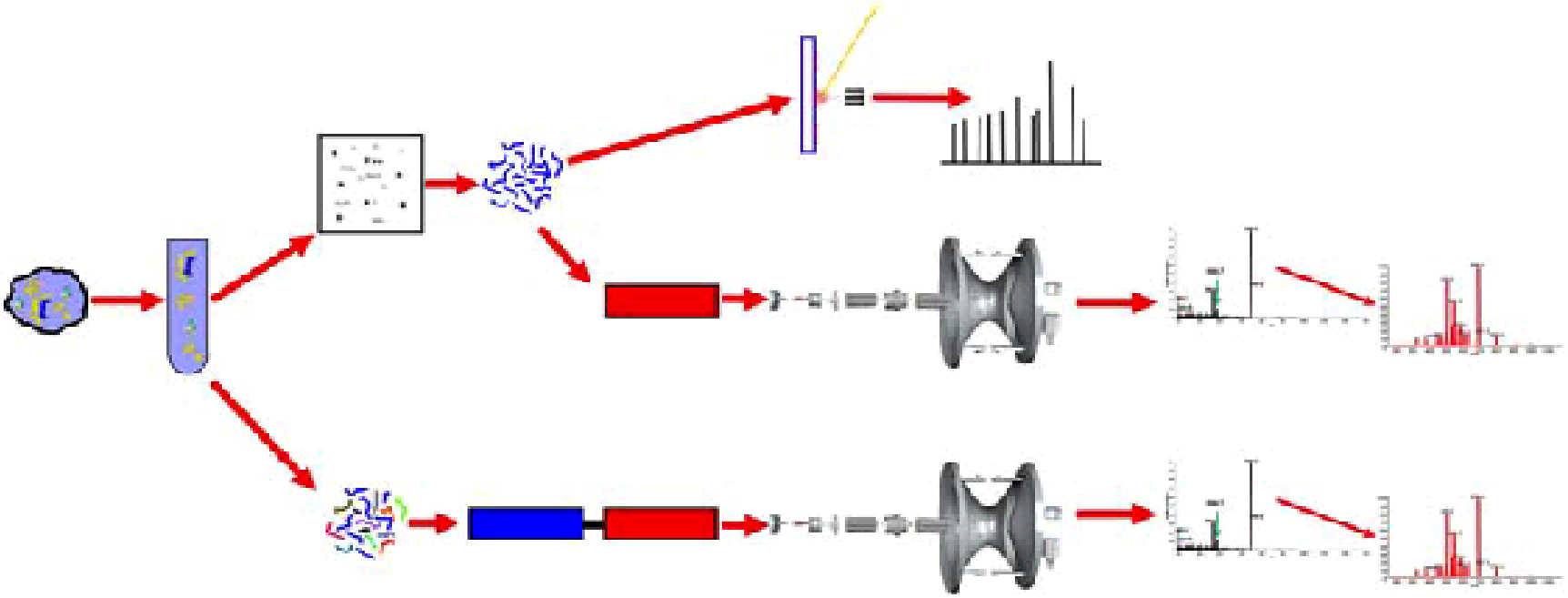
Large scale protein identification by Bi-Phasic column

Paul Shieh
Column Technology Inc.,

Large scale protein identification

Goal: to identified as many proteins as possible in tissues, cells or blood.

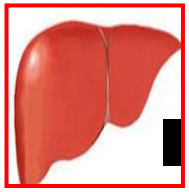
General approach: 1. 2D gel electrophoresis followed by LC/MS/MS analysis.
2. Shot-sequence with two dimensional LC/MS/MS.



