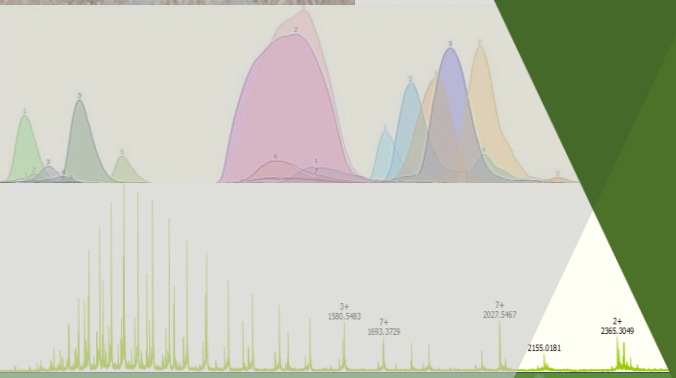
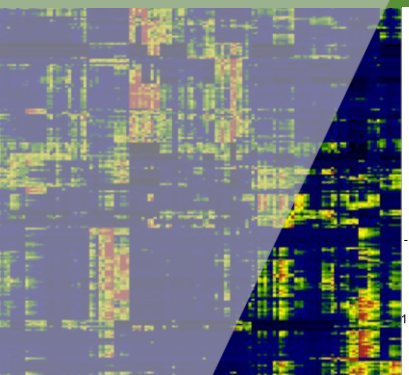
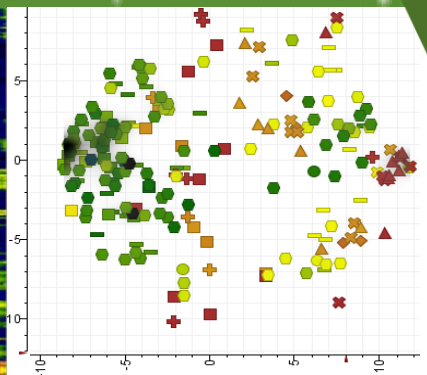






Milk top-down proteomics

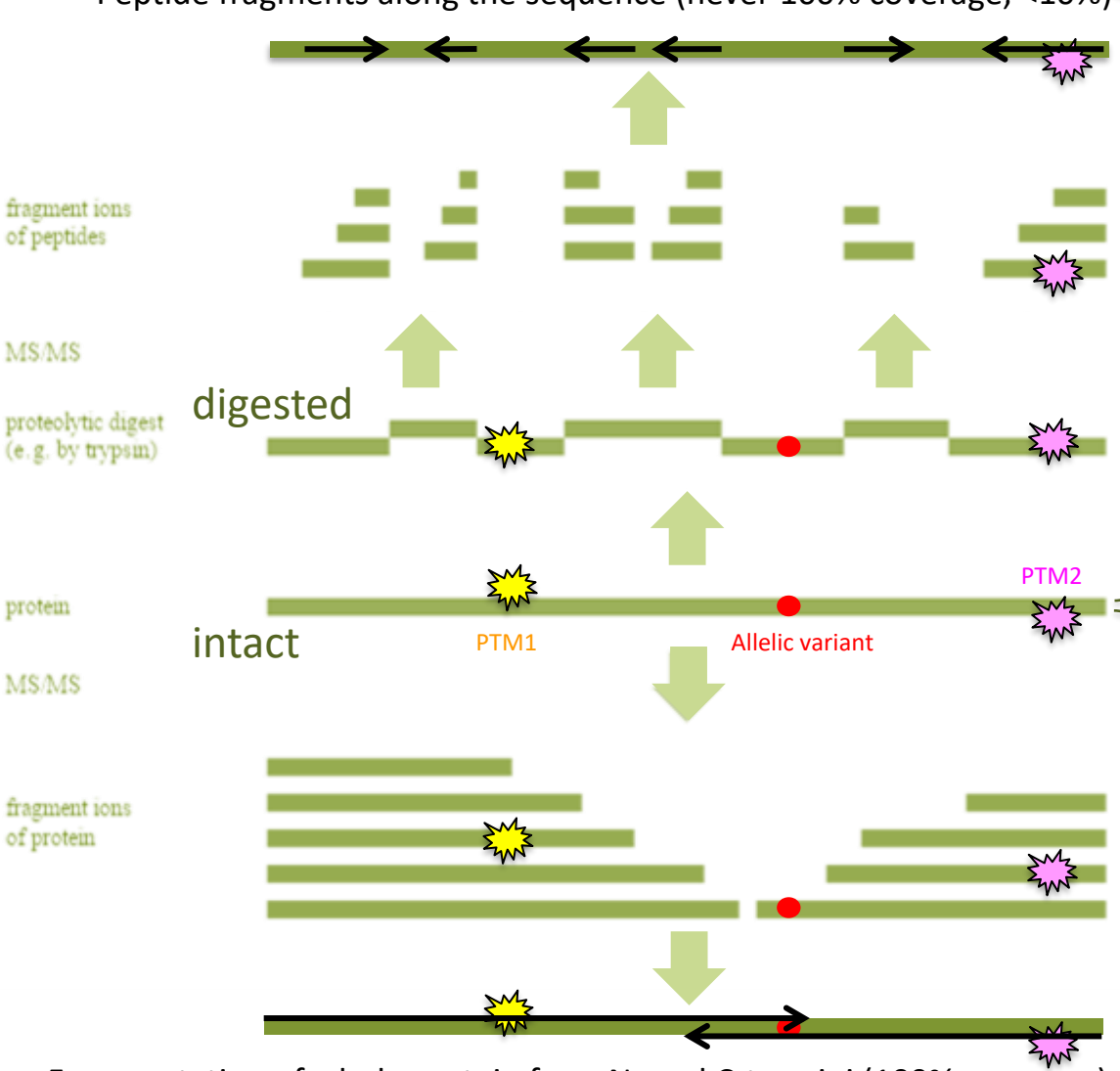
Dr Delphine Vincent

10 Mar 2018

Top-down and bottom-up proteomics

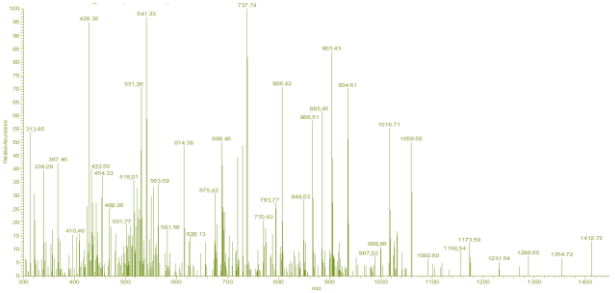
Highly complementary

Peptide fragments along the sequence (never 100% coverage, <10%)



Fragmentation of whole protein from N- and C-termini (100% coverage)

Top-down sequencing (TDS)



Bottom-up proteomics (peptides)

- \+ all MWs
- \- Dynamic range
- \+ Most mature
- \+ Widely used, sufficient tools
- \+ Can handle complex samples
- \- PTM analysis
- \- Protein processing
- \- Allelic variation
- \- Protein quantitation

up

bottom

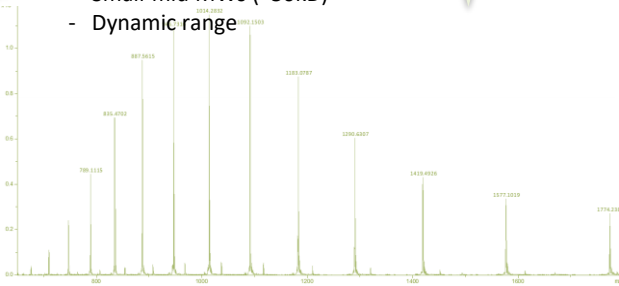
Top-down proteomics (proteins)

- \+ Protein quantitation
- \+ Allelic variation
- \+ PTM analysis on protein level
- \+ Protein processing
- \- Relatively new approach
- \- Complexity of the sample
- \- Small-mid MWs (<50kD)
- \- Dynamic range

