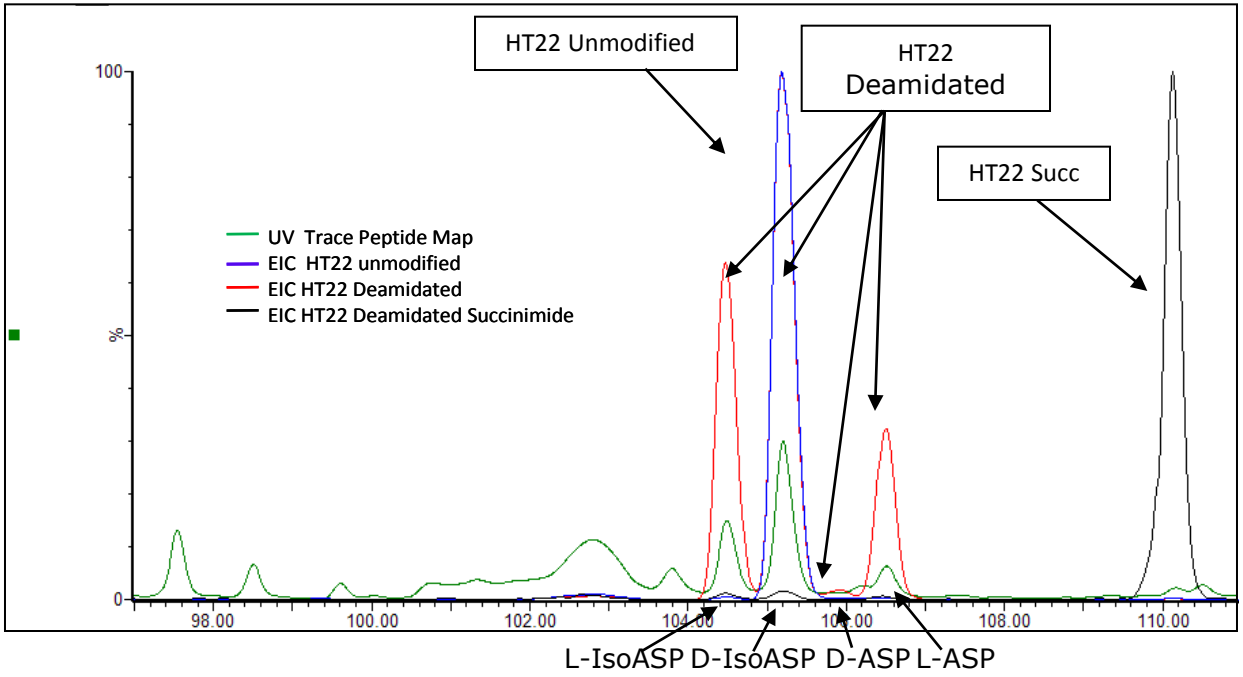




Deamidation



Characterizing Five Deamidation Products from a Single Asparagine Residue in Recombinant mAb's Through LC-MS Peptide Mapping.
Lucka A. et al. ASMS 2009



Additional Characterization Workflows

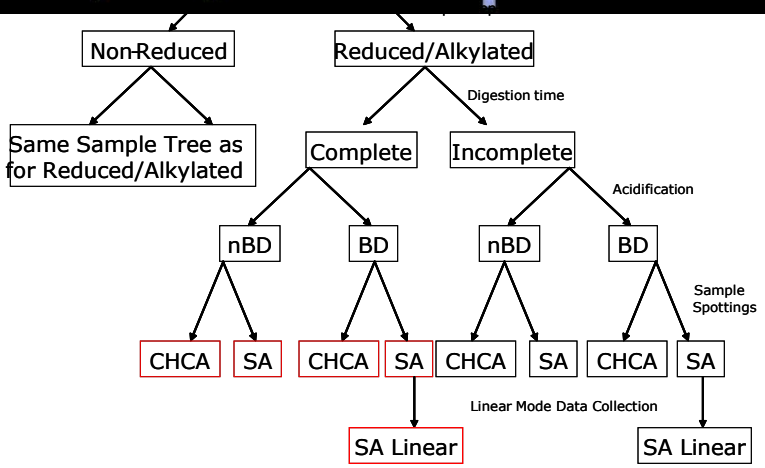
Rise to Greatness™

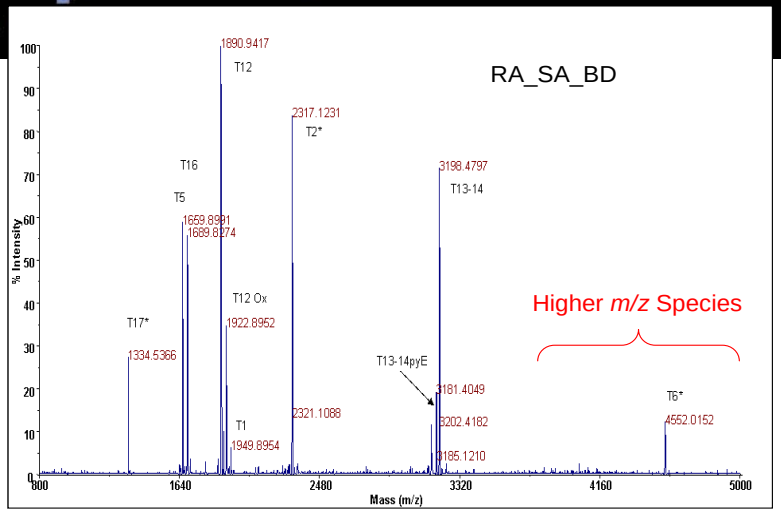
Additional Strategies to get 100% primary sequence coverage

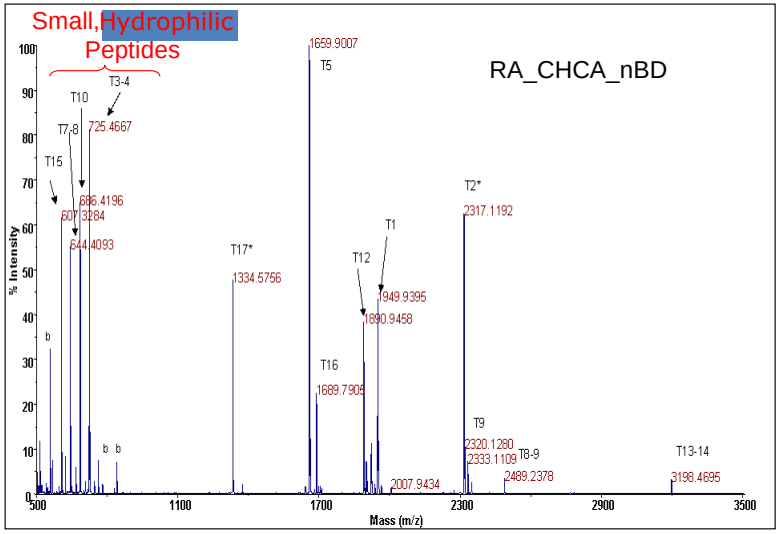
- MALDI Strategies for PMF – multiple sample preps + composite searches
- Sub-ambient trapping for nano/cap LC – Peptide Mapping
- Alexion's Bioinformatics Strategy for rapid data reduction

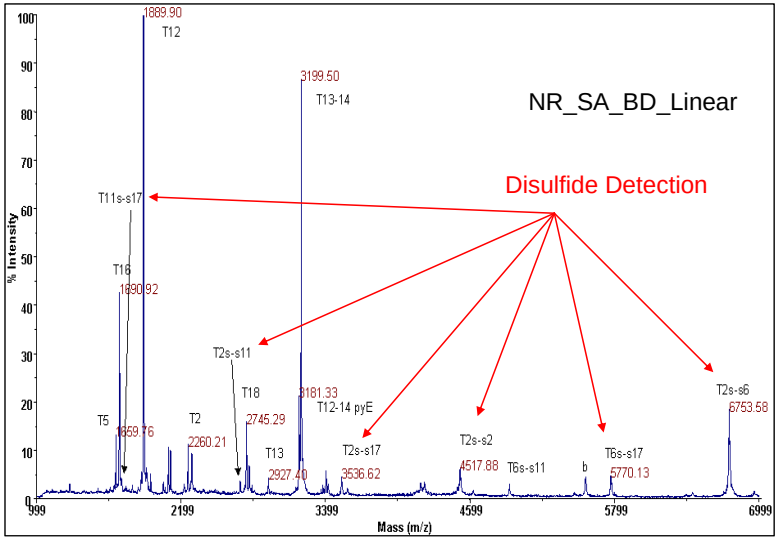
Rapigest Digestions via MALDI

Rise to Greatness™









Sub-ambient Trapping of Hydrophilic Peptides

Every Day Matters™

Research Peptide Maps
Nano/Cap HPLC

1100 HPLC

LC/MSD
ESI-ToF

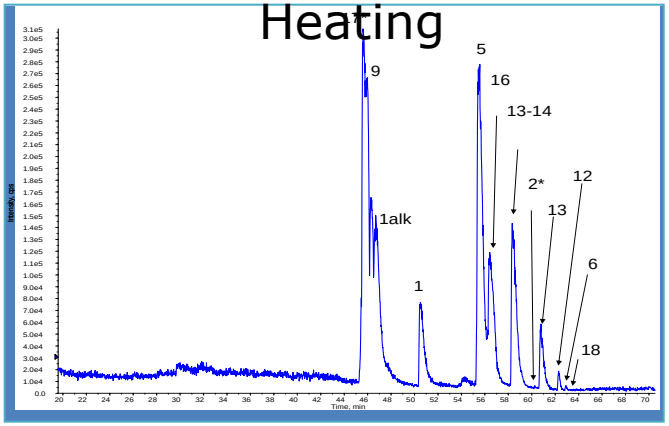
Synapt HDMS
QToF-IMS

Nano/Cap RP C-18 column

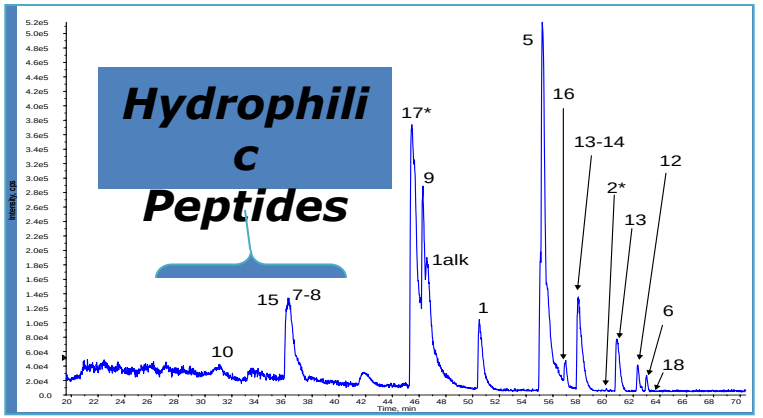
Solvent A: 0.1% FA in H_2O
Solvent B: 0.1% FA in ACN

Peptides were eluted via linear gradient from 2-50% Solvent B in 90 min flow rate: 0.3-5 µL/min.
ESI 3.5 KV and a *m/z* range of 50-3000 was analyzed.