Protein identification by 2D gel electrophoresis followed by LC/MS/MS

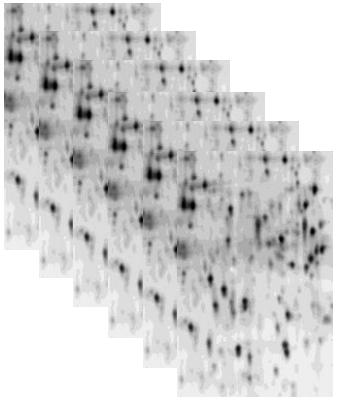
2D gel electrophoresis

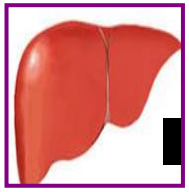
LC/MS/MS analysis



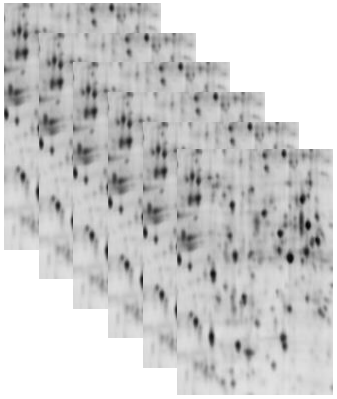
tumorous tissue

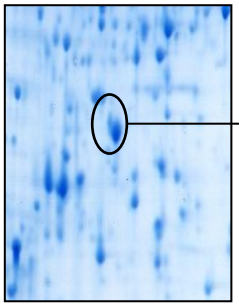
Lysis





paired nontumorous tissue





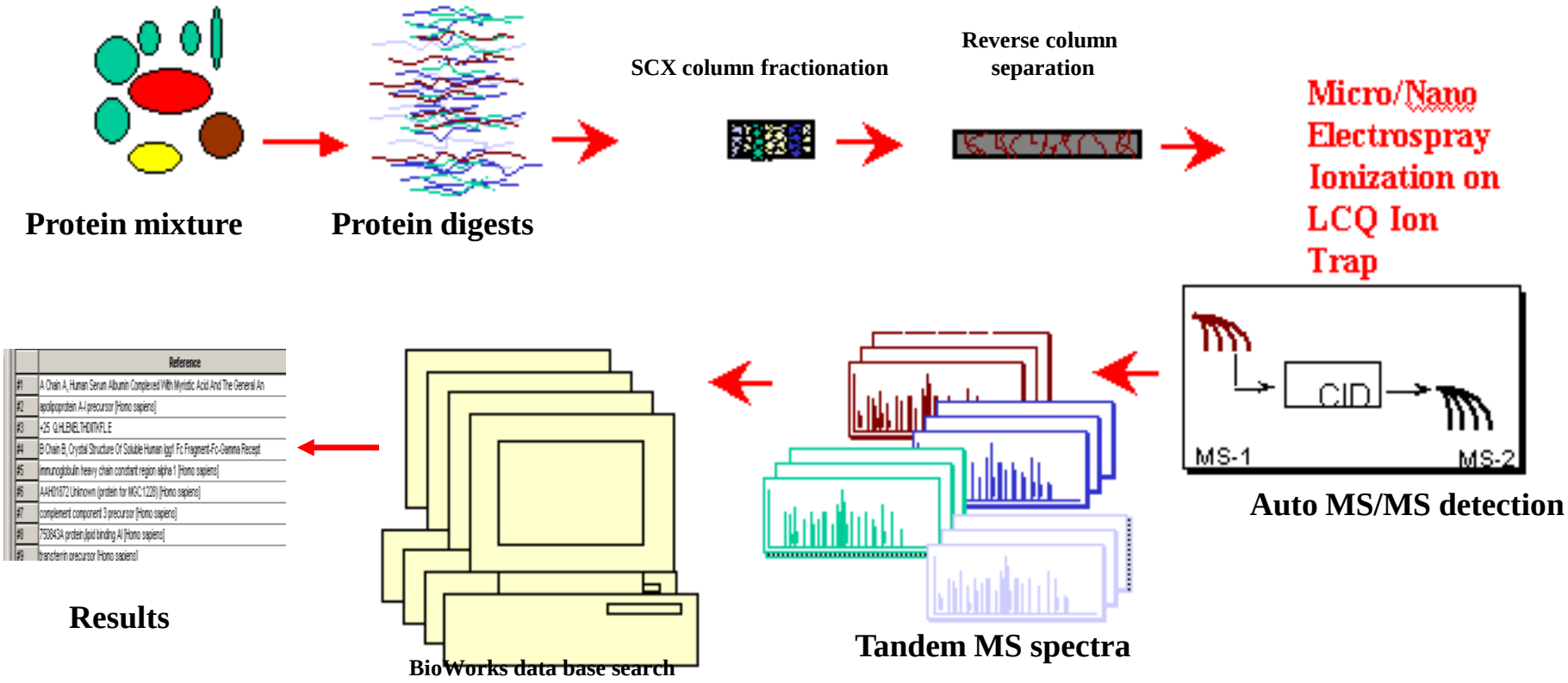
Digest

2D gel electrophoresis method

- Advantages:
  - High resolution protein separation based on their pI and size.
  - Traditional separation method, mature technology.
  - Visually identify the gel spots.
- Disadvantages:
  - Poor reproducibility.
  - Timely, labor intensive.
  - Not able to work on hydrophobic or high pI proteins
  - Can not automated.
  - Short dynamic range.

Two dimensional LC/MS/MS for protein identification

Common approach: SCX fractionation followed by reverse phase chromatography.