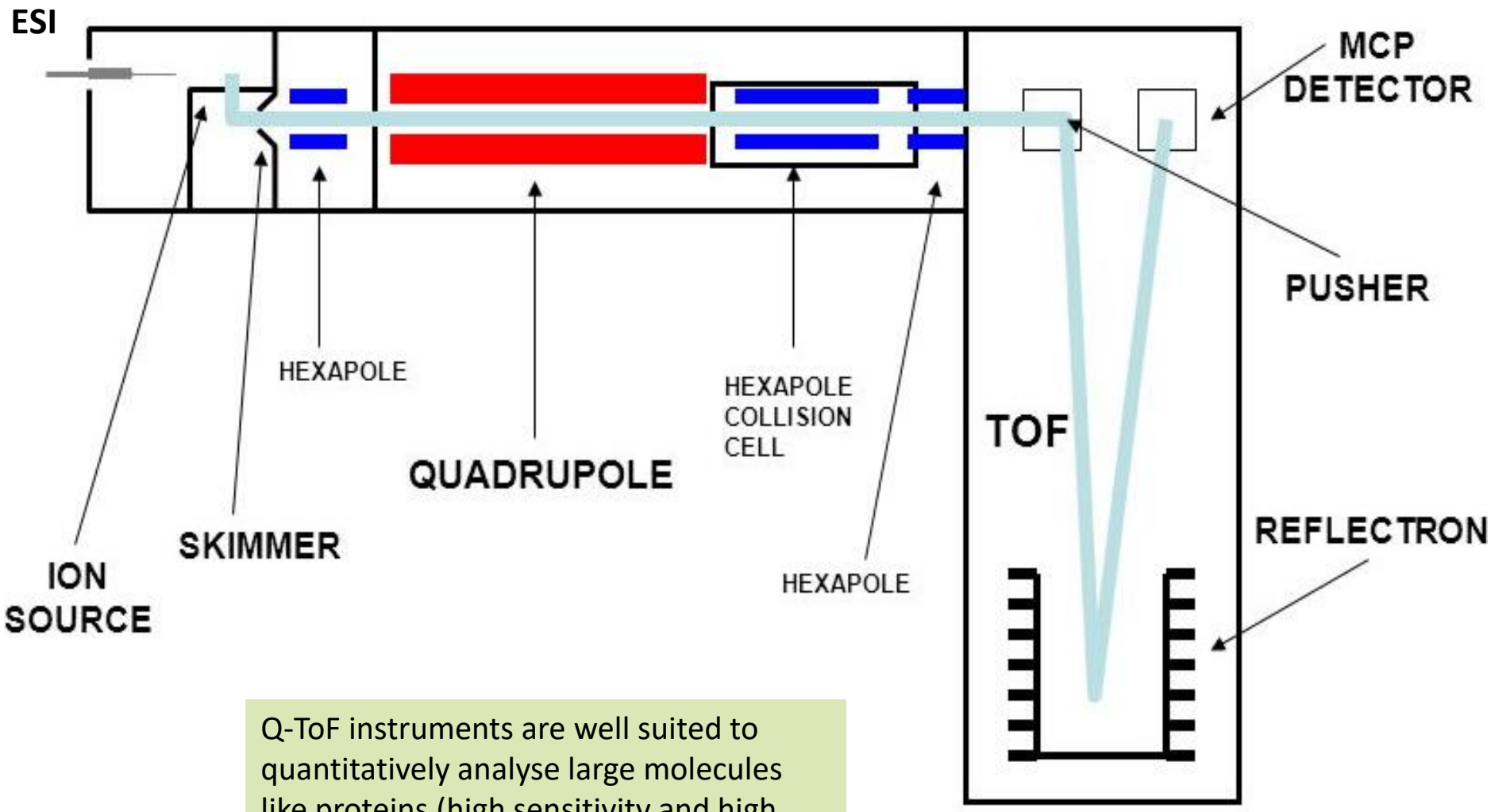
top

proteoforms

down

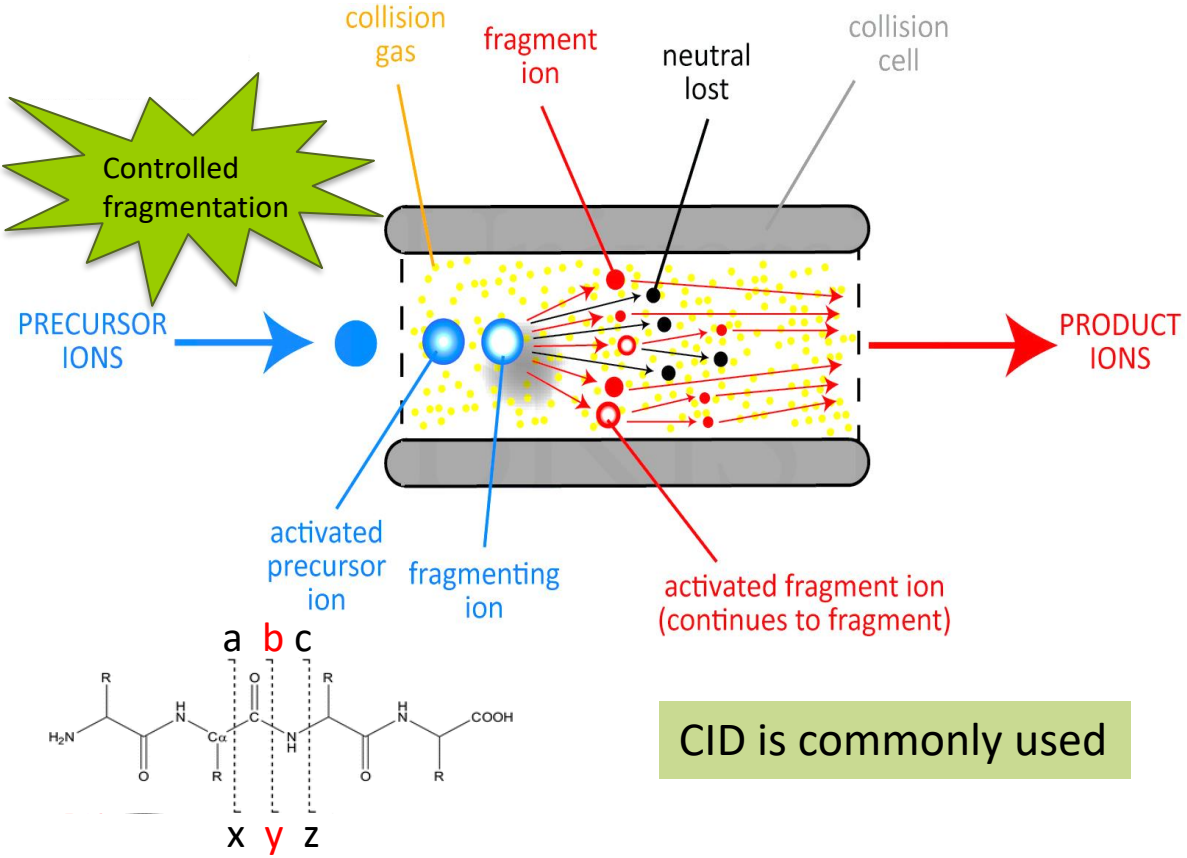


Qq-ToF mass spectrometer



Q-ToF instruments are well suited to quantitatively analyse large molecules like proteins (high sensitivity and high resolution).

MS/MS ion fragmentation modes: CID (collision induced dissociation)



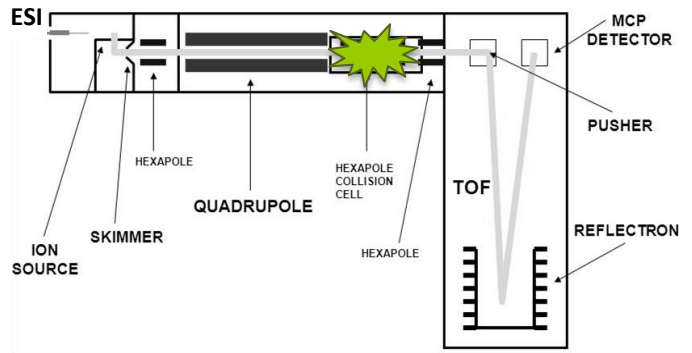
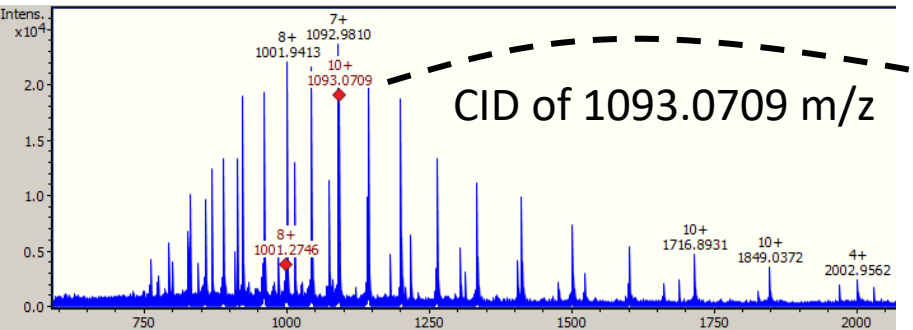
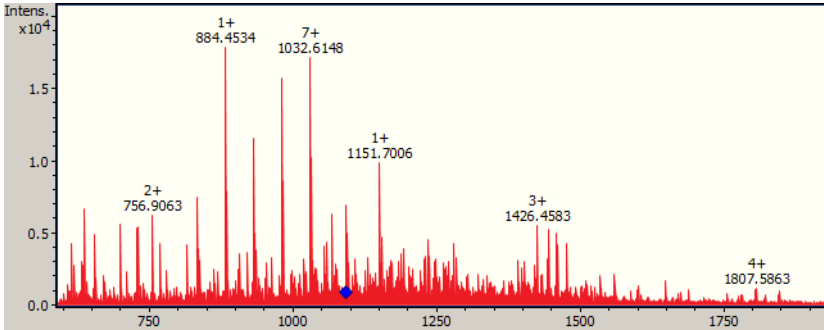


Figure of merit	"Low-energy" CID (slow activation)
Instruments used	Tandem quadrupoles, quadrupole hybrids (e.g., QqTOF)
Collision energy	1–200 eV
Collision number	10–100
Activation time scale	0.5–1 ms
Instrument time scale (kinetic window)/minimum observable reaction rate	0.1–1 ms/10^4–10^3 s^{-1}
Distribution of internal energy	Centered at few eV, no high energy tail
Variability of internal energy	Readily variable with collision energy to obtain energy resolved info.
Efficiency	5–50%
General results	Lower energy processes only, isomerization of precursor may occur, sequential dissociation observed

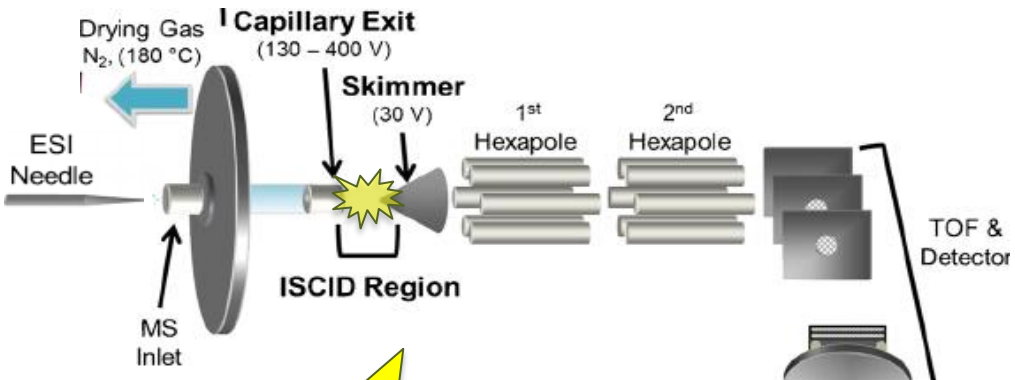
(Wells & McLuckey, 2005)

CID is commonly used

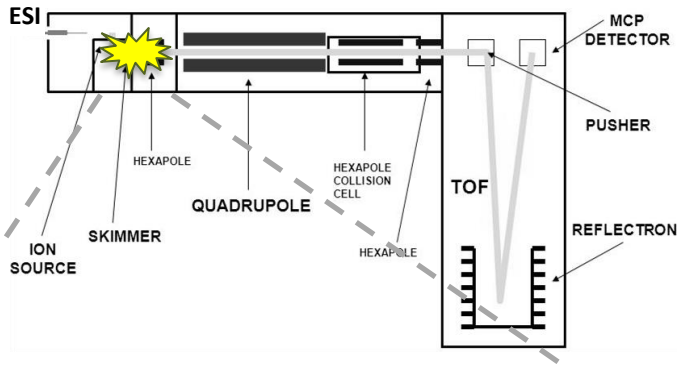


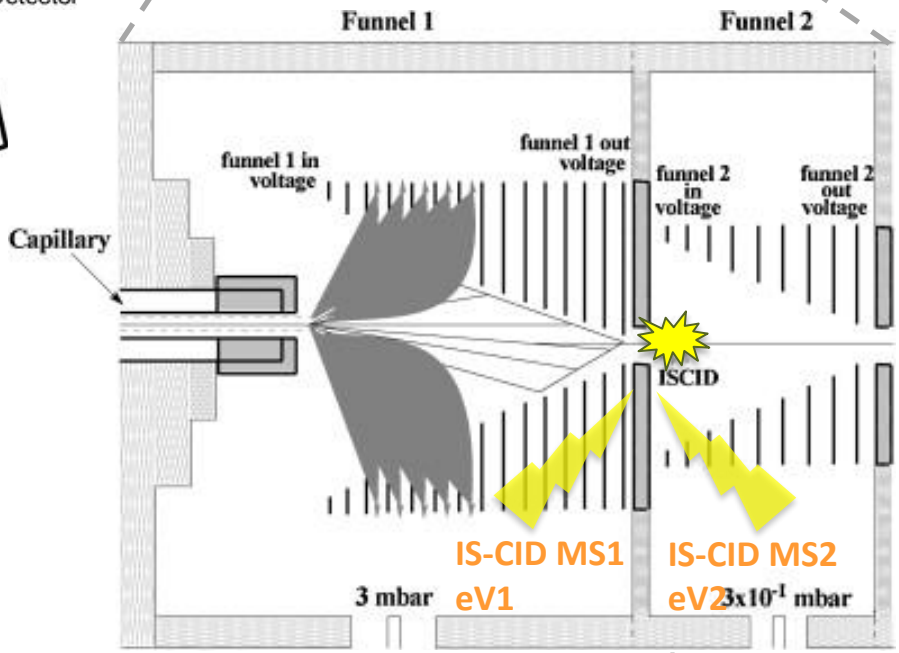


in- source-CID (IS-CID)