Room Temperature Enrichment no Column Heating



Sub-ambient Enrichment



Proteomic Search Engines are Inadequate for Protein Characterization

- System Issues
  - Uses single search mode mono *vs.* average mass
- Identification Issues
  - Incorrect Id's with too many mods pre-selected
- Missing Features
  - Neglect retention time and archival information
  - Neglect known assay artifacts
- Data Presentation Issues
  - No back association to the LC-MS chromatogram/MALDI spectra

Heterogeneity / PTM, CAM	Δ Mass Change (Da)
C-terminal Lysine	-128.1726
Disulfide	-2.0158
Pyroglutamation	-17.0306
G0F Glycan	1445.3355
Carboxymethylation of Cys/Met	58.0055
SNP	Varies
Oxidation	15.9949
Acetylation	42.0106
Deamidation	0.984

- false positive protein identifications

- false positive peptide identifications for correct protein

- missed peptides for a correctly identified protein

Data Analysis is the Bottleneck for Protein Characterization

Alexion's Bioinformatics Strategy

Rise to Greatness™

- Fully document each peptide RT, Accurate Mass, and fragmentation confirmation, after which only Accurate Mass and RT time pairs are screened
- Single search to identify all known peptides, regardless of signal intensity, currently requires manual interrogation of the data
- Identifies all potential peptides that meet criteria
- Protein specific database for sequence id/confirmation:
  - normal, mis-cleaved, nonspecific cleavages, single point mutations, PTM, CAM, cross-linked, and assay related peaks
- Correlates peptide *m/z*, charge, retention time, intensity and fragmentation. Tracks and trends peptide information in an archival Master Table which is chromatography and molecule specific.
- Suggests mods for unknowns based on the specific protein
- Visual Results output

Distilled BiopharmaLynx Data

Peptide	Fragment Number	Modifiers	Control RT (Min)	Control Intensity (Counts)	Control m/z	Control Charge State	Control Mass (Da)	b/y Possible	Control b/y Found	Control b/y %	Control b/y List	Analyte RT (Min)	Analyte Mass (Da)	Delta RT (Min)	Delta Intensity (%)	Comment
CKVSNK	1:T024-025*	Carboxymethyl C(1)	17.8	11979	736.3651	1	735.3571	10	0	0		17.7	735.3548	0.1	10.6	
			18.7	2130	618.3457	1	617.3378					18.7	617.3371	0	9.8	
DYEK	2:T014/y4		19.4	4889	554.2432	1	553.2353	6	4	66.7	y1;y2;(y4b2);(y4b3)	19.2	553.2368	0.1	12.6	
ADYEK	2:T014x2		19.4	7876	1249.553	1	1248.545	8	0	0		19.2	1248.549	0.1	7	
ADYEK	2:T014		19.4	113777	625.2786	1	624.2707	8	6	75	b3;b4;y1;y2;y3;y4	19.3	624.2691	0.1	7.9	
			19.5	2089	647.2622	1	646.2543									
KPGASVK	1:T002x2		20.9	4015	1371.829	1	1370.821	12	0	0		20.9	1370.826	0	3.4	
KPGASVK	1:T002-NH3		20.9	2064	335.1991	2	668.3824	12	7	58.3	b1;b2;b4;b5;y1;y3;y4					
PGASVK	1:T002/y6		20.9	8737	558.3215	1	557.3136	10	8	80	y1;y3;y4;y5;(y6b2);(y6b3);(y6b4);(y6b5)	20.9	557.3146	0	17.9	
KPGASVK	1:T002		21	210469	686.4184	1	685.4105	12	10	83.3	b1;b2;b4;b5;b6;y1;y3;y4;y5;y6	21	685.4113	0	11.6	
			21.1	1431	739.3319	1	738.324									
			21.1	6890	708.4005	1	707.3925					21.1	707.3922	0	6.4	
SFNR	2:T017x2		25.4	8243	1045.515	1	1044.507	6	0	0		25.3	1044.509	0.1	1.8	
SFNR	2:T017		25.4	28977	523.2538	1	522.2459	6	2	33.3	y1;y3	25.3	522.2458	0.1	2.7	
			25.5	1743	1067.497	1	1066.49									
			26.1	1289	868.4534	1	867.4455									
VDKTVER	1:T014-015		26.1	6845	846.468	1	845.4601	12	1	8.3	y5	26.1	845.4593	0.1	30.1	

RT, Intensity, m/z, and Charge State

Alexion's Bioinformatics – Mastra

Every Day Matters™

Molecular Species ID

m/z Stats

Retention Time Stat

Trypsin #	AA Range	Mono MH+	Ave. MH+	Partials	Sequence	Ave RT	Min RT	Max RT	SD RT
1	[1- 19]	1949.923	1951.102	0	ADIQMTQSPSSLSASVGDR	31.15	30.70	31.60	0.18
1-2 C* (gel)	[1- 39]	4247.013	4249.64	1	ADIQMTQSPSSLSASVGDRVTITCGASENIY GALNWYQR (1 * Carbamidomethyl (C))	43.40	43.40	43.40	
2 C* (gel)	[20- 39]	2316.108	2317.561	0	VTITCGASENIYGALNWYQR (1 * Carbamidomethyl (C))	42.34	41.60	45.10	1.23
2-4 C* (gel)	[20- 46]	3022.557	3024.439	2	VTITCGASENIYGALNWYQRKPGKAPK (1 * Carbamidomethyl (C))	38.02	38.00	38.10	0.04
3-4	[40- 46]	725.467	725.901	1	KPGKAPK	3.03	2.90	3.10	0.06
3-5	[40- 62]	2366.355	2367.77	2	KPGKAPKLLIYGATNLADGVPSR	37.85	37.50	38.50	6.25
4	[44- 46]	315.203	315.389	0	APK	2.62	2.40	2.70	0.13
5 isoD	[47- 62]	1659.906	1660.892	0	LLIYGATNLA isoDGVPSR	36.44	36.00	36.90	0.20
5	[47- 62]	1659.906	1660.892	0	LLIYGATNLADGVPSR	37.14	36.80	37.80	0.22
6 C* (gel)	[63-104]	4550.134	4552.922	0	FSGSGSGTDFTLTISSLQ.....YYCQNVLNTPL TFGQGTK (1 * Carbamidomethyl (C))	50.55	50.50	50.70	0.08
7	[105-108]	488.308	488.599	0	VEIK	10.63	9.70	11.20	0.42
7-8	[105-109]	644.409	644.785	1	VEIKR	10.48	9.90	11.50	0.62
8-9	[109-137]	2488.213	2489.597	1	RTGGGGSGGGGSGGGGSQVQLVQSGAEV K				
8-10	[109-144]	3155.615	3157.395	2	RTGGGGSGGGGSGGGGSQVQLVQSGAEVK KPGASVK				
9	[110-137]	2332.112	2333.411	0	TGGGGSGGGGSGGGGSQVQLVQSGAEVK	27.24	26.60	28.50	0.55
2 C*(gel) ox	[20- 39]	2332.103	2333.56	0	VTITCGASENIYGALNWYQR (1 * Carbamidomethyl (C))	40.35	39.71	41.61	0.57
9-10	[110-144]	2999.514	3001.209	1	TGGGGSGGGGSGGGGSQVQLVQSGAEVK KPGASVK	26.15	26.10	26.20	0.07
10	[138-144]	686.42	686.822	0	KPGASVK	4.44	3.30	4.90	0.65
13 PG	[164-190]	2909.33	2911.142	0	Digest Artifact QAPGQGLEWMGEILPGSGSTEYTENFK (1 * Pyro-glu (N-term peptide))	45.90	45.70	46.10	0.28
13-14 PG	[164-192]	3180.458	3182.416	1	Digest Artifact QAPGQGLEWMGEILPGSG.....EIM YTENFKDR (1 * Pyro-glu (N-term peptide))	44.67	44.40	44.80	0.11
1 Acet	[1- 19]	1991.934	1993.139	0	ADIQMTQSPSSLSASVGDR (1 * Acetyl (N-term))	34.03	34.00	34.10	0.05
2s-s17	[20-39s-s213-223]	3533.61	3535.857	0	VTITCGASENIYGALNWYQRSEDTAVYYCAR (Disulfide bonded, 1 * Additional C & N Term)	39.10	39.00	39.20	0.14
6s-s11	[63-104s-s145-148]	4926.312	4929.395	0	FSGSGSGTDFTLTISSLQ.....NVLNTPLTFGQ GTKVSCK (Disulfide bonded, 1 * Additional C & N Term)	46.75	46.60	46.90	0.15
11s-s11	[145-148s-s145-148]	869.422	870.073	0	VSCKVSCK (Disulfide bonded, 1 * Additional C & N Term)	6.80	6.30	7.30	0.71
11s-s17	[145-148s-s213-223]	1710.746	1711.896	0	VSCKSEDTAVYYCAR (Disulfide bonded, 1 * Additional C & N Term)	22.77	22.40	22.90	0.14
-	Assay Artifact	253	253.2	-	DTT Dimer	6.25	6.20	6.30	0.04
-	Assay Artifact	259	259.1	-	HQ Buffer, RapiGest, Trypsin Digest Assay Unknowns (RT~2.5 min.)	2.53	2.30	3.00	0.24
-	Assay Artifact	280.99	281.19	-	HQ Buffer, RapiGest, Trypsin Digest Assay Unknowns (RT~23.7 min.)	24.00	23.90	24.20	0.14
-	Assay Artifact	310.95	311.15	-	DTT Dimer + Alk	16.87	16.70	17.20	0.16