2D chromatography

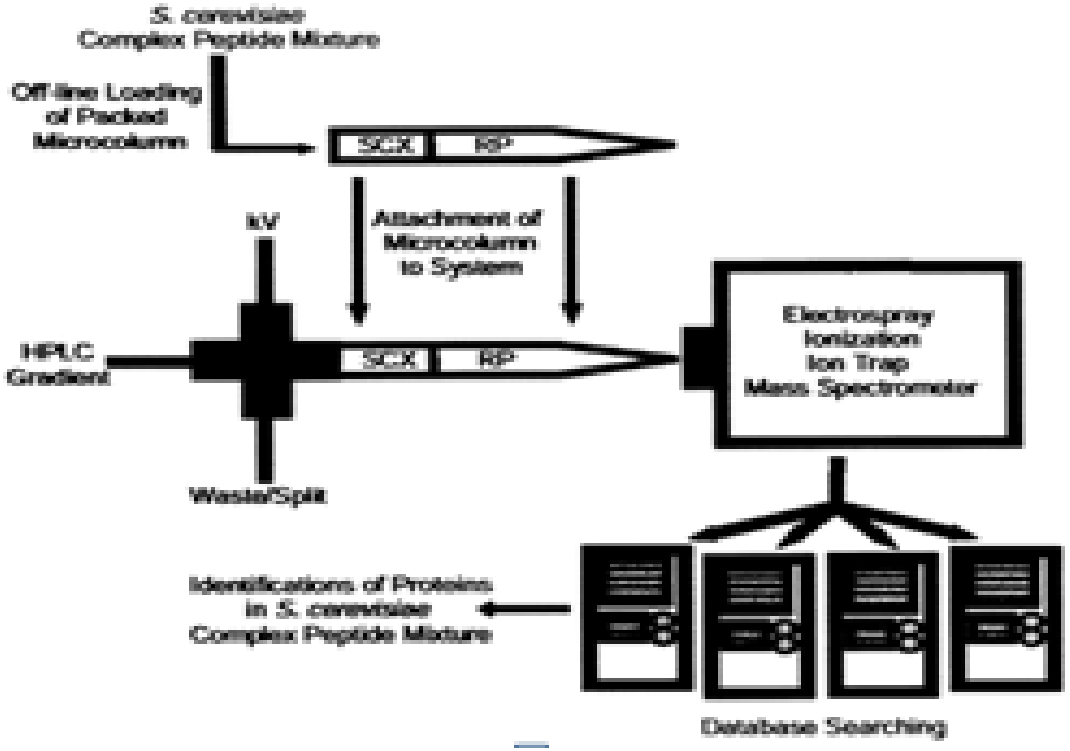
- Gidding's theory:

  For two orthogonal separation methods, the peak capacity is equal to the multiple of individual peak capacities.

$$P = P1 \times P2 \times P3$$

The first 2D LC/MS/MS set up (MUDPIT)

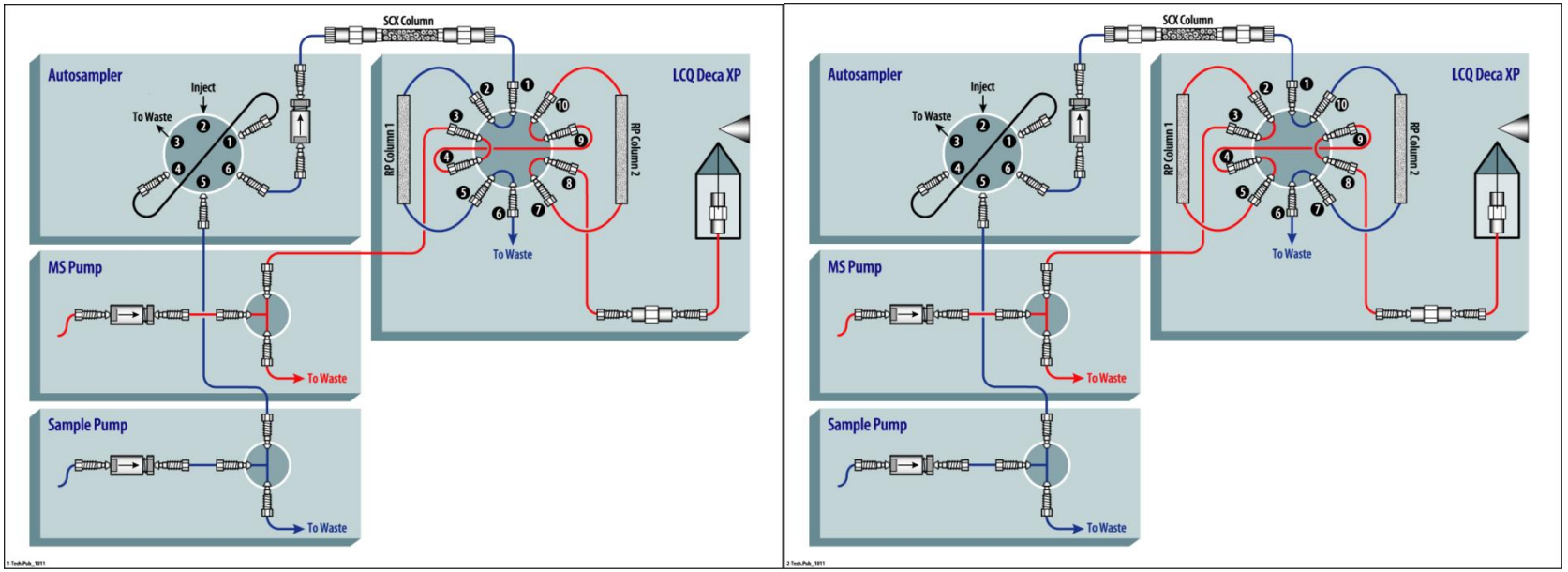
Spray tips packed by both SCX and reverse phase material . Hundreds to thousands of proteins were identified but still a serious problem with the salt issue.