uncontrolled fragmentation

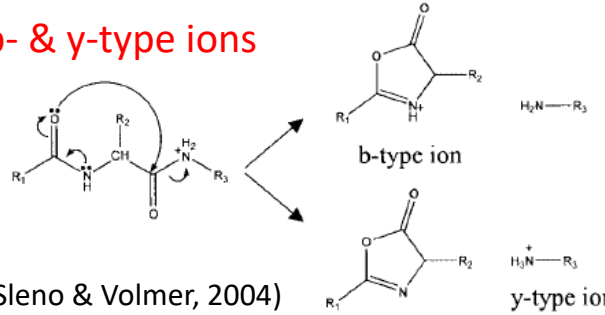




<u>IS-CID</u>:
The DC potential difference between the exit of funnel 1 (IS-CID MS1) and the entrance of funnel 2 (IS-CID MS2) causes uncontrolled fragmentation.
Very fast process (msec).

IS-CID is under-utilised.

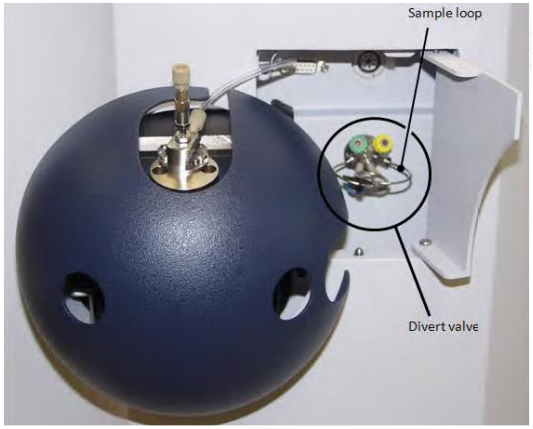
b- & y-type ions

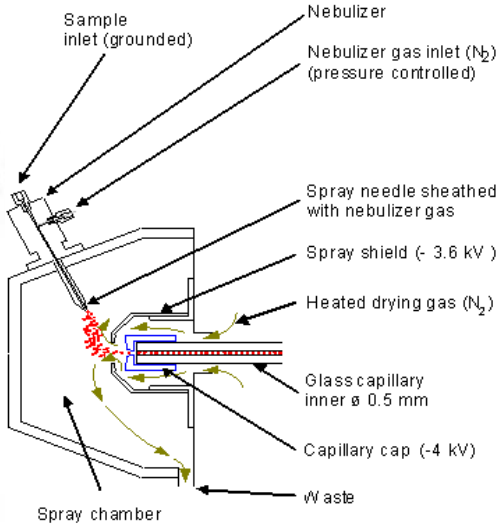


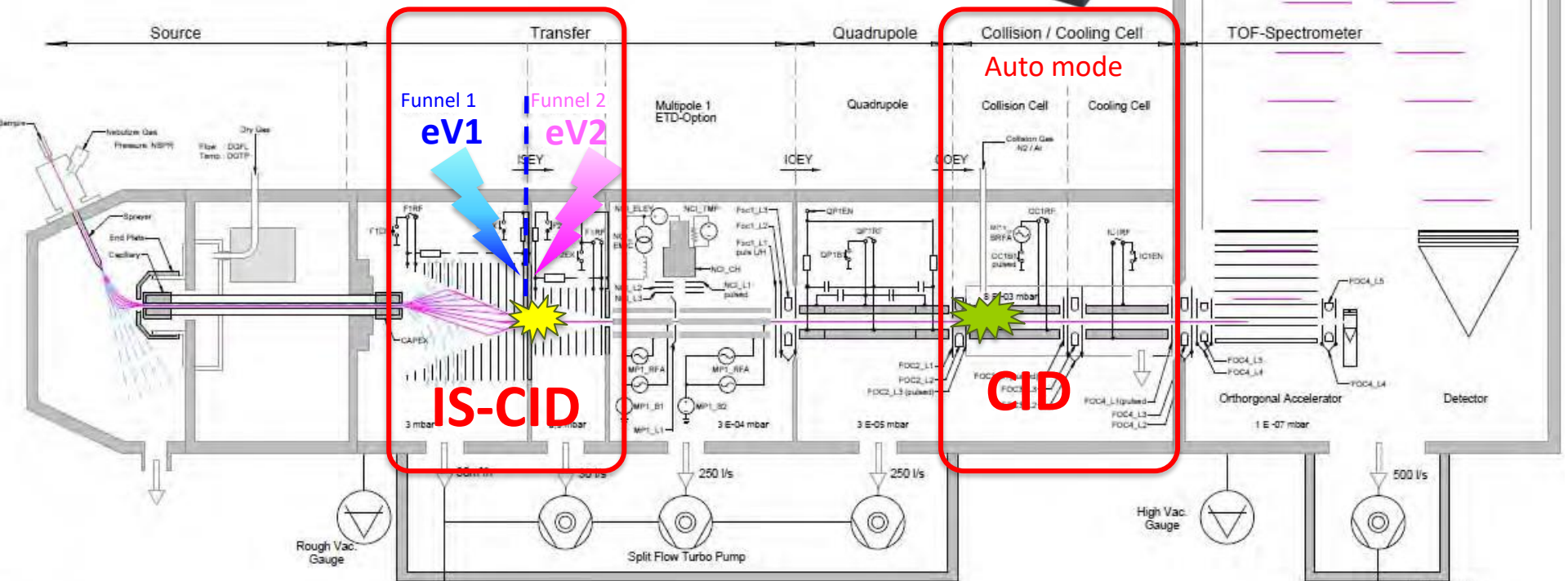
(Sleno & Volmer, 2004)

maXis Qq-ToF mass analyser (Bruker)

ESI source





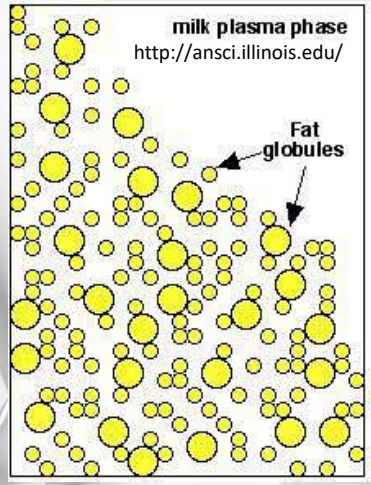


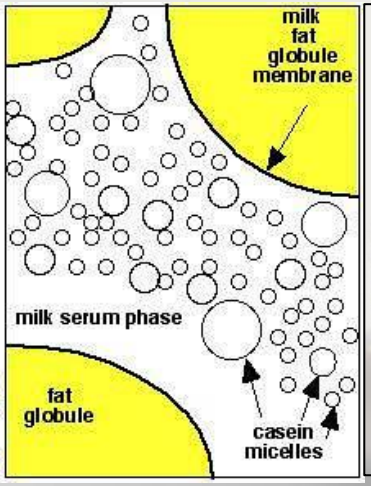
Milk components

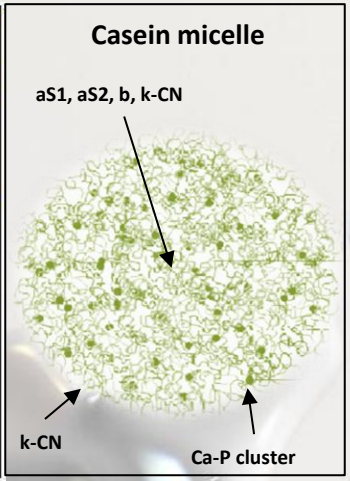
Milk: emulsion of fat globules, a suspension of casein micelles (casein, calcium, phosphorous), suspended in an aqueous phase, with solubilized lactose, whey proteins, and some minerals.

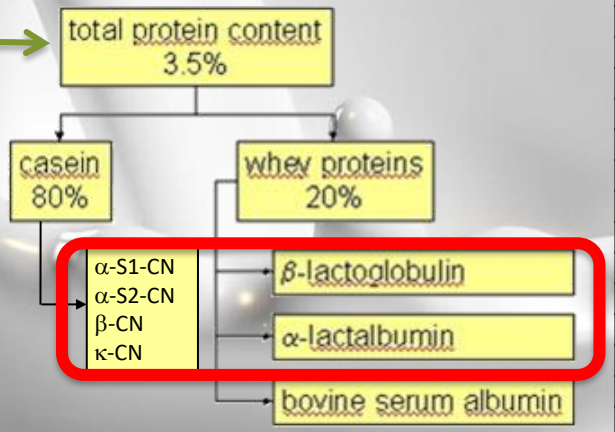
Nutrition Facts
Milk, whole, 3.25% fat

Amount Per 100 grams	
Calories 61	
Total Fat 3.2 g	
Saturated fat 1.9 g	
Polyunsaturated fat 0.2 g	
Monounsaturated fat 0.8 g	
Cholesterol 10 mg	
Sodium 43 mg	
Potassium 132 mg	
Total Carbohydrate 4.8 g	
Dietary fiber 0 g	
Sugar 5 g	
Protein 3.2 g	
Vitamin A 3%	Vitamin C
Calcium 11%	Iron
Vitamin D 12%	Vitamin B-6
Vitamin B-12 6%	Magnesium









Protein	Concentration (g/L)
aS1-CN	10.0
b-CN	9.3
k-CN	3.3
b-LG	3.2
aS2-CN	2.6
a-LA	1.2
g-CN	0.8
Ig	0.8
PP/8F/8S	0.5
BSA	0.4
MFGM	0.4
PP3	0.3
Lactoferrin	0.1
Transferrin	0.1
Total	33.0