Isomers

Tryptic peptides ~ MW / Fragmentation

Modified & Miscleavages

Disulfide Linkages

Assay Artifacts

Alexion's – Master Table Output

Rise to Greatness™

Peak (#)	Ret Time (min)	m/z ()	charge (+)	MH+ (Da)	% Base Peak	Trypsin Frag (#)	Range	Comments: Sequence ($MH+_{mono}$: $MH+_{ave}$)
1	3.2	494.3	1	494.30	12.9	11 C* (sol)	[145-148]	VSCK (1 * Carboxymethyl (C)) (494.228 : 494.584)
2	4.8	343.75	2	686.49	12.6	10	[illegible]	[illegible]
3	10.1	322.75	2	644.49	5.2	7-8	[illegible]	[illegible]
4	10.6	488.4	1	488.40	2.4	7	[illegible]	[illegible]
5	11.2	607.45	1	607.45	2.8	15	[193-197]	VTMTR (607.323 : 607.746)
6	16.8	311.02	1	311.02	2.1	-	Assay Artifact	DTT Dimer + Alk (310.95 : 311.15)
7	18	665.5	1	665.50	1.0	15 N-Term CoMe	[193-197]	Digest Artifact VTMTR (1 * Carboxymethyl (N-term) (Sol)) (665.329 : 665.782)
8	21.4	440.15	1	440.15	3.3	-	Assay Artifact	[illegible]
[illegible]	[illegible]	[illegible]	1	447.08	1.0	-	Assay Artifact	[illegible] 7.05)
[illegible]	[illegible]	[illegible]	3	1711.97	2.3	11s-s17	[145-148s-s213-223]	VSCKSEDTAVYYCAR (Disulfide bonded, 1 * Additional C & N Term) (1710.746 : 1711.896)
11	25.6	668.65	2	1336.29	76.9	17 C* (sol)	[213-223]	SEDTAVYYCAR (1 * Carboxymethyl (C)) (1335.552 : 1336.407)
12	26	830.61	3	2489.81	2.9	8-9	[109-137]	RTGGGGSGGGGSGGGGSQVQLVQSGAEVK (2488.213 : 2489.597)
13	26.7	670.36	3	2009.06	16.6	1 N-Term CoMe	[1- 19]	Digest Artifact ADIQMTQSPSSLSASVGDR (1 * Carboxymethyl (N-term) (Sol)) (2007.929 : 2009.138)
[illegible]	[illegible]	[illegible]	3	2333.39	56.8	9	[110-137]	TGGGGSGGGGSGGGGSQVQLVQSGAEVK (2332.112 : 2333.411)
15	27.8	797.85	3	2391.53	4.4	9 N-Term or K CoMe	[110-137]	[illegible] (N-term or K) (Sol)) (
16	31.2	651.09	3	1951.25	65.3	1	[1- 19]	ADIQMTQSPSSLSASVGDR (1949.923 : 1951.102)
17	34	665.15	3	1993.43	1.8	1 Acet	[1- 19]	ADIQMTQSPSSLSASVGDR (1 * Acetyl (N-term)) (1991.934 : 1993.139)
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]4	3-5 : 11-12 C* (sol)	[40- 62] : [145-163]	KPGKAPKLLIYGATNLADGVPSR (2366.355 : 2367.77) : VSCKASGYIFSNYWIQWVR (1 * Carboxymethyl (C)) (2365.143 : 2366.677)
19	36.2	554.36	3	1661.06	3.0	5 isoD	[47- 62]	LLIYGATNLAisoDGVPSR (1659.906 : 1660.892)
20	37.1	554.37	3	1661.09	100.0	5	[47- 62]	LLIYGATNLADGVPSR (1659.906 : 1660.892)
21	38	605.99	5	3025.92	8.5	2-4 C* (sol)	[20- 46]	VTITCGASENIYGALNWYQRKPGKAPK (1 * Carboxymethyl (C)) (3023.541 : 3025.424)
[illegible]	[illegible]	[illegible]	3	1690.91	57.7	16	[198-212]	DTSTSTVYMELSSLR (1689.8 : 1690.851)
[illegible]	[illegible]	[illegible]	3	2318.39	8.2	2 C* (sol)	[20- 39]	VTITCGASENIYGALNWYQR (1 * Carboxymethyl (C)) (2317.092 : 2318.546)
24	42.4	800.75	4	3199.90	28.5	13-14	[164-192]	QAPGQGLEWMGEILPGSGSTEYTENFKDR (3197.484 : 3199.446)
25	44	976.68	3	2928.02	6.2	13	[164-190]	QAPGQGLEWMGEILPGSGSTEYTENFK (2926.356 : 2928.173)
26	44.7	1061.40	3	3182.18	1.4	13-14 PG	[164-192]	Digest Artifact QAPGQGLEWMGEILPGSG.....EILPGSGSTEYTENFKDR (1 * Pyro-glu (N-term peptide)) (3180.458 : 3182.416)
[illegible]	[illegible]	[illegible]	3	1891.43	57.0	12	[149-163]	ASGYIFSNYWIQWVR (1889.933 : 1891.115)
[illegible]	[illegible]	[illegible]	4	4553.78	23.9	6 C* (sol)	[63-104]	FSGSGS[illegible] (C)) (4551.118 : 4553.906)
29	50.9	853.75	8	6822.94	4.4	Unknown		
30	51.4	916.01	3	2746.01	10.8	18	[224-247]	YFFGSSPNWYFDVWGQGTLVTVSS (2744.267 : 2745.973)
31	54.7	1200.73	4	4799.90	1.2	12-13	[149-190]	ASGYIFSNYWIQWVRQAPGQGLEWMGEILPGSGSTEYTENFK (4797.272 : 4800.265)

Green: Mass, RT, Fragmentation

White: Assay Artifact

Yellow: Mass and/or RT

Red: Unknown

Cross-Links

Modifications

Different peptides ~MW

Isomerization

Mis-cleavages

Conclusions

- Protein Characterization Goals are for 100% Primary Sequence Coverage

- Alexion's Biotherapeutic Characterization group uses additional strategies

  - Branched Sample Prep / Composite Search MALDI-ToF
  - Sub-Ambient Trapping LC-MS
  - MasterTable for Data Reduction

Acknowledgments

Waters Contacts

Jeff Mazzeo
Scott Berger
Ken Eglinton
Ying Qing Yu
Dustin Yaworsky

Alexion Pharmaceuticals

Bruce Andrien
Rekha Patel
Christine Nowak
Helena Prieto

Examination of an a Biotherapeutic Characterization Strategy.

GEN Webinar Series 2009

Adam Lucka, Rekha Patel, Christine Nowak, and Bruce Andrien
Alexion Pharmaceuticals
Protein Characterization Group

Waters
THE SCIENCE OF WHAT'S POSSIBLE.™

Integration of Bioinformatics for Biotherapeutic Protein Characterization

Yan Wang, Ph.D
Analytical Biochemistry Department
Biogen Idec, Cambridge

1. Introduction

- Workflow for protein characterization for Biotherapeutic proteins
- Bioinformatics tools for protein identification
- Bioinformatics tools for protein characterization

2. Examples of protein characterization with bioinformatics

- Intact mass analysis
- Peptide mapping
- Disulfide bond mapping
- MS^E analysis
- MS/MS analysis

3. Conclusion

MS-based Protein Characterization for Biotherapeutic Proteins (of known protein sequence)

- Intact mass analysis
  - Does the detected mass match the theoretical mass?
  - Are there other species present besides the expected one?
- Peptide mapping
  - Do peptides conform to the predicted sequence? Any mutations involved?
  - What post-translational modifications are involved?
  - Are disulfide bonds in the proper arrangement?
- Tandem mass analysis
  -Identify mutations and post-translational modification, sequence unknown peptides, determine sites of mutation and post-translational modifications

Extensive Bioinformatics Tools for Protein Identification (Proteomics)

Identification by Peptide Mass Fingerprint

MS-Fit: (UCSF)
http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm

Profound: (Rockefeller University)
http://129.85.19.192/profound_bin/WebProFound.exe?FORM=1

Mascot PMF: (Matrixscience)
http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF

Identification by Tandem MS/MS

SEQUEST: (ThermoElectron)
http://fields.scripps.edu/sequest/

MASCOT: (Matrixscience)
http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=MIS

OMSSA:
NCBI--Open Mass Spectrometry Search Algorithm
http://pubchem.ncbi.nlm.nih.gov/omssa/omssacgi.cgi?searchsettings=iontrap.xml

Bioinformatics Tools for Biotherapeutic Protein Characterization

- **Protein characterization works with relatively pure molecules having known sequences**
- **Analytical data processing is the bottleneck of protein characterization**
- **Bioinformatics tools are less developed**

Data Processing Softwares:

1. Software by Agilent

- Mass Hunter matches protein sequences and finds post-translational modifications; Useful for biotherapeutic protein characterization
- Mass Profiler is good for comparing the statistical significance of certain m/z values and for quantitative analysis.