First commercial available 2D LC/MS/MS (Proteome X)

Using salt gradient followed by reverse phase chromatography

Plumbing Diagrams for Proteome X.



Advantage: Fully automated 2D LC MS/MS system.
Disadvantages: require additional pump, valve, and software. Salt is still an issue.

CTI revolutional online 2D LC/MS/MS peptide/protein separation

Bi-phasic column with pH buffer

pH and reverse phase based 2D separation

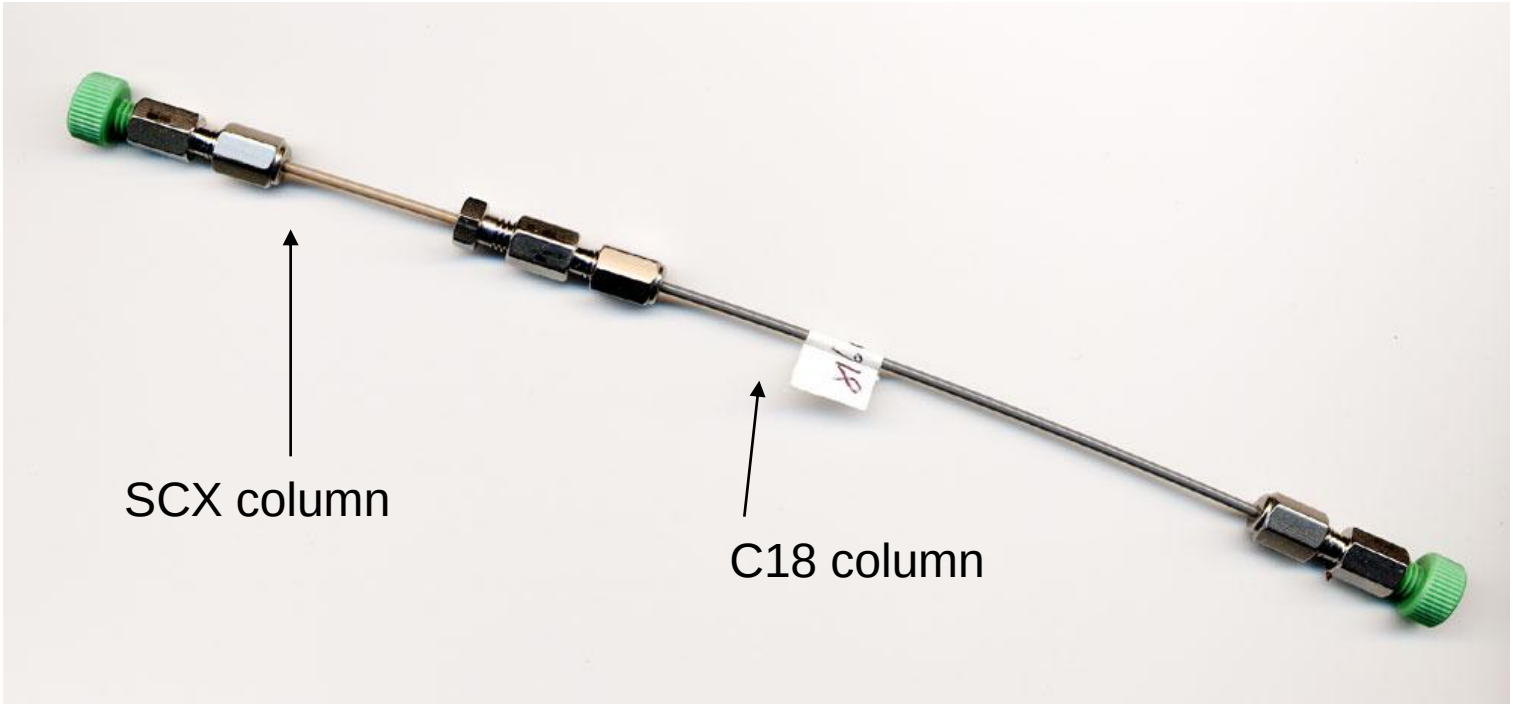
Advanced technology for large scale protein identification