LC-MS/MS analyses of intact milk proteins

Jersey
Holstein

Chromatogram
Separation over time by hydrophobicity

Deconvoluted mass
aLA B

MS

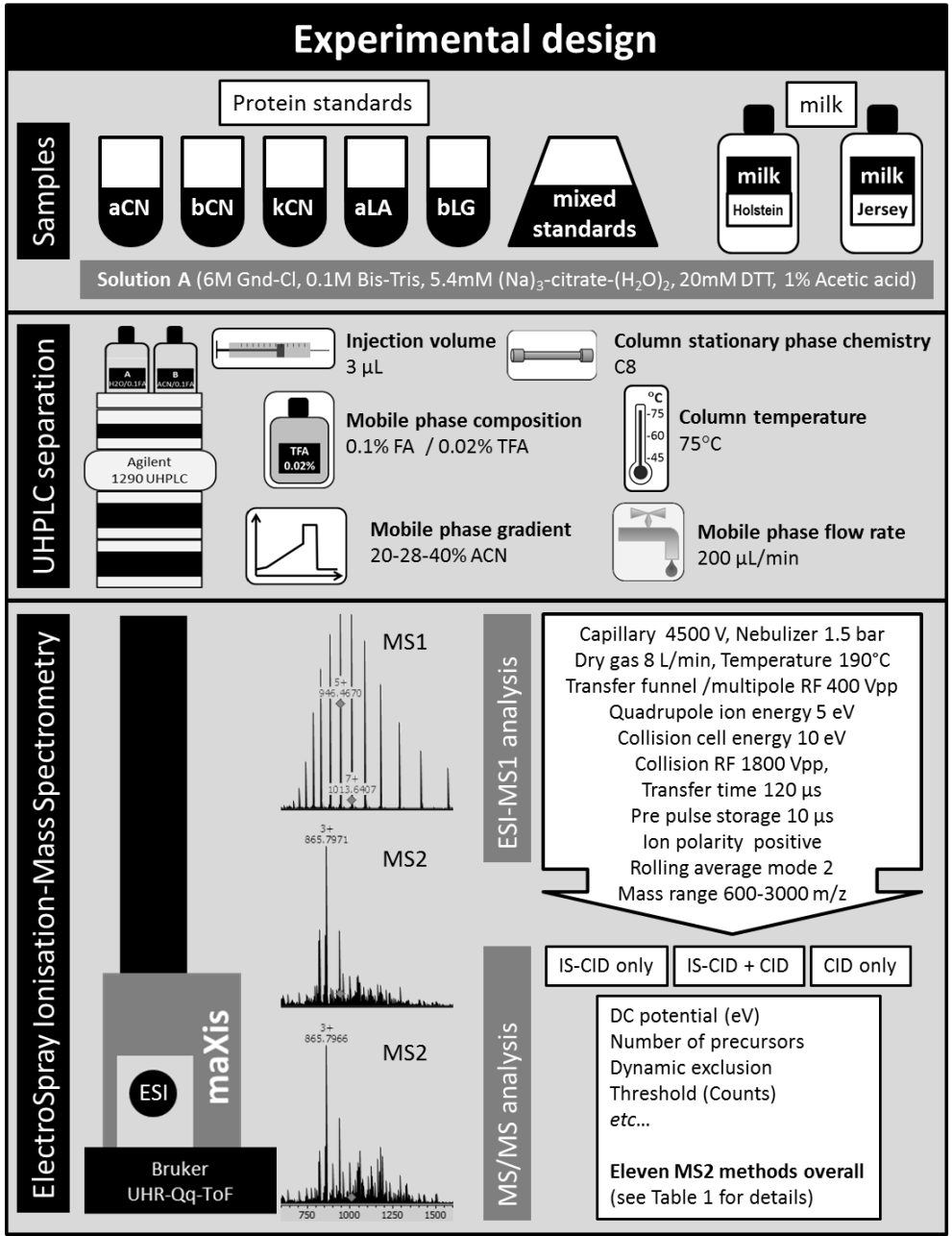
Mass spectra @ RT 16.5 min
Separation by mass

MS/MS

aLA B [1-123] top-down sequencing

EQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQAIVQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVGINYWLAHKALCSEKLDQWLCEKL

Improving top-down sequencing (TDS) of milk major proteins



Aim:

Improving sequencing coverage by top-down proteomics using a LC-ESI-Q-ToF platform.

Problem:

Need to develop high throughput datamining tools to assess methods. Annotation activities do not exist yet in Genedata Refiner.

Solution:

Collaboration with Dr Dominik Mertens (Genedata Expressionist) to create relevant processing workflows.

MS/MS Methods tested

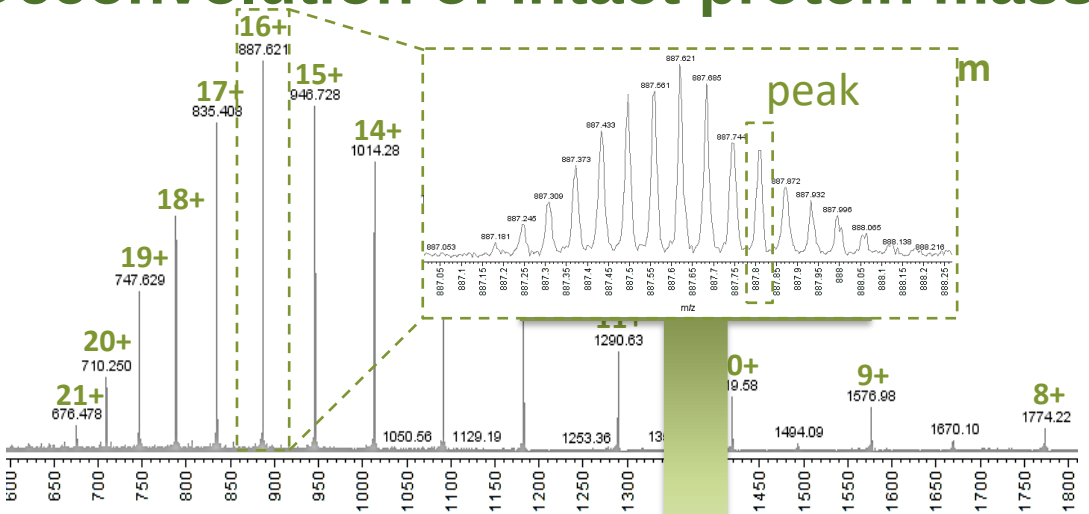
Parameter	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Method 9	Method 10	Method 11	Method 12
End plate offset	500 V	500 V	500 V	500 V	500 V	500 V	500 V	500 V	500 V	500 V	500 V	500 V
Capillary	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V	4500 V
Nebulizer	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar	1.5 bar
Dry gas	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min	8 L/min
Temperature	190°C	190°C	190°C	190°C	190°C	190°C	190°C	190°C	190°C	190°C	190°C	190°C
Waste (Na-formate infusion)	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min	0-2.5 min
Source	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min	2.5-40 min
Funnel RF	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp	400 Vpp
ISC												0 eV
Tr												Vpp
Io												5 eV
Lo												m/z
Cc												0 eV
Cc												Vpp
Tr												0 µs
Pr												0 µs
Io												tive
M												m/z
Su												321
Rc												2
isC												YES
isC												0 eV
isC												5 eV
Acquisition time factor				1	1	1	1	1	1	1	1	1
Auto MS/MS mode	OFF	ON	ON	OFF	ON	ON	ON	ON	ON	ON	ON	OFF
preference activated	NO	NO	NO	NO	NO	NO	NO	NO	YES	NO	NO	NO
charge state preferred range									7 - 15			
exclude singly									YES			
Acquisition	NO	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO
MS/MS low (per 1000 sum.)		10000 cts	10000 cts		10000 cts	10000 cts	10000 cts	10000 cts	10000 cts	10000 cts	10000 cts	
MS/MS high (per 1000 sum.)		1000000 cts	1000000 cts		1000000 cts	1000000 cts	1000000 cts	1000000 cts	1000000 cts	1000000 cts	1000000 cts	
Exclude		100-600 m/z	100-600 m/z		100-600 m/z	100-600 m/z	100-600 m/z	100-600 m/z	100-600 m/z	100-600 m/z	100-600 m/z	
No. of precursors		2	2		2	2	2	3	4	4	4	
Absolute (per 1000 sum.)		1153 cts	170 cts		10 cts	10 cts	10 cts	10 cts	10 cts	10 cts	10 cts	
Absolute		17302 cts	2550 cts		183 cts	183 cts	183 cts	183 cts	183 cts	183 cts	183 cts	
ON/OFF		ON	ON		ON	ON	ON	ON	ON	ON	ON	
Exclude after		3 spectra	3 spectra		1 spectrum	1 spectrum	1 spectrum	1 spectrum	1 spectrum	1 spectrum	1 spectrum	
Release after		5 min	5 min		5 min	5 min	5 min	5 min	5 min	5 min	5 min	

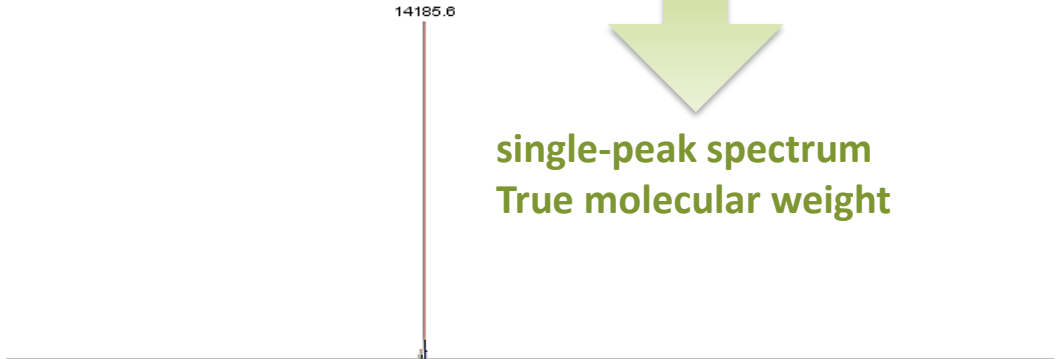
Parameter	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Method 9	Method 10	Method 11	Method 12
Transfer time (µs)	120	120	120	120	120	120	120	120	150	120	120
Pre pulse storage (µs)	10	10	10	10	10	10	10	10	15	10	10
isCID	OFF	OFF	ON	ON	ON	ON	ON	ON	ON	ON	ON
isCID MS (eV)			3	30	50	70	5	3	3	3	10
isCID MS/MS (eV)			15	10	10	10	20	10	30	10	15
Auto MS/MS mode	ON	ON	OFF	ON	ON	ON	ON	ON	ON	ON	OFF
No. of precursors	2	2		2	2	2	3	4	4	4	
Absolute (cts)	17302	2550		183	183	183	183	183	183	183	
summary	CID	CID	IS-CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID

MS

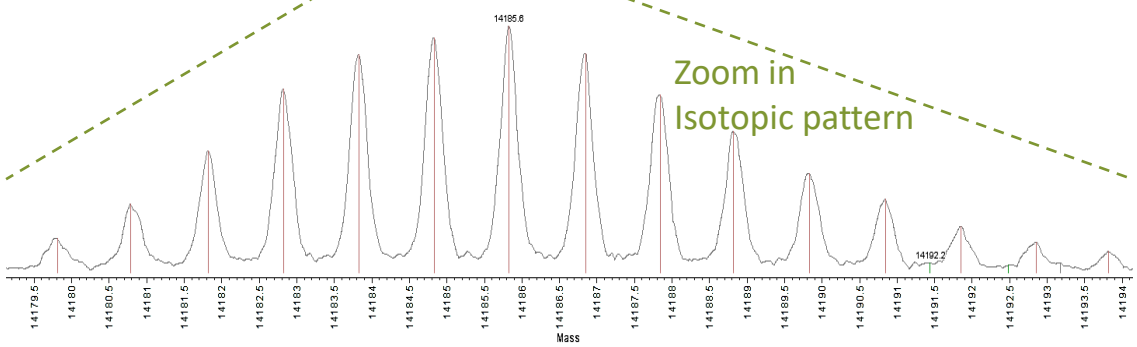
Protein deconvolution

Deconvolution of intact protein mass spectra

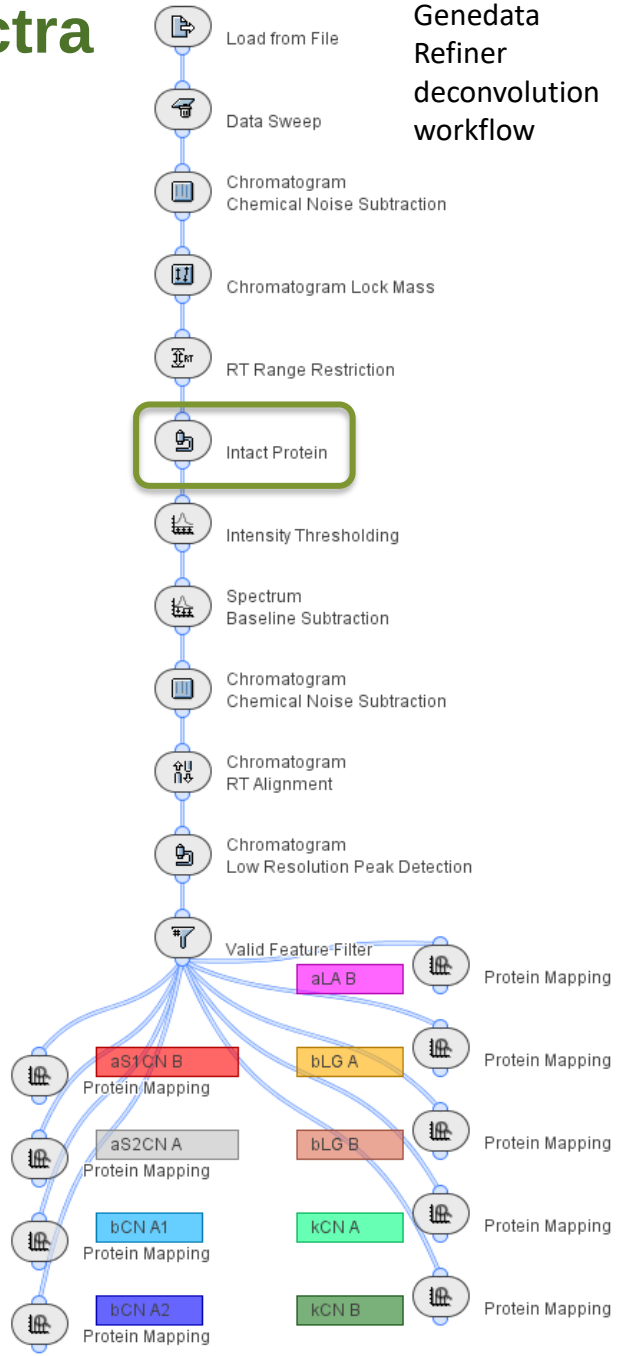




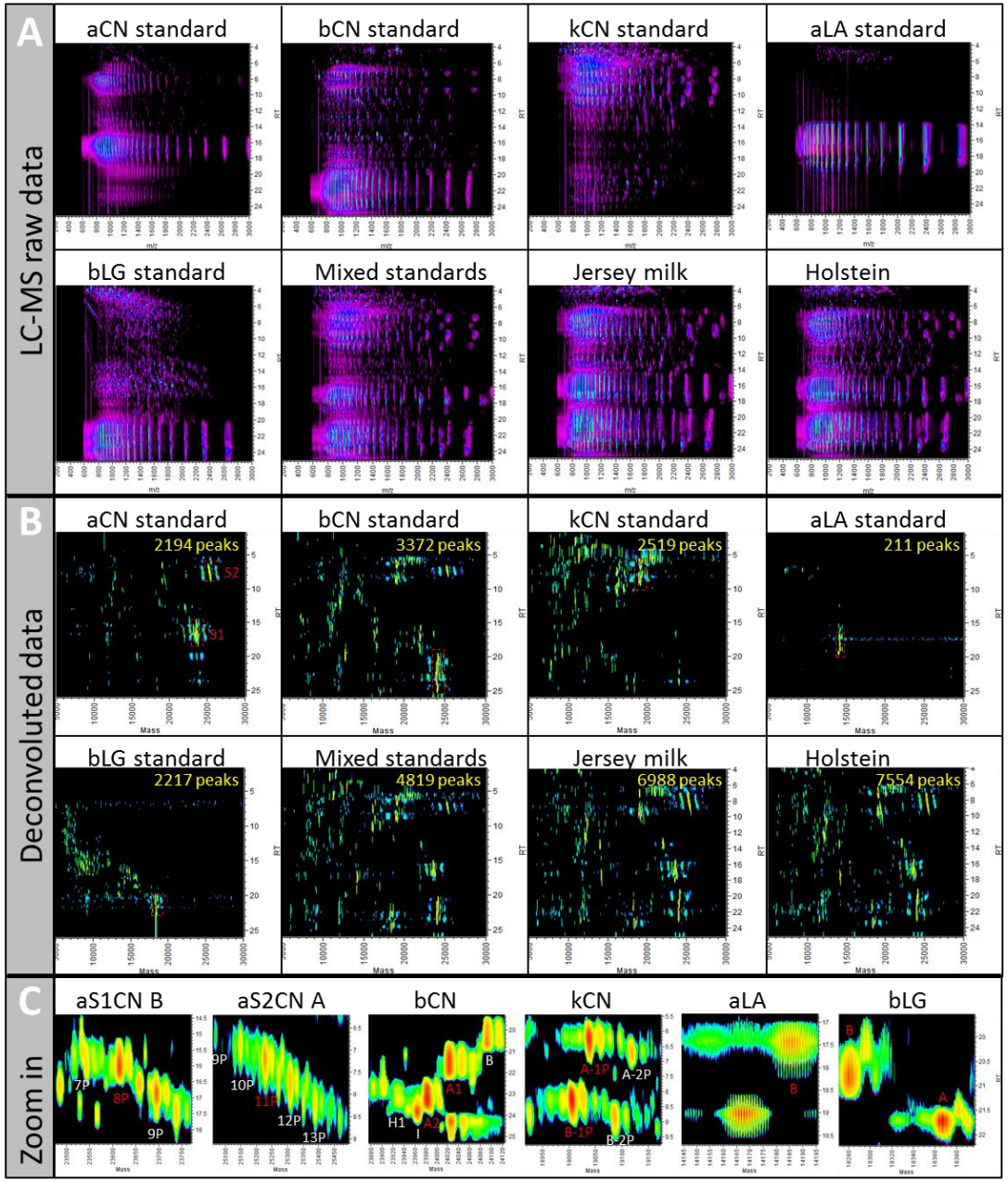
single-peak spectrum
True molecular weight

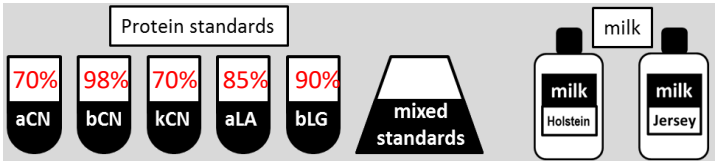


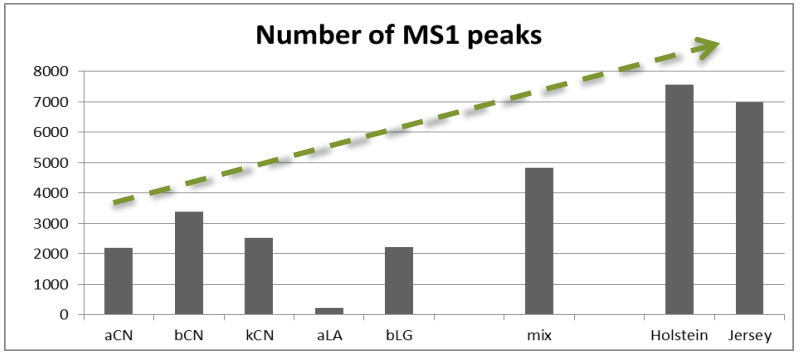
Genedata Refiner deconvolution workflow



Samples of increasing complexity







Complexity:

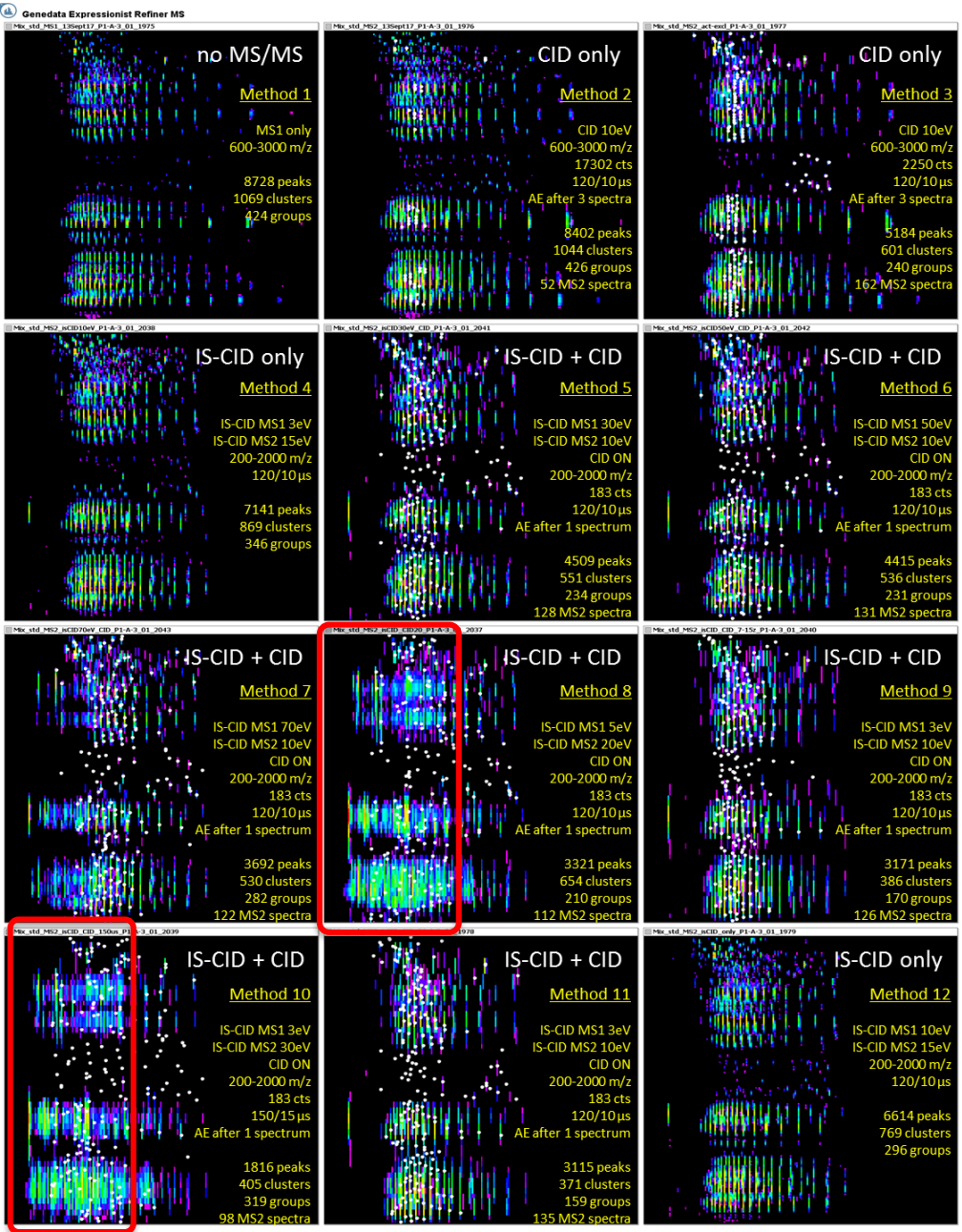
aLA < aCN < bLG < kCN < bCN < mix < milk

Successful sample processing by Genedata regardless of its level of complexity.