2. Software by Waters

- Useful for biotherapeutic protein characterization
- Can process MSe data to confirm peptide sequences
- Can show the peak intensity difference, but no statistical visualization
- user friendly

Bioinformatics Can Greatly Increase Data Analysis Productivity

No Bioinformatics

Manually match m/z values of every spectrum across the chromatogram with the theoretical peptides

Identify the matched m/z in each spectrum,
assign the annotation

Perform ms/ms analysis for unknown m/z values

Manually input all the identified peptide sequences, m/z values, charge states, post-translational modifications….

Months of work

With Bioinformatics

Automate batch data processing, annotate and compare the results among reference standard and experimental samples

Process MSe data to validate peptide mapping assignment

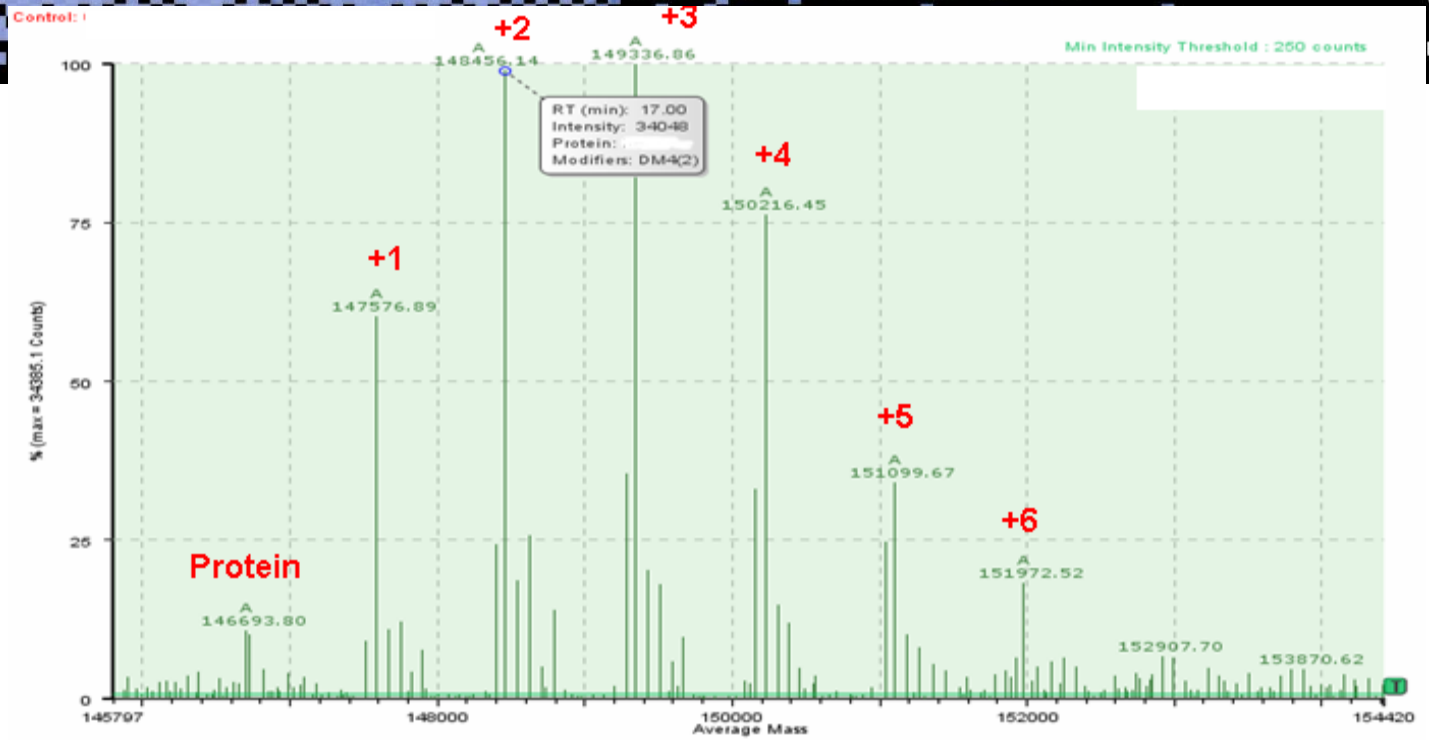
Produces annotated spectra, chromatograms, coverage maps and tabular data

Export peptide/fragment data for searching with other bioinformatics tools

Weeks to days

Intact Mass Analysis

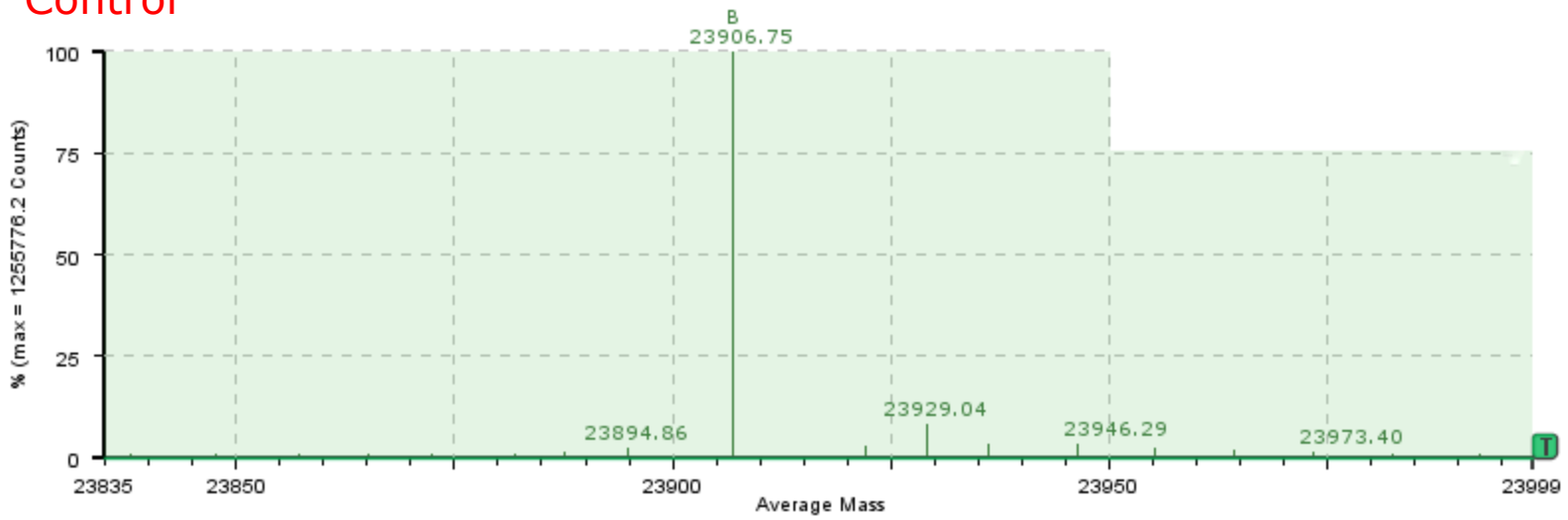
Example 1: Single sample analysis showed protein and its conjugates



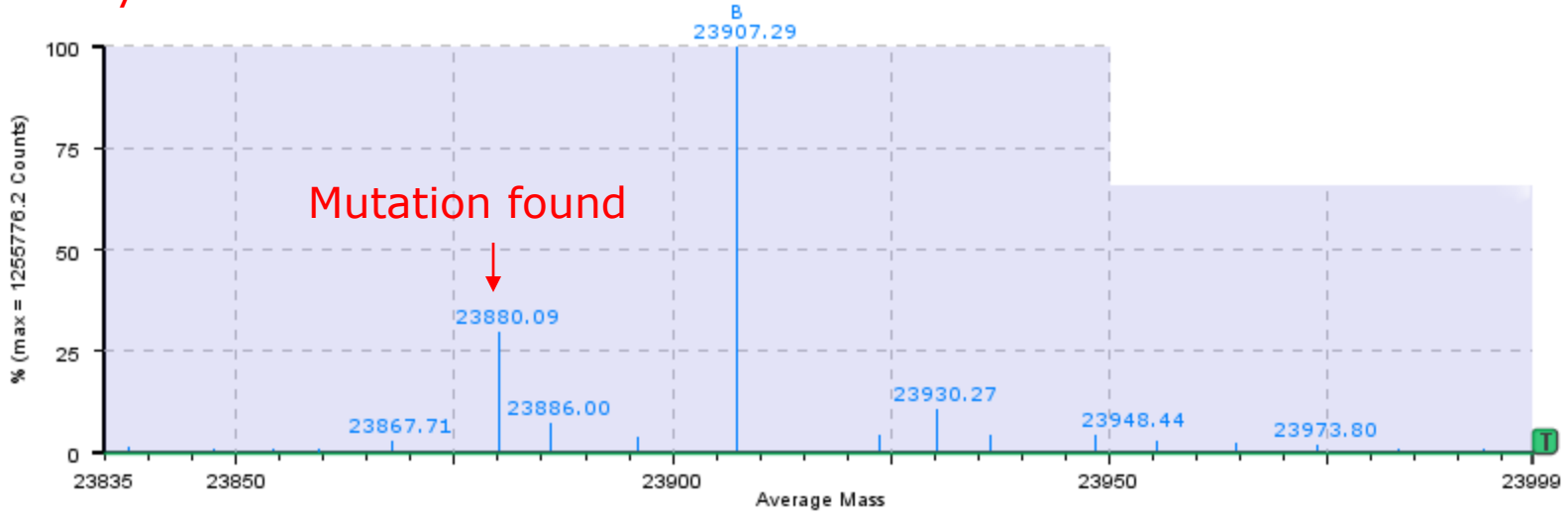
	Modifiers	▼1 Calculated Protein Mass (Da)	Control RT (Min)	Control Mass (Da)	Control Intensity (Counts)
1	DM4(9)	154612.9531	17.0	154622.2344	344.9
2	DM4(7)	152852.9531	17.0	152846.0938	1325.3
3	DM4(6)	151972.9531	17.0	151972.5156	6286.0
4	DM4(5)	151092.9531	17.0	151099.6719	11700.1
5	DM4(4)	150212.9531	17.0	150216.4531	26240.6
6	DM4(3)	149332.9531	17.0	149336.8594	34385.1
7	DM4(2)	148452.9531	17.0	148456.1406	34047.6
8	DM4(1)	147572.9531	17.0	147576.8906	20736.7
9		146692.9531	17.0	146693.7969	3725.2