Anal. Chem. most accessed paper in 2005

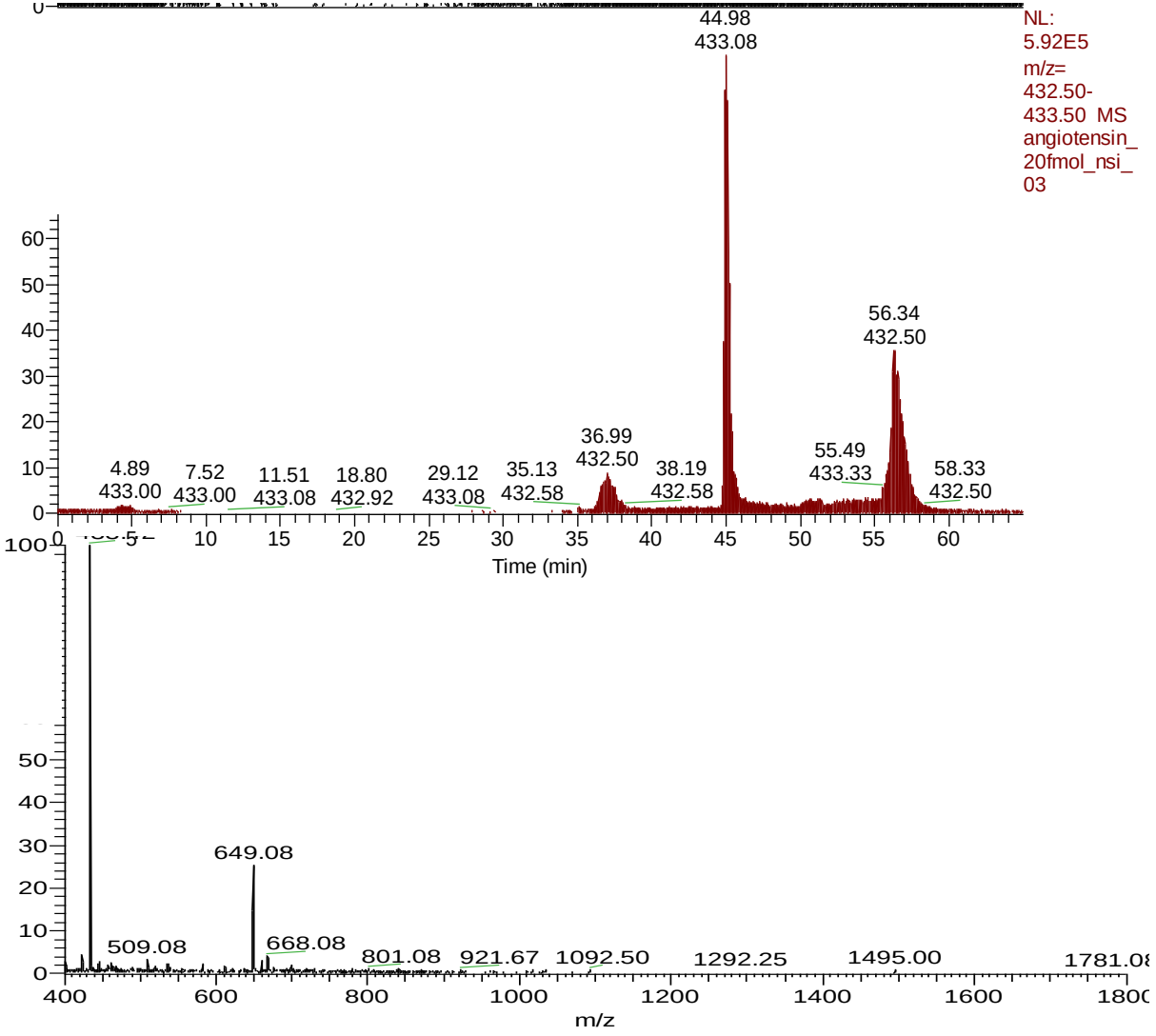
The simplest and most reproducible 2D LC/MS/MS method