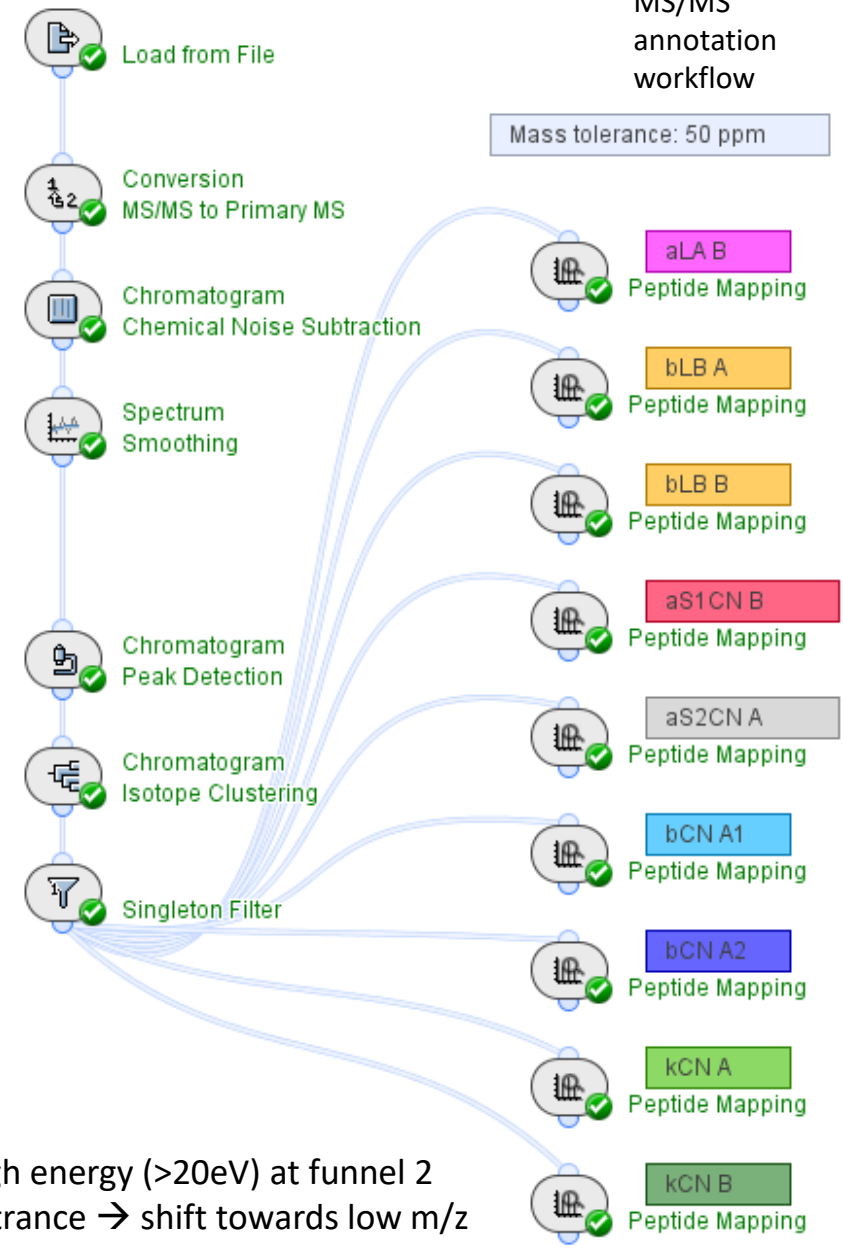
MS/MS

Protein annotation

Effect of the methods on the LC/MS patterns

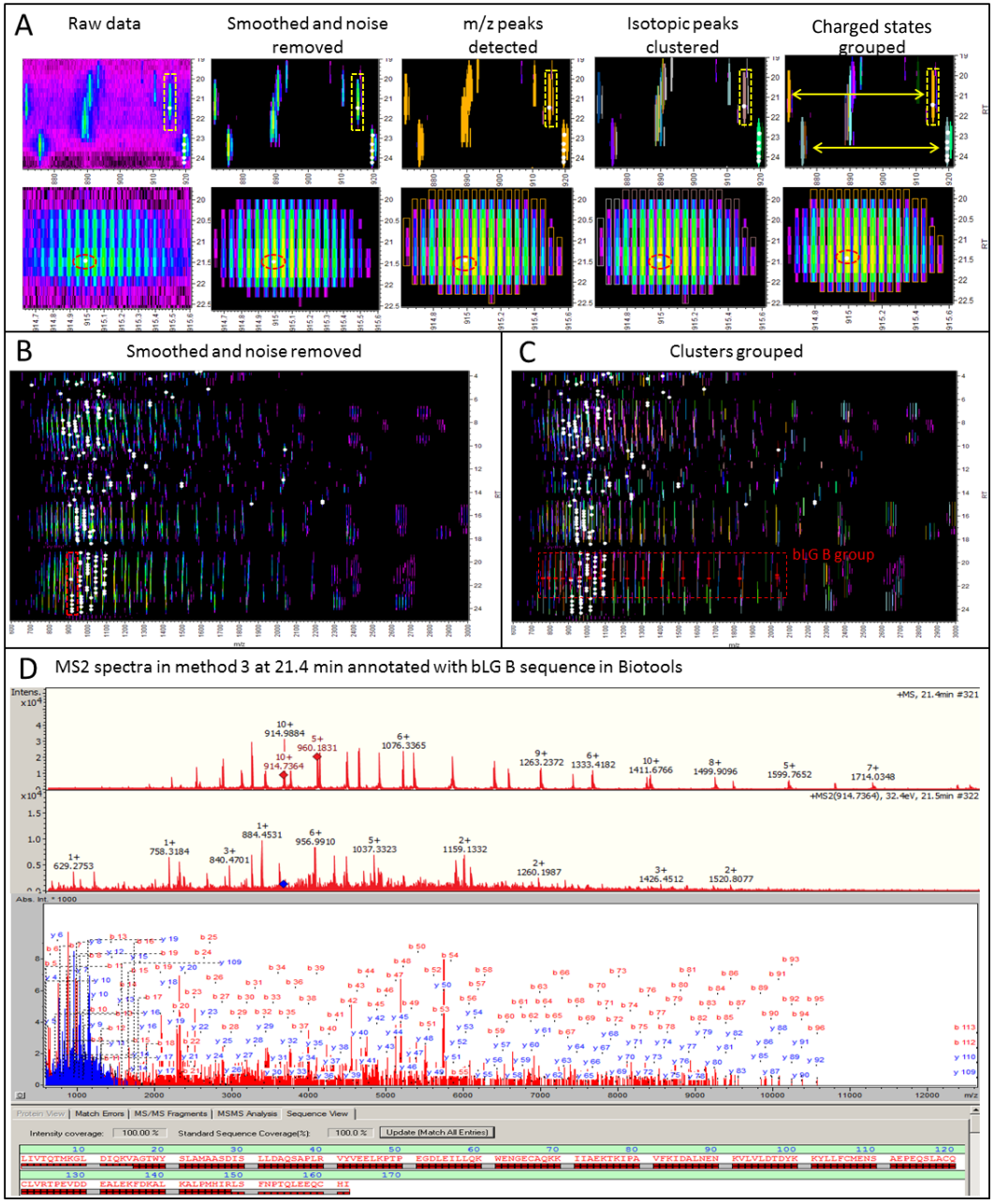


Genedata Refiner MS/MS annotation workflow



High energy (>20eV) at funnel 2 entrance → shift towards low m/z

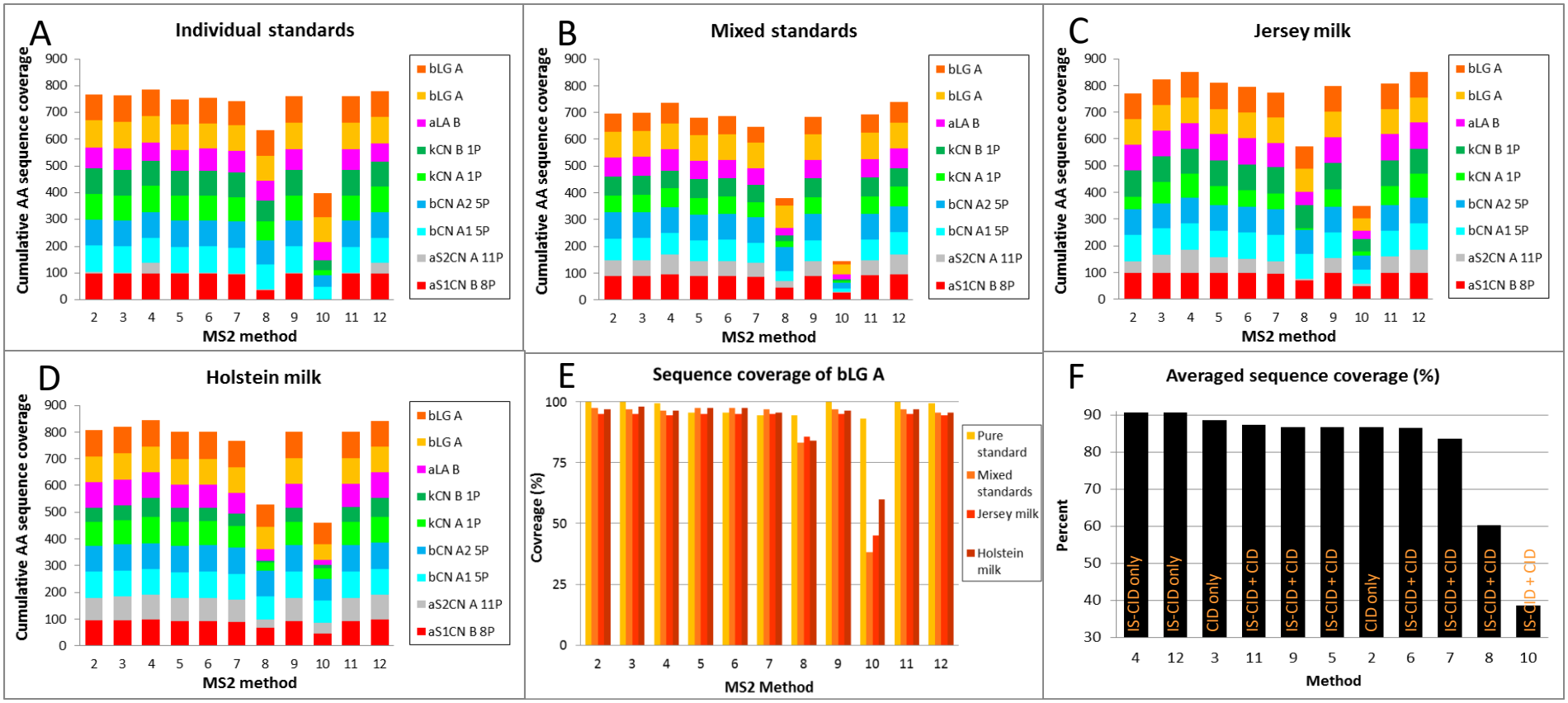
Annotation of MS/MS spectra (e.g. bLG B)



Upon (IS-)CID fragmentation, the intact protein is cut into peptides of various sizes anywhere along the AA sequence.

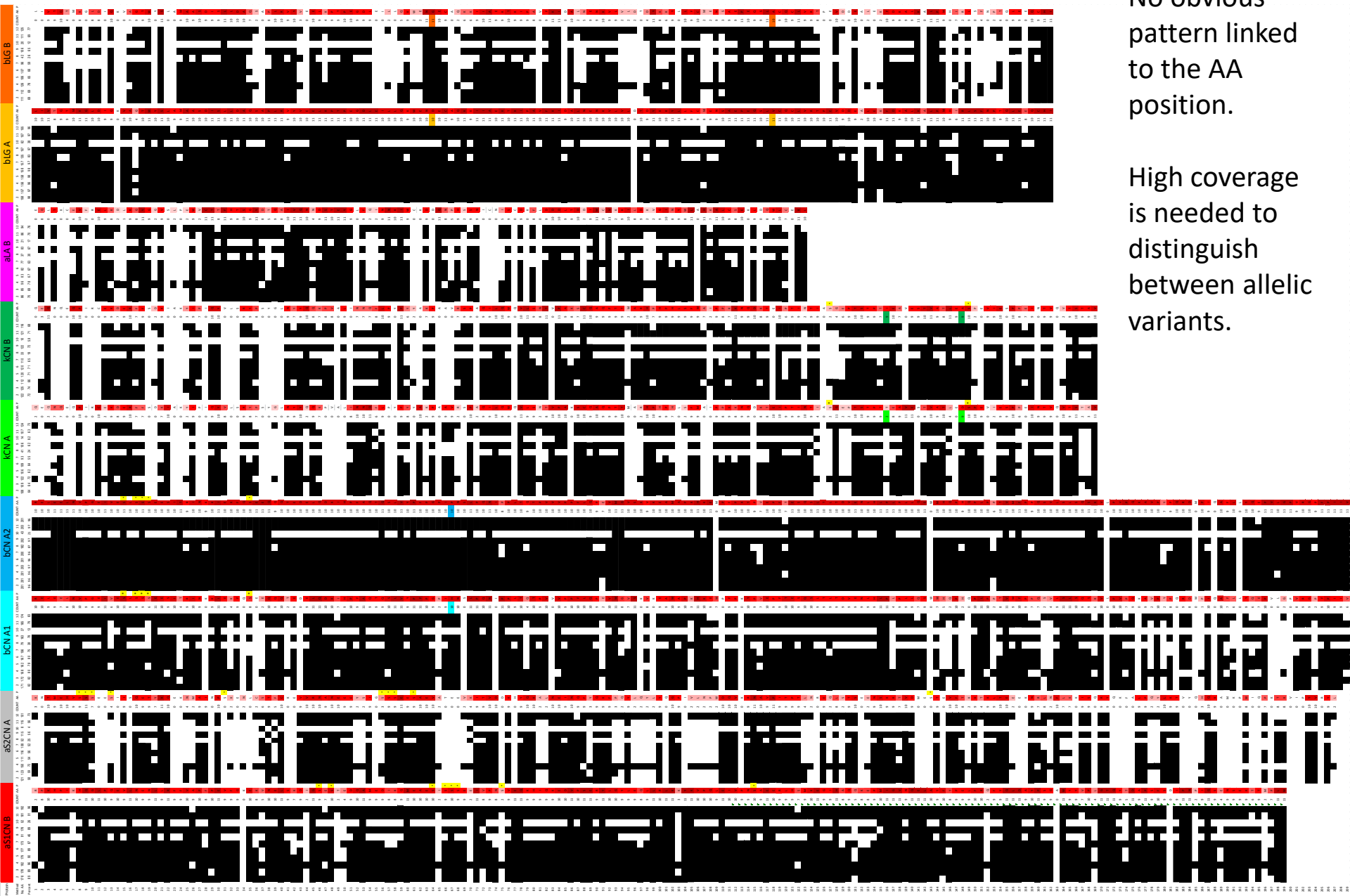
TDS is successful if annotation produces a “ladder” pattern (1 AA shift at a time).

Comparison of MS/MS methods



Parameter	Method 4	Method 12	Method 3	Method 11	Method 9	Method 5	Method 2	Method 6	Method 7	Method 8	Method 10
Transfer time (µs)	120	120	120	120	120	120	120	120	120	120	150
Pre pulse storage (µs)	10	10	10	10	10	10	10	10	10	10	15
isCID	ON	ON	OFF	ON	ON	ON	OFF	ON	ON	ON	ON
isCID MS (eV)	3	10		3	3	30		50	70	5	3
isCID MS/MS (eV)	15	15		10	10	10		10	10	20	30
Auto MS/MS mode	OFF	OFF	ON	ON	ON	ON	ON	ON	ON	ON	ON
No. of precursors			2	4	4	2	2	2	2	3	4
Absolute (cts)			2550	183	183	183	17302	183	183	183	183
summary	IS-CID	IS-CID	CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	CID	IS-CID+CID	IS-CID+CID	IS-CID+CID	IS-CID+CID
sequence coverage(%)	90.64	90.64	88.63	87.37	86.80	86.74	86.72	86.57	83.61	60.25	38.61

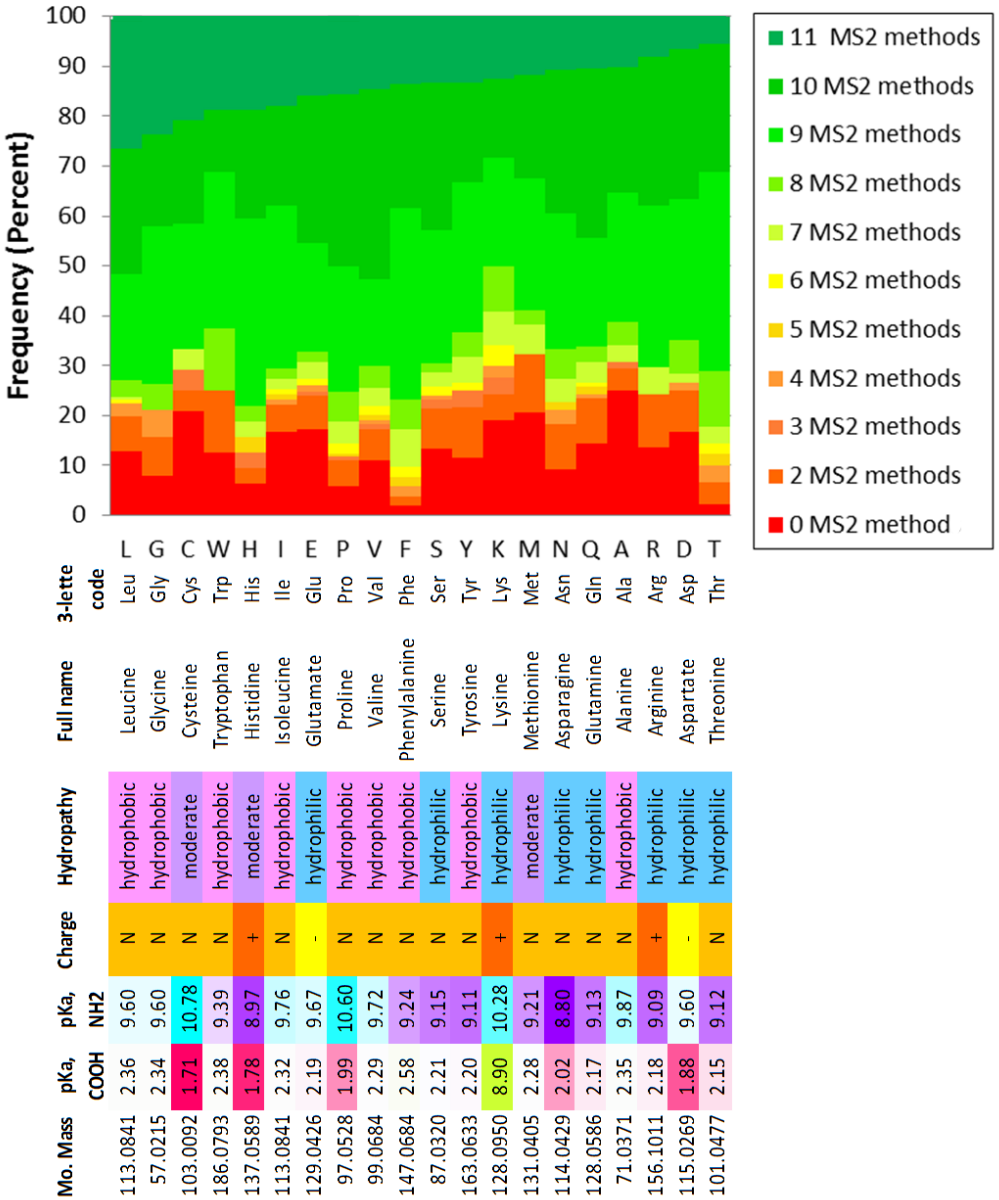
AA position does not impact TDS efficacy



No obvious pattern linked to the AA position.