Example 2 : Batch samples can be processed and compared for high-throughput

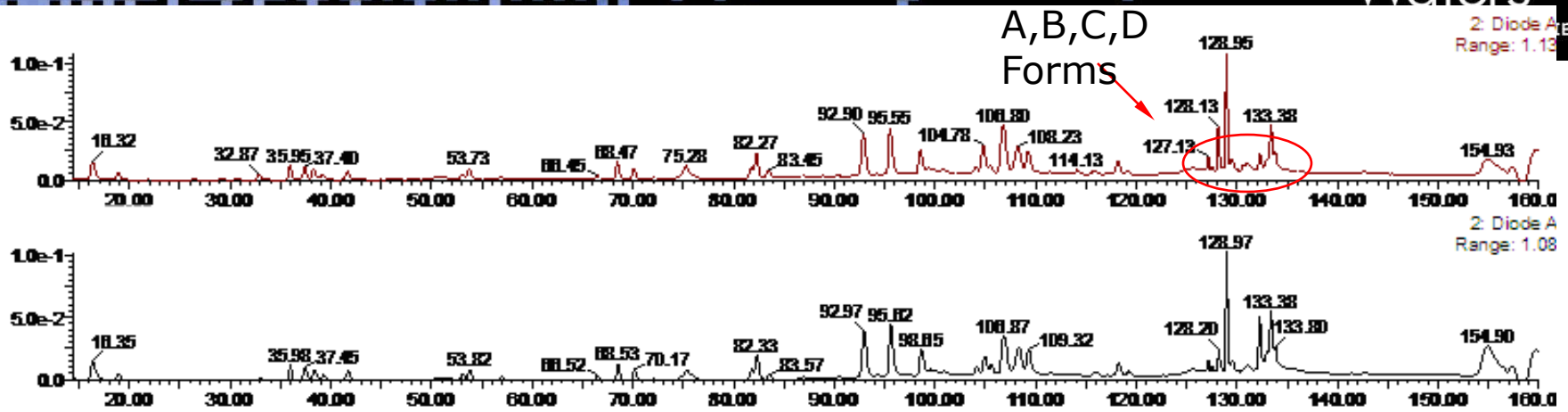
Control

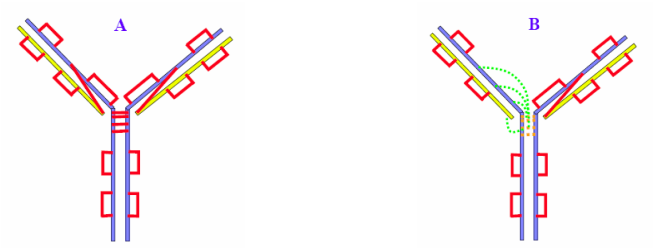


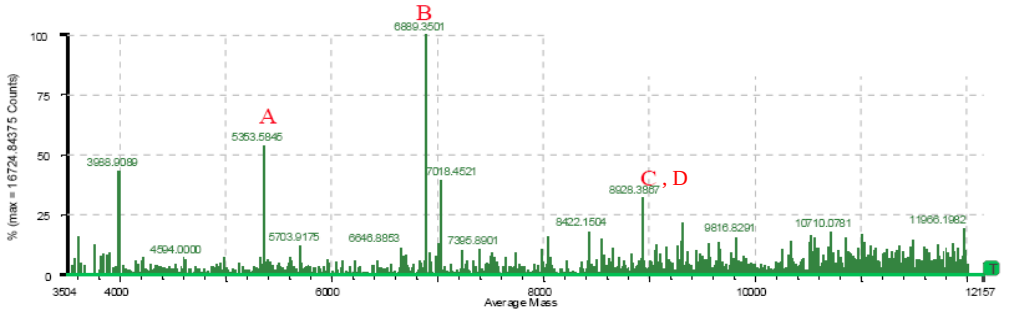
Analyte

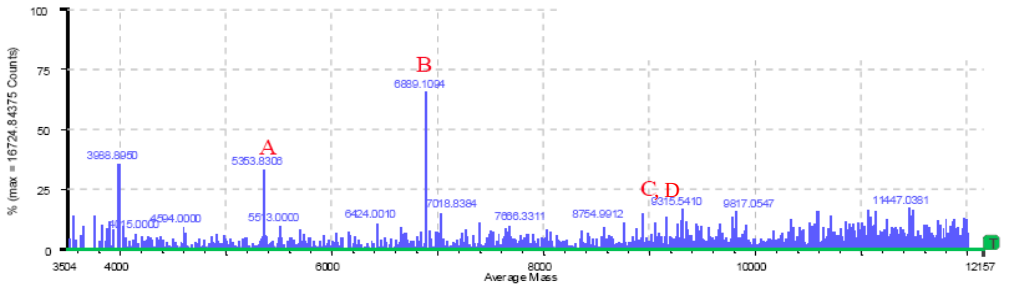


Example 3 : IgG2 antibody with three different disulfide-bond structures can be resolved