Bi-Phasic column

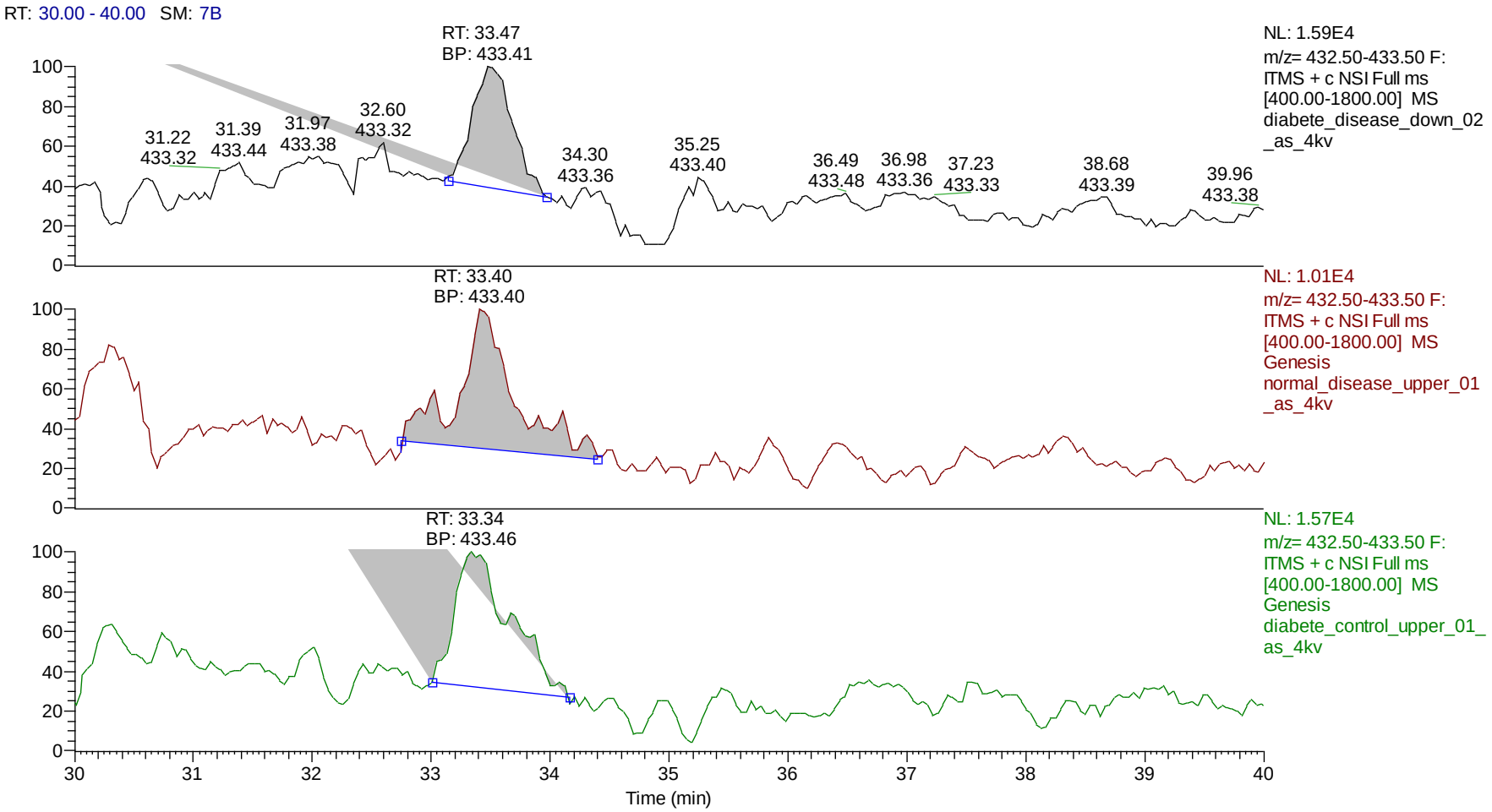
Integrated Multi-Dimensional LC (IMDL) based on pH elution.



LTQ sensitivity using CTI column:20 fmol Angiotensin (flow rate:2.5 ul/min)