High coverage is needed to distinguish between allelic variants.

AA hydrophobocity impacts TDS efficacy



Leucine, Glycine, Cysteine, Tryptophan, and Histidine, Isoleucine displayed the best response to TDS.

These AAs present moderate to high hydrophobicity.

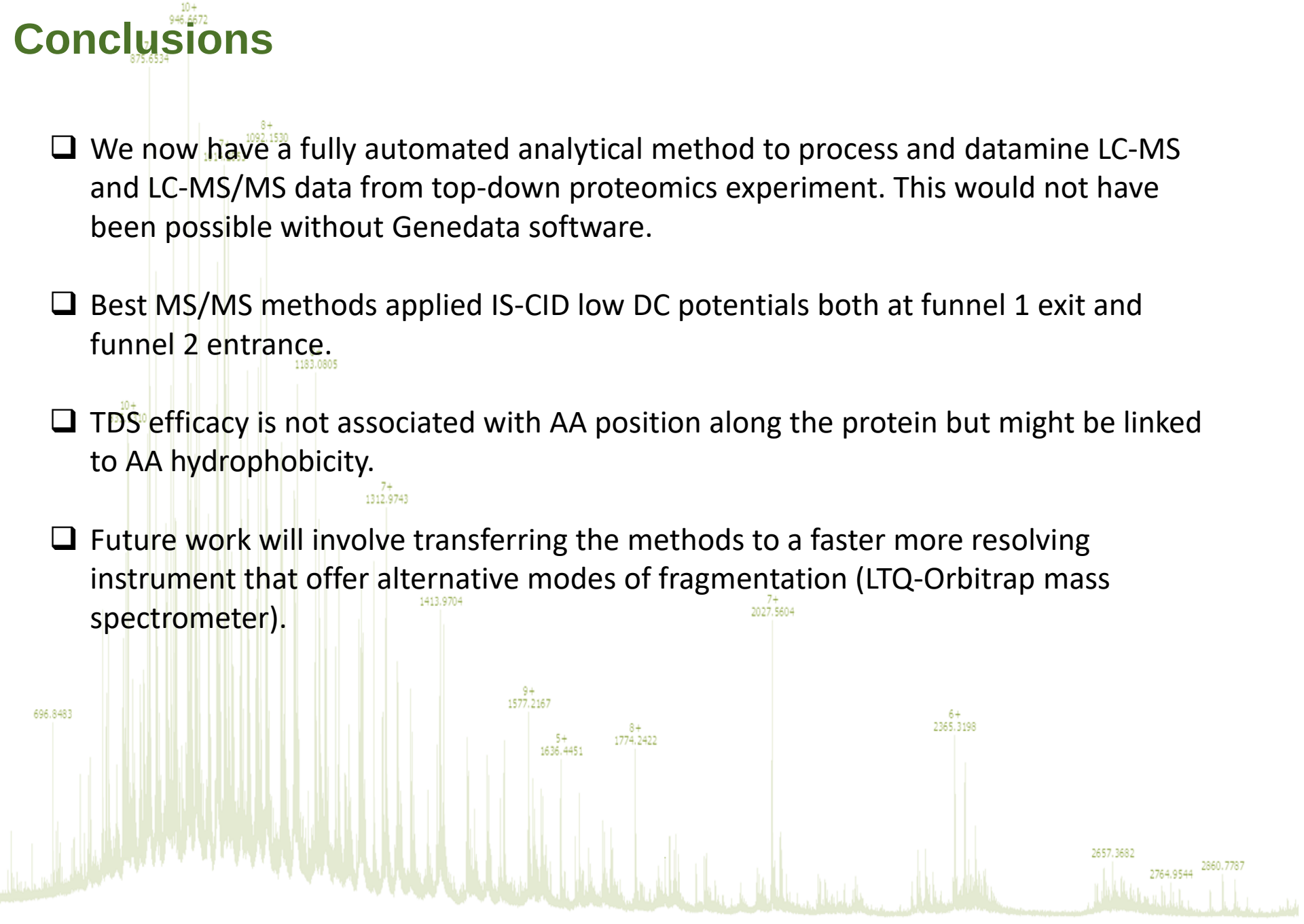
Conversely, Threonine, Asparagine and Arginine showed the least success rate.

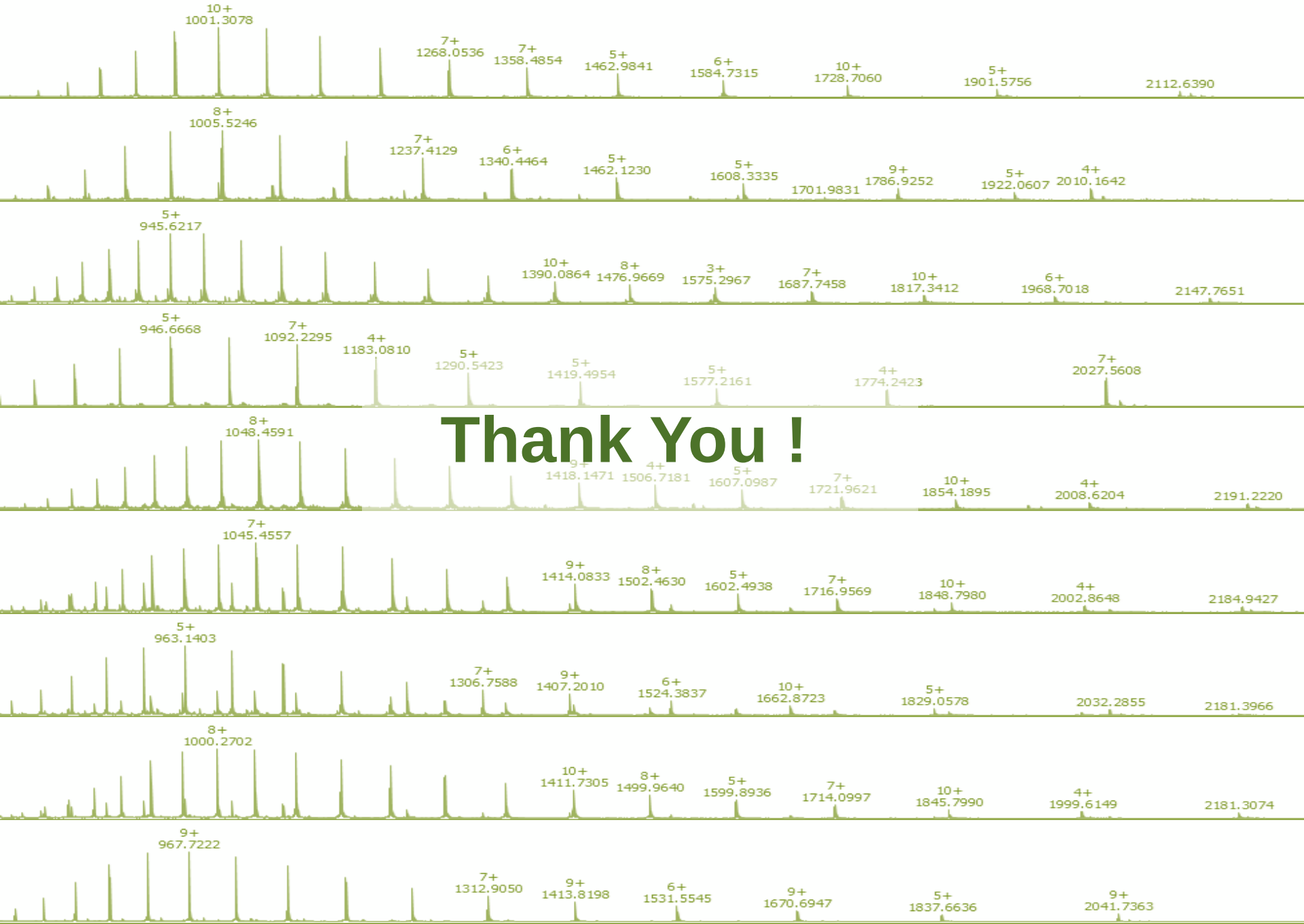
These AAs are hydrophilic.

AA hydrophobicity level would have an impact on fragmentation efficiency.

Conclusions

- We now have a fully automated analytical method to process and datamine LC-MS and LC-MS/MS data from top-down proteomics experiment. This would not have been possible without Genedata software.
- Best MS/MS methods applied IS-CID low DC potentials both at funnel 1 exit and funnel 2 entrance.
- TDS efficacy is not associated with AA position along the protein but might be linked to AA hydrophobicity.
- Future work will involve transferring the methods to a faster more resolving instrument that offer alternative modes of fragmentation (LTQ-Orbitrap mass spectrometer).




10+
1001.3078
7+
1268.0536
7+
1358.4854
5+
1462.9841
6+
1584.7315
10+
1728.7060
5+
1901.5756
2112.6390
8+
1005.5246
7+
1237.4129
6+
1340.4464
5+
1462.1230
5+
1608.3335
1701.9831
9+
1786.9252
5+
1922.0607
4+
2010.1642
5+
945.6217
10+
1390.0864
8+
1476.9669
3+
1575.2967
7+
1687.7458
10+
1817.3412
6+
1968.7018
2147.7651
5+
946.6668
7+
1092.2295
4+
1183.0810
5+
1290.5423
5+
1419.4954
5+
1577.2161
4+
1774.2423
7+
2027.5608
8+
1048.4591
9+
1418.1471
4+
1506.7181
5+
1607.0987
7+
1721.9621
10+
1854.1895
4+
2008.6204
2191.2220
7+
1045.4557
9+
1414.0833
8+
1502.4630
5+
1602.4938
7+
1716.9569
10+
1848.7980
4+
2002.8648
2184.9427
5+
963.1403
7+
1306.7588
9+
1407.2010
6+
1524.3837
10+
1662.8723
5+
1829.0578
2032.2855
2181.3966
8+
1000.2702
10+
1411.7305
8+
1499.9640
5+
1599.8936
7+
1714.0997
10+
1845.7990
4+
1999.6149
2181.3074
9+
967.7222
7+
1312.9050
9+
1413.8198
6+
1531.5545
9+
1670.6947
5+
1837.6636
9+
2041.7363

Thank You !