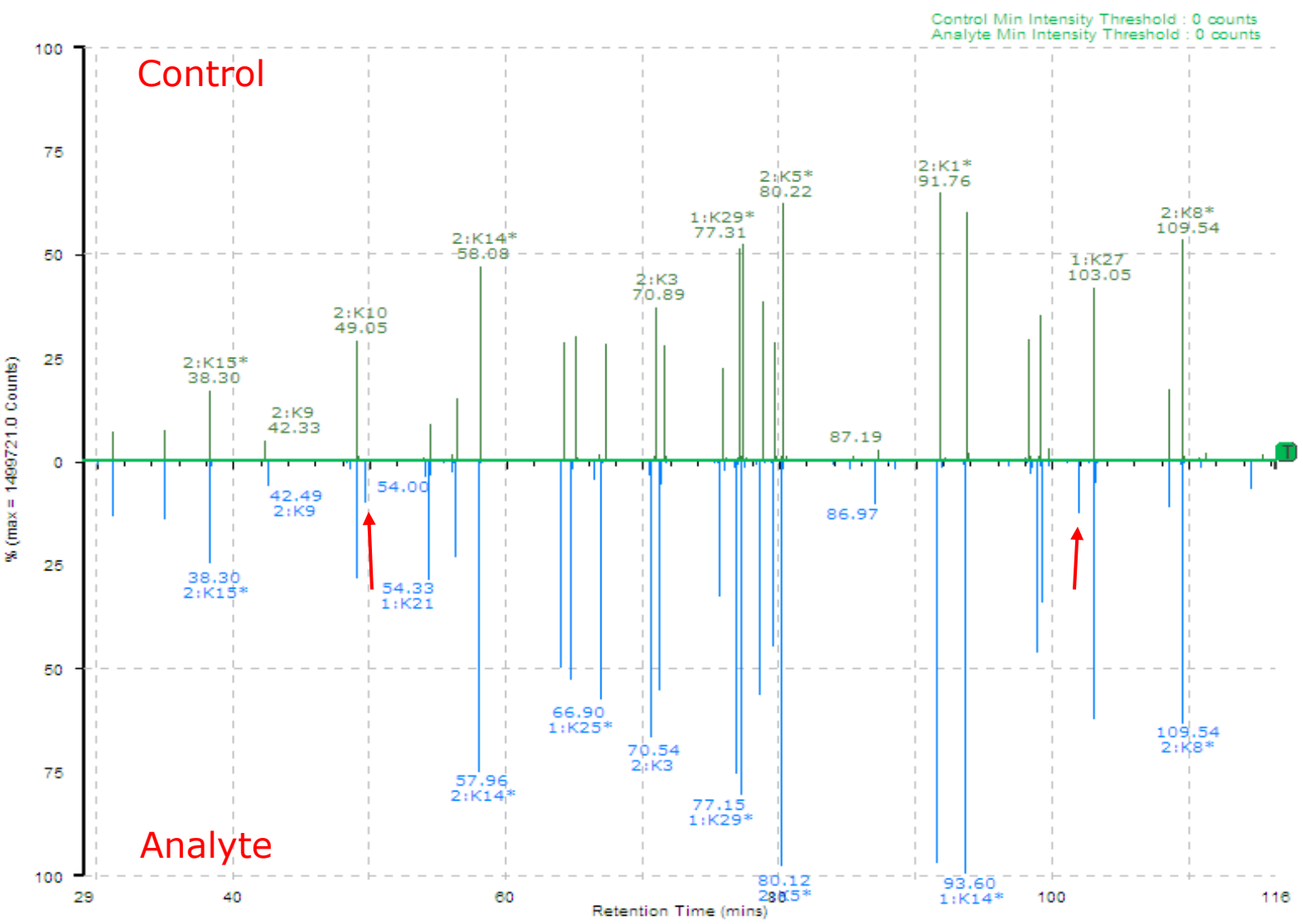






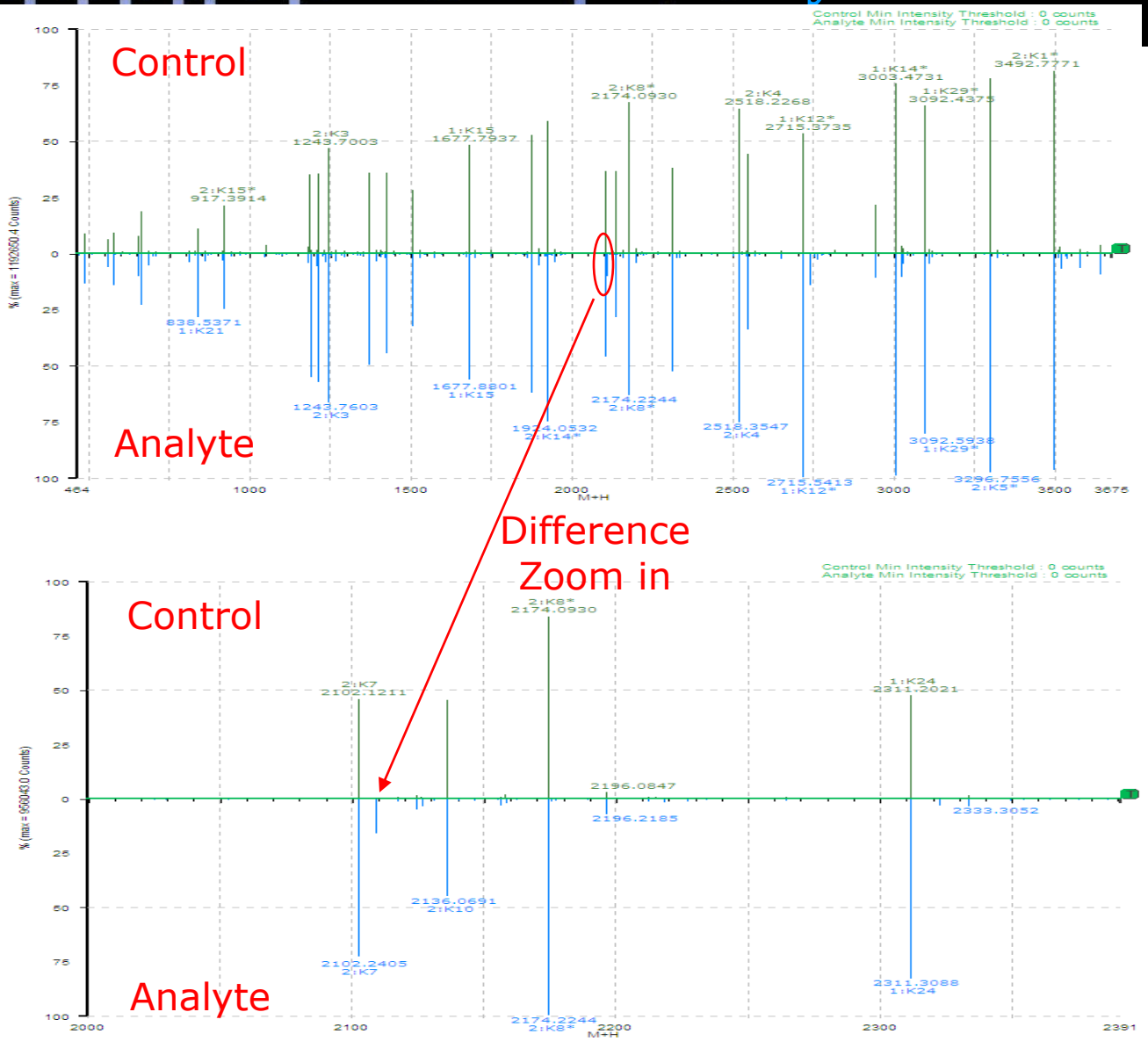
Peptide Mapping Data Processing

Processed Chromatogram Showed Differences between Control vs Analyte



Difference can be Observed in Spectrum:

Control *vs* Analyte



Comparison are Summarized in Tabular Data

Control *vs* Analyte

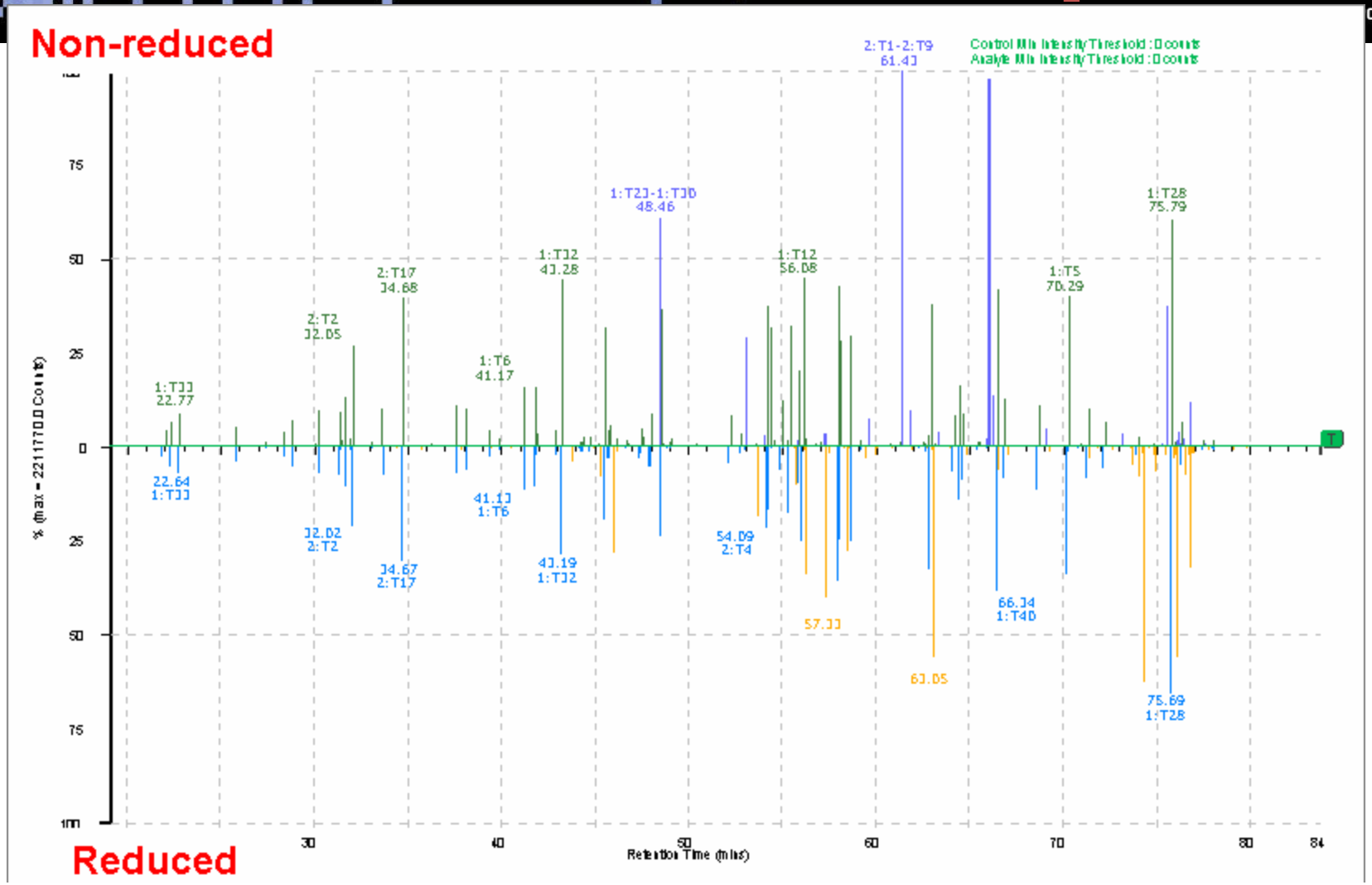
Chromatogram | Spectrum | Coverage Map | Protein Digests | Peak Match Data

	Fragment Number	Start	End	Calculated ...	Control RT...	Control m/z	Co...	Control Mass (Da)	Control Inten...	▲1 An...	Analyte m/z	A...	Analyte Mass ...	Analyte Intensi...	Delta...
1	2:K6	109	112	487.3006	31.1	488.3078	1	487.2999	107309.0	31.4	488.3298	1	487.3219	119451.3	10.2
2	1:K28/y4	419	422	461.2486	34.9	462.2583	1	461.2504	33426.0	35.1	462.2788	1	461.2709	25914.7	29.0
3	1:K28	418	422	574.3326	34.9	575.3392	1	574.3313	110118.0	35.1	575.3642	1	574.3563	124050.1	11.2
4	2:K15*/y6*	214	219	829.3541	38.3	415.6849	2	829.3540	13896.0	38.3	415.7037	2	829.3915	10847.4	28.1
5	2:K15*	213	219	916.3862	38.3	459.1997	2	916.3836	256552.0	38.3	459.2220	2	916.4281	215683.9	18.9
6					38.4	939.3770	1	938.3691	10097.0	38.4	939.4203	1	938.4124	13350.0	24.4
7	2:K9	151	154	559.3118	42.3	560.3181	1	559.3102	75155.0	42.4	560.3428	1	559.3349	55926.1	34.4
8										48.5	787.4412	2	1572.8665	20703.4	
9	2:K10x2	155	174	4269.9229	49.0	1424.3254	3	4269.9526	28803.0	49.0	1424.3901	3	4270.1465	17204.4	67.4
10	2:K10	155	174	2134.9614	49.1	1068.4880	2	2134.9602	436132.0	49.1	1068.5386	2	2135.0613	247822.4	76.0
11	2:K8*	132	150	2157.0566	49.1	1079.4825	2	2156.9492	17648.0	49.1	1079.5333	2	2157.0508	13305.8	32.6
12	2:K10/y6	169	174	706.3134						49.1	707.3522	1	706.3443	12904.6	
13	2:K10 155-174 mutant									49.6	1055.0325	2	2108.0491	89437.6	
14					53.9	404.7409	2	807.4659	13147.0	54.1	808.5115	1	807.5036	36418.8	63.9
15	1:K21/y6	337	342	653.3748	54.4	654.3783	1	653.3704	95390.0	54.4	654.4099	1	653.4020	90391.6	5.5
16	1:K21	335	342	837.4960	54.4	838.4990	1	837.4911	134017.0	54.4	838.5372	1	837.5293	249582.3	46.3
17					54.4	860.4831	1	859.4752	14944.0	54.5	860.5236	1	859.5157	33845.1	55.8
18					56.1	432.2263	1	431.2184	24704.0	56.1	432.2457	1	431.2378	27374.4	9.8
19	2:K6	109	112	487.3006						56.4	488.2944	1	487.2865	12240.4	
20	1:K30	448	454	659.3490	56.3	660.3523	1	659.3444	225395.0	56.4	660.3835	1	659.3756	202608.2	11.2
21					56.4	682.3373	1	681.3294	13375.0	56.4	682.3674	1	681.3595	49072.6	72.7