10fmol angiotensin as Internal Standard in complicate biological sample

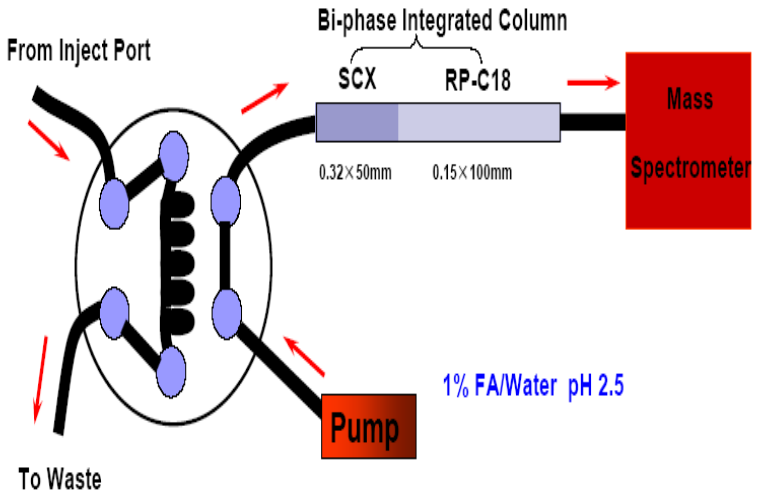


IMDL method

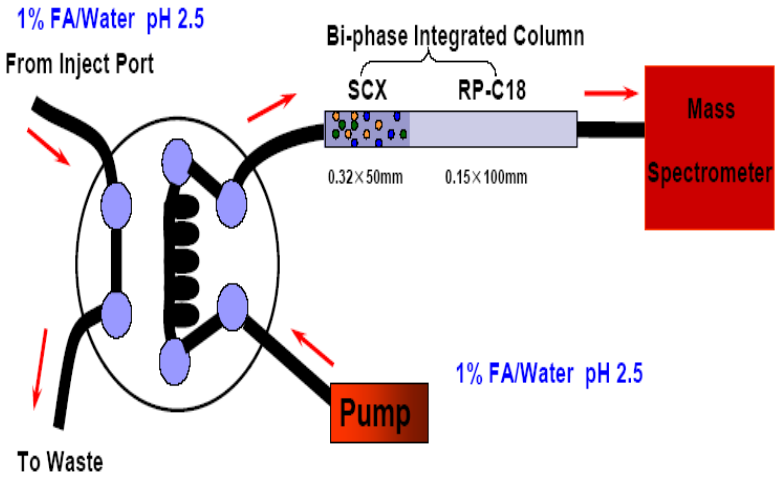
- Bi-phasic column contains both ion exchange and reverse phase material for the 2D LC separation.
- Inject buffer plugs by autosampler.
- Peptide fractionation controlled by pH buffers.

Integrated multi-dimensional LC (IMDL) based on pH elution

Sample Loading

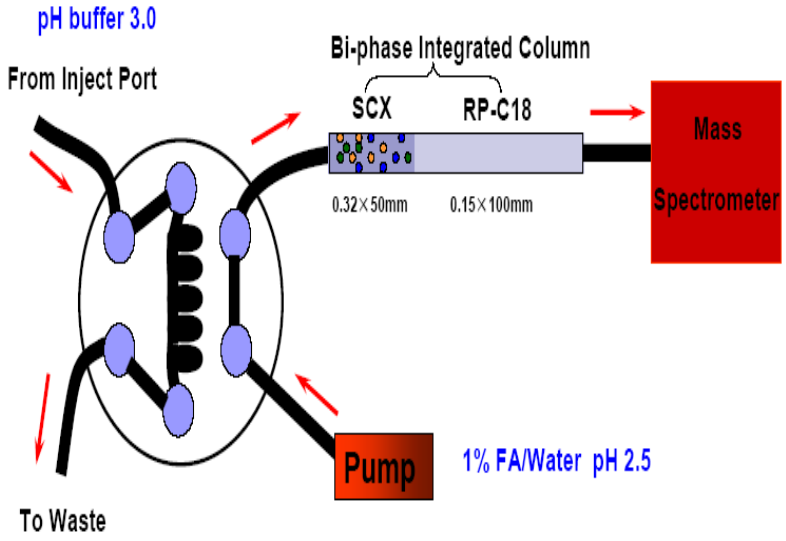


Sample Injection

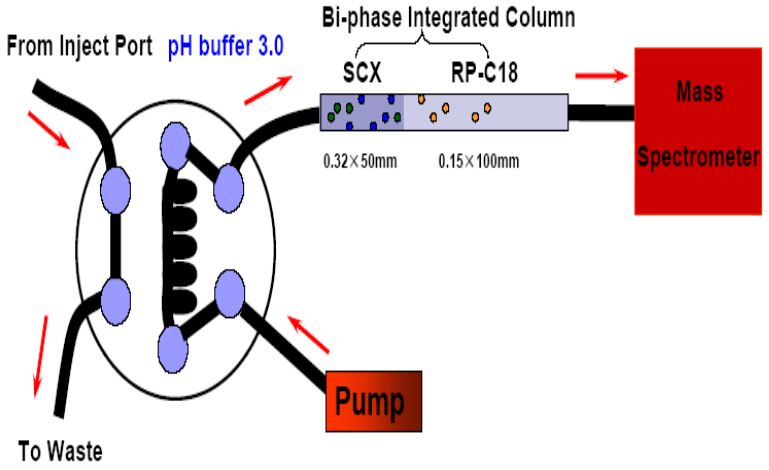


Integrated multi-dimensional LC (IMDL) based on pH elution

pH Buffer Loading 1

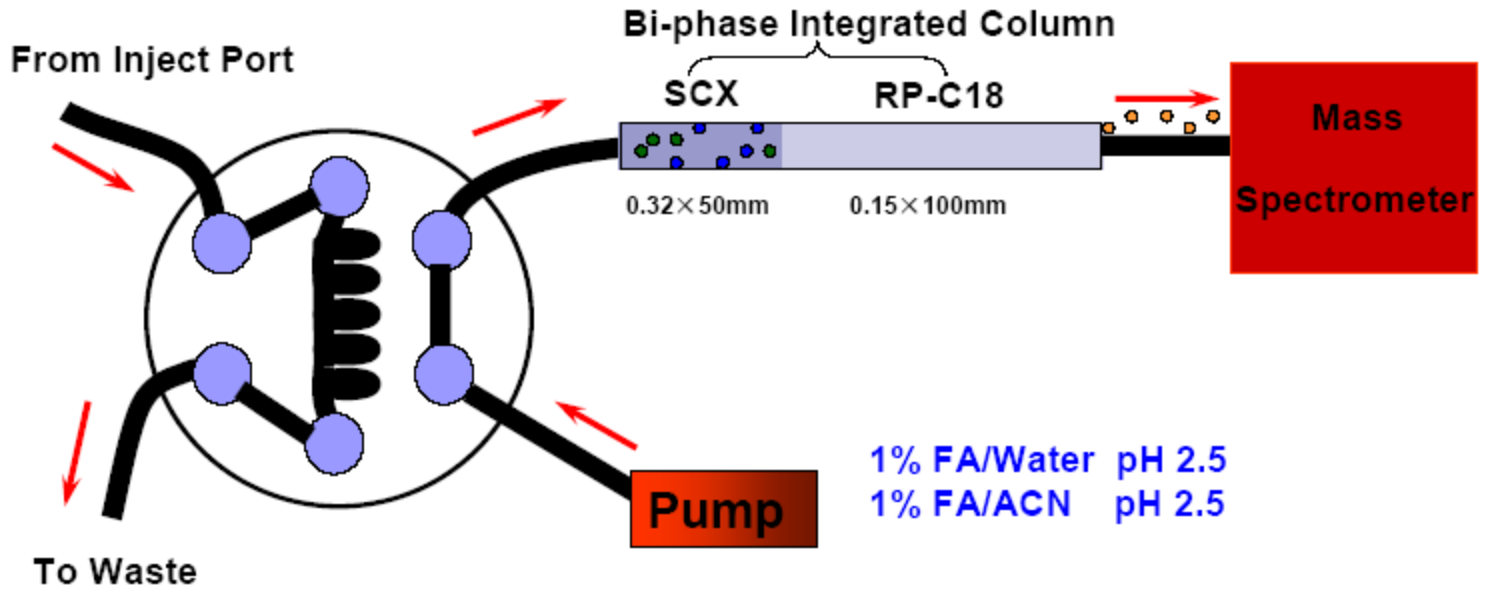


pH Buffer Elution 1



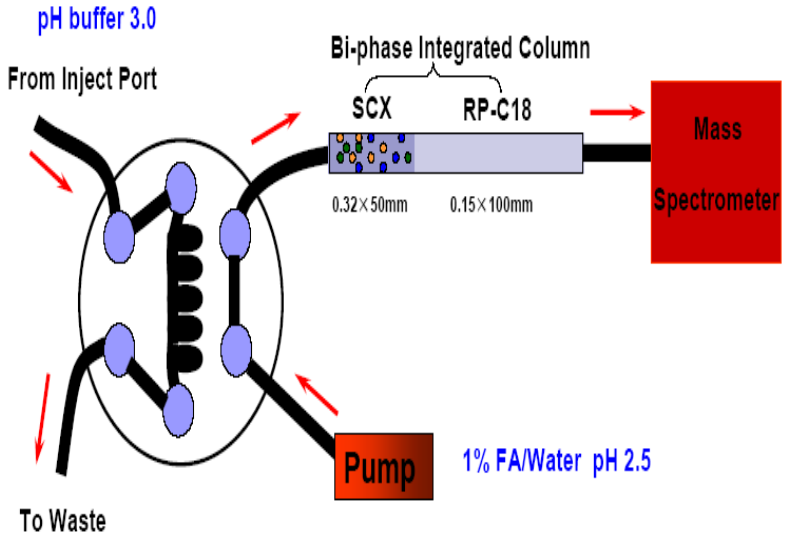
Integrated multi-dimensional LC (IMDL) based on pH elution

RP-Gradient Elution 1

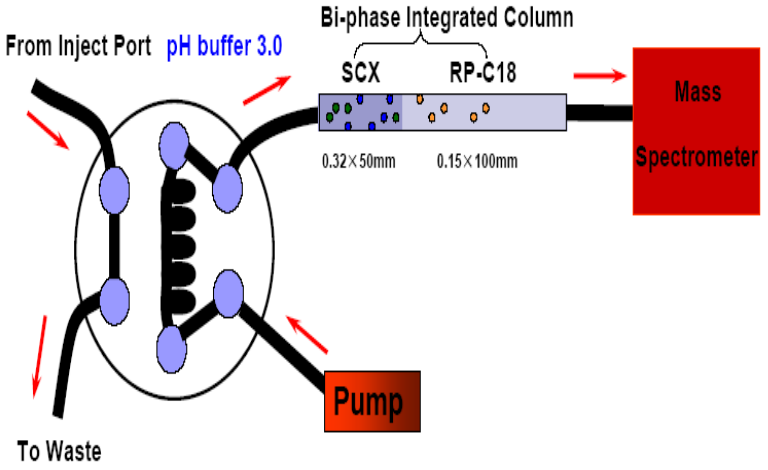


Integrated multi-dimensional LC (IMDL) based on pH elution

pH Buffer Loading 2



pH Buffer Elution 2



Repeat for other pH buffers