Disulfide Mapping: Comparison of the chromatograms of non-reduced and reduced samples

Waters

OSSIBLE.™

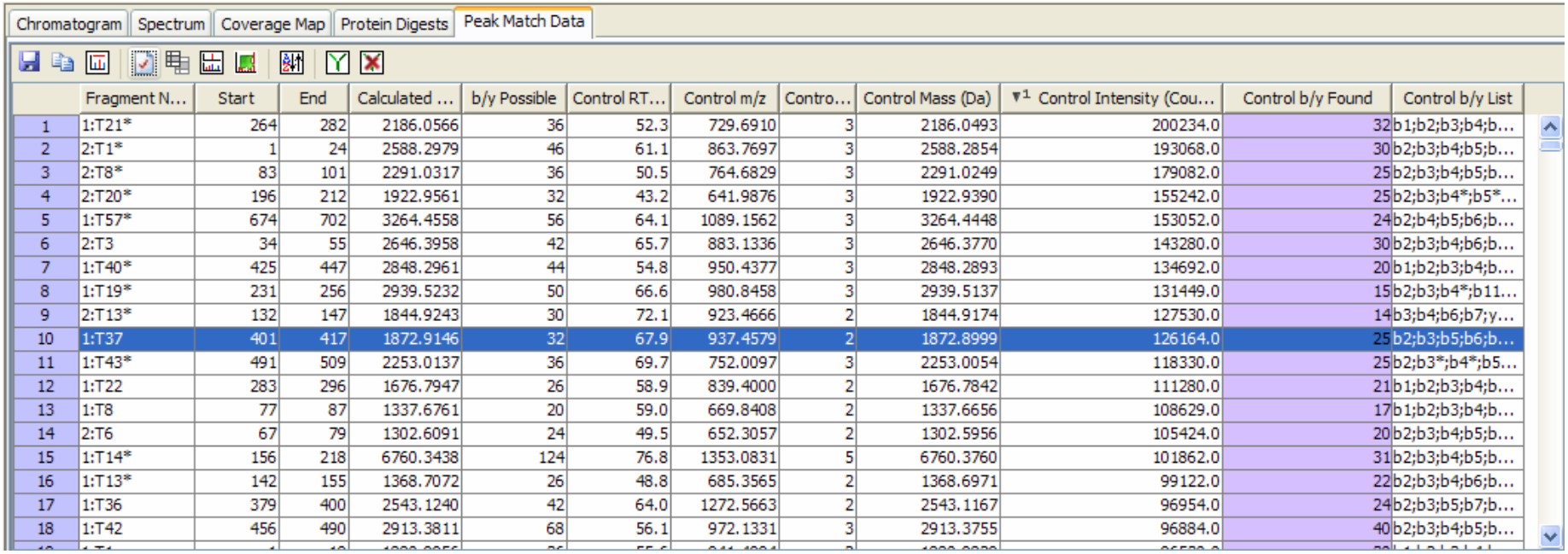


Disulfide-linked Peptides in the Map of Non-reduced Protein

Fragment Number	Modifiers	Calculated Peptide Mass (Da)	Control RT (Min)	Control m/z	Control Charge State	Control Mass (Da)	Control Intensity (Counts)
3:T001-3:T009		4667.198	59.7	1167.7939	4	4667.144	565863
3:T001-3:T009*	Oxidation M(1)	4683.193	58.5	1171.7952	4	4683.1489	196462
1:T002-1:T009		2504.11	51	835.7098	3	2504.1055	186185
3:T014-3:T021		3555.749	63.6	1186.2482	3	3555.7207	759713
3:T014-015-3:T021		3883.924	57.1	1295.5906	3	3883.748	1912
1:T019-3:T022-023		1260.486	27.1	631.2525	2	1260.4891	17914
1:T020-021-2:T020-021		5454.783	74.1	1364.6843	4	5454.7056	271024
1:T020-2:T020-021		5229.636	75.2	1308.3971	4	5229.5566	44785
1:T023-1:T030		2328.098	46.7	777.0377	3	2328.0894	386328
1:T023-024-1:T030		3986.882	16	1994.2504	2	3986.4849	12200
1:T023-1:T029-030		2748.299	15.1	1375.1942	2	2748.3726	8846
1:T038-1:T043		3844.824	57.1	1282.6047	3	3844.7903	207856
1:T038-1:T042-043		4087.957	56.5	1363.6487	3	4087.9221	12174
1:T046-1:T052		3378.478	63.7	1127.1567	3	3378.4463	98158
1:T046-1:T052*	Oxidation M(1)	3394.473	59.7	1132.491	3	3394.449	458
1:T054-1:T061		3848.751	60	1283.9138	3	3848.7178	55271

Caution: False positives can be found in the assignment!

MSe Data Processing to Validate Peptide Mapping Assignment



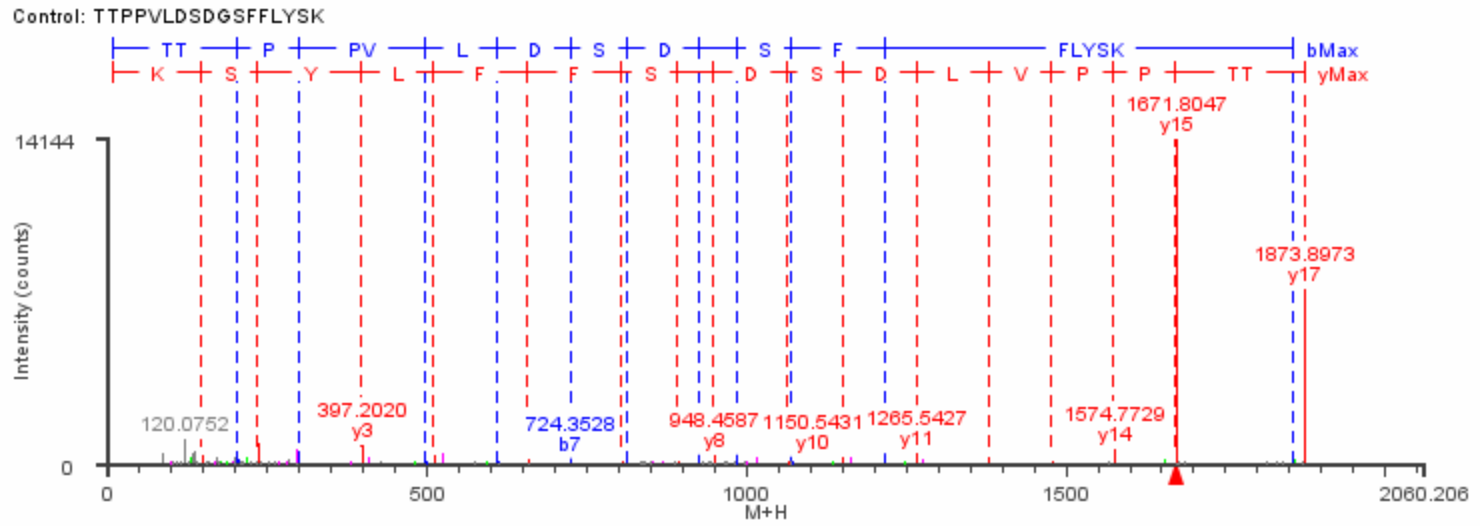
Chromatogram | Spectrum | Coverage Map | Protein Digests | Peak Match Data

	Fragment N...	Start	End	Calculated ...	b/y Possible	Control RT...	Control m/z	Contro...	Control Mass (Da)	▼1 Control Intensity (Cou...	Control b/y Found	Control b/y List
1	1:T21*	264	282	2186.0566	36	52.3	729.6910	3	2186.0493	200234.0	32	b1;b2;b3;b4;b...
2	2:T1*	1	24	2588.2979	46	61.1	863.7697	3	2588.2854	193068.0	30	b2;b3;b4;b5;b...
3	2:T8*	83	101	2291.0317	36	50.5	764.6829	3	2291.0249	179082.0	25	b2;b3;b4;b5;b...
4	2:T20*	196	212	1922.9561	32	43.2	641.9876	3	1922.9390	155242.0	25	b2;b3;b4*;b5*...
5	1:T57*	674	702	3264.4558	56	64.1	1089.1562	3	3264.4448	153052.0	24	b2;b4;b5;b6;b...
6	2:T3	34	55	2646.3958	42	65.7	883.1336	3	2646.3770	143280.0	30	b2;b3;b4;b6;b...
7	1:T40*	425	447	2848.2961	44	54.8	950.4377	3	2848.2893	134692.0	20	b1;b2;b3;b4;b...
8	1:T19*	231	256	2939.5232	50	66.6	980.8458	3	2939.5137	131449.0	15	b2;b3;b4*;b11...
9	2:T13*	132	147	1844.9243	30	72.1	923.4666	2	1844.9174	127530.0	14	b3;b4;b6;b7;y...
10	1:T37	401	417	1872.9146	32	67.9	937.4579	2	1872.8999	126164.0	25	b2;b3;b5;b6;b...
11	1:T43*	491	509	2253.0137	36	69.7	752.0097	3	2253.0054	118330.0	25	b2;b3*;b4*;b5...
12	1:T22	283	296	1676.7947	26	58.9	839.4000	2	1676.7842	111280.0	21	b1;b2;b3;b4;b...
13	1:T8	77	87	1337.6761	20	59.0	669.8408	2	1337.6656	108629.0	17	b1;b2;b3;b4;b...
14	2:T6	67	79	1302.6091	24	49.5	652.3057	2	1302.5956	105424.0	20	b2;b3;b4;b5;b...
15	1:T14*	156	218	6760.3438	124	76.8	1353.0831	5	6760.3760	101862.0	31	b2;b3;b4;b5;b...
16	1:T13*	142	155	1368.7072	26	48.8	685.3565	2	1368.6971	99122.0	22	b2;b3;b4;b6;b...
17	1:T36	379	400	2543.1240	42	64.0	1272.5663	2	2543.1167	96954.0	24	b2;b3;b5;b7;b...
18	1:T42	456	490	2913.3811	68	56.1	972.1331	3	2913.3755	96884.0	40	b2;b3;b4;b5;b...

MS^E Sequencing Output

Sequence: TTPPVLDSDGSFFLYSK Fragment Number: 1:T37 Modifiers:

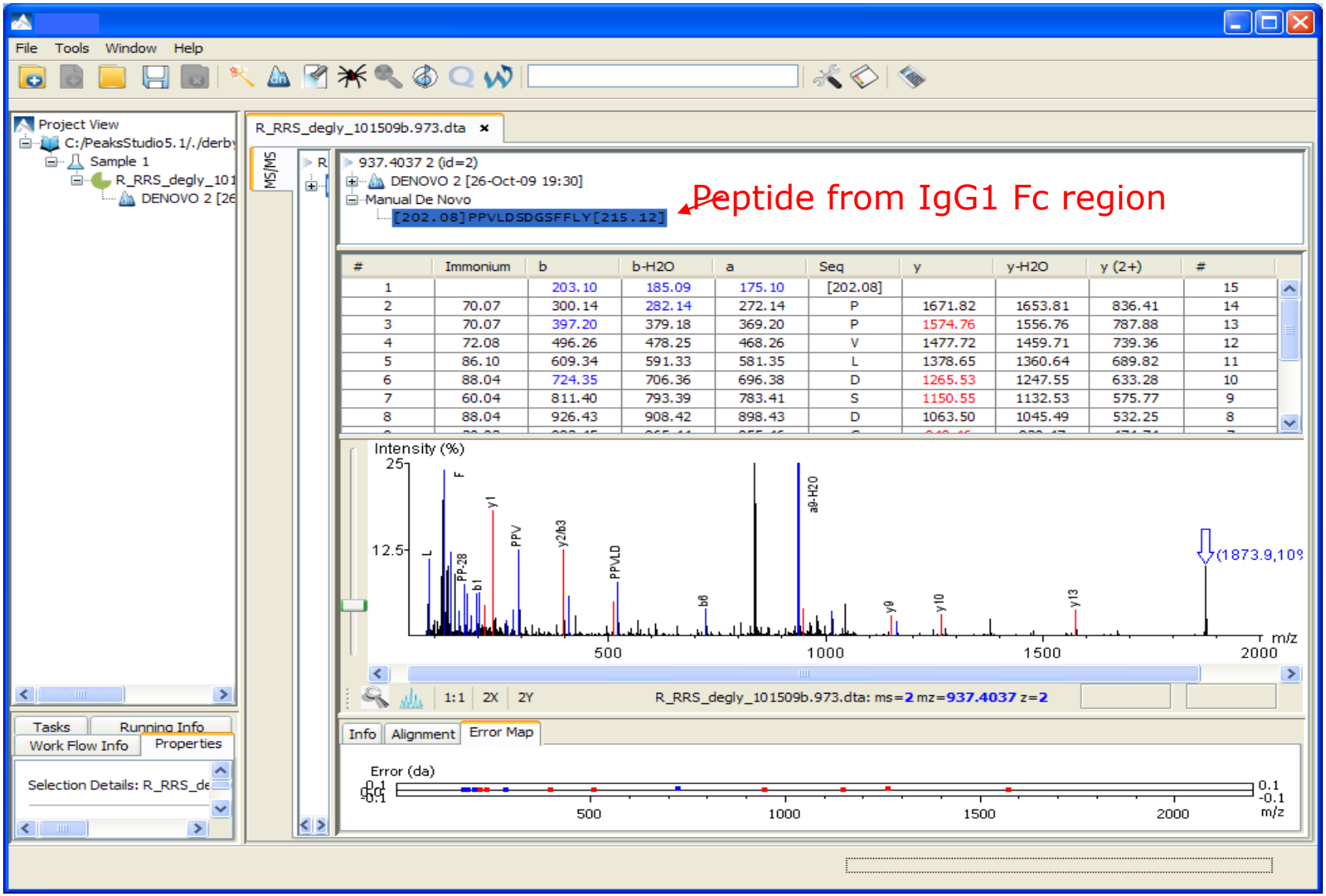
Peptide from IgG1 Fc region



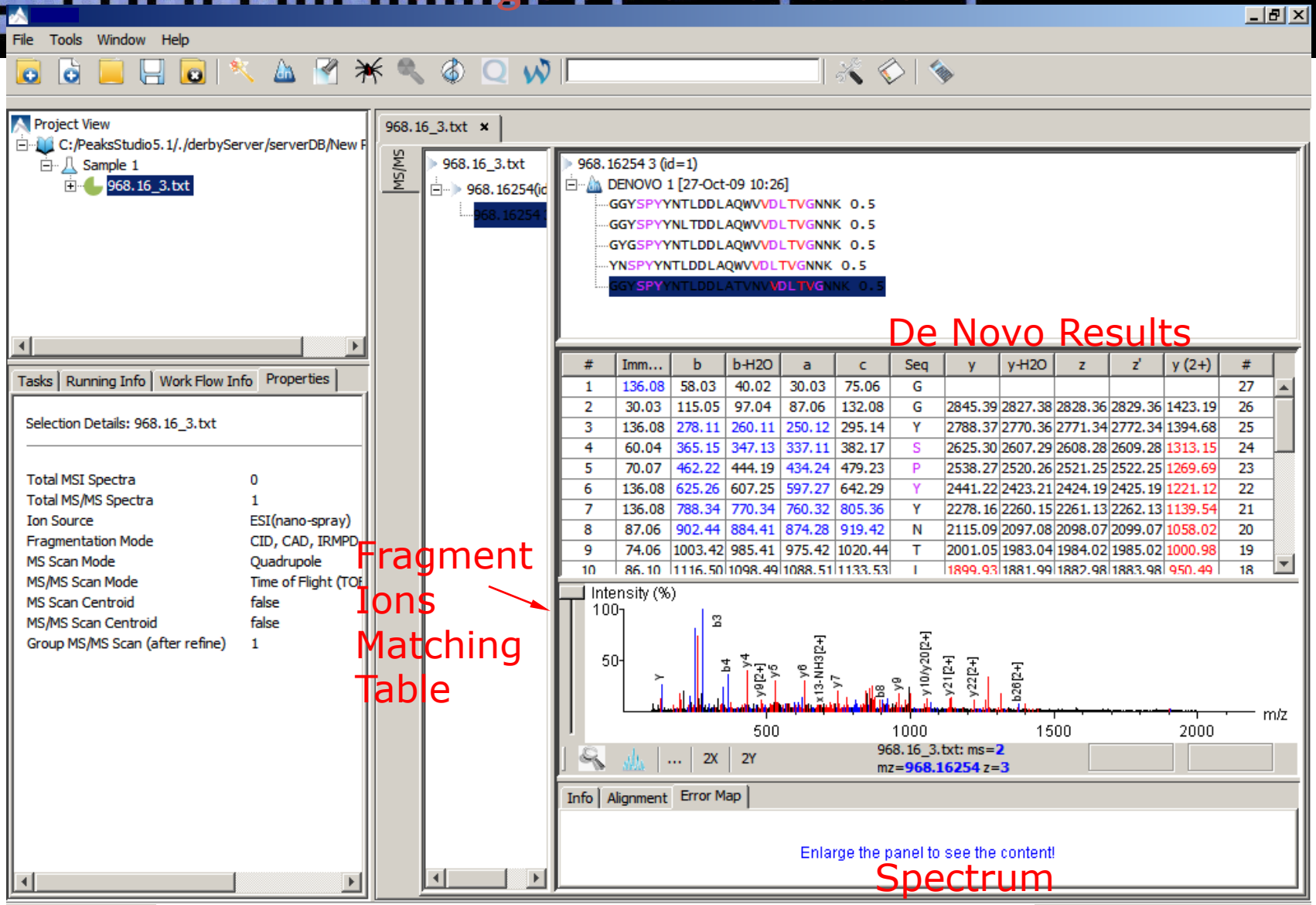
Assignment	Fragment Mass (Da)	Control Peak Mass (Da)	Control Mass Error (Da)	Control RT (mins)	▼1 Control Intensity (cou...	Cor
y15	1671.8270	1671.8047	0.0223	67.9	14144.0	
y17	1873.9224	1873.8973	0.0250	67.9	7644.0	
y2	234.1454	234.1382	0.0072	67.9	992.0	
y3	397.2087	397.2020	0.0067	67.9	860.0	
y14	1574.7742	1574.7729	0.0012	67.9	697.0	
y11	1265.5690	1265.5427	0.0262	67.9	510.0	
y8	948.4831	948.4587	0.0244	67.9	448.0	
y1	147.1133	147.1081	0.0052	67.9	418.0	

Export MS or MS^E Data for Use with Other Bioinformatics Tool

de novo sequencing for exported MS^E data using another bioinformatics tool



Another Example: MS/MS for *de novo* sequencing using bioinformatics tool



SPIDER Search Identified as Cytokeratin peptide

Conclusions

- Integration of bioinformatics can dramatically increase the efficiency of biotherapeutic protein characterization

- Caution still needs to be taken when using the bioinformatics tools. Expertise is required for in-depth data analysis.

Acknowledgements

- Dr. Dingyi Wen
- Dr. Sheng Gu
- Susan Foley
- Monika Vecchi
- Craig Wildes
- Yaping Sun

- Dr. Alex Buko

Overcoming LC/MS Data Overload in Biotherapeutic Peptide Mapping

Q&A

Thank you for attending

Overcoming LC/MS Data Overload in Biotherapeutic Peptide Mapping

Sponsored by

Waters

THE SCIENCE OF WHAT'S POSSIBLE.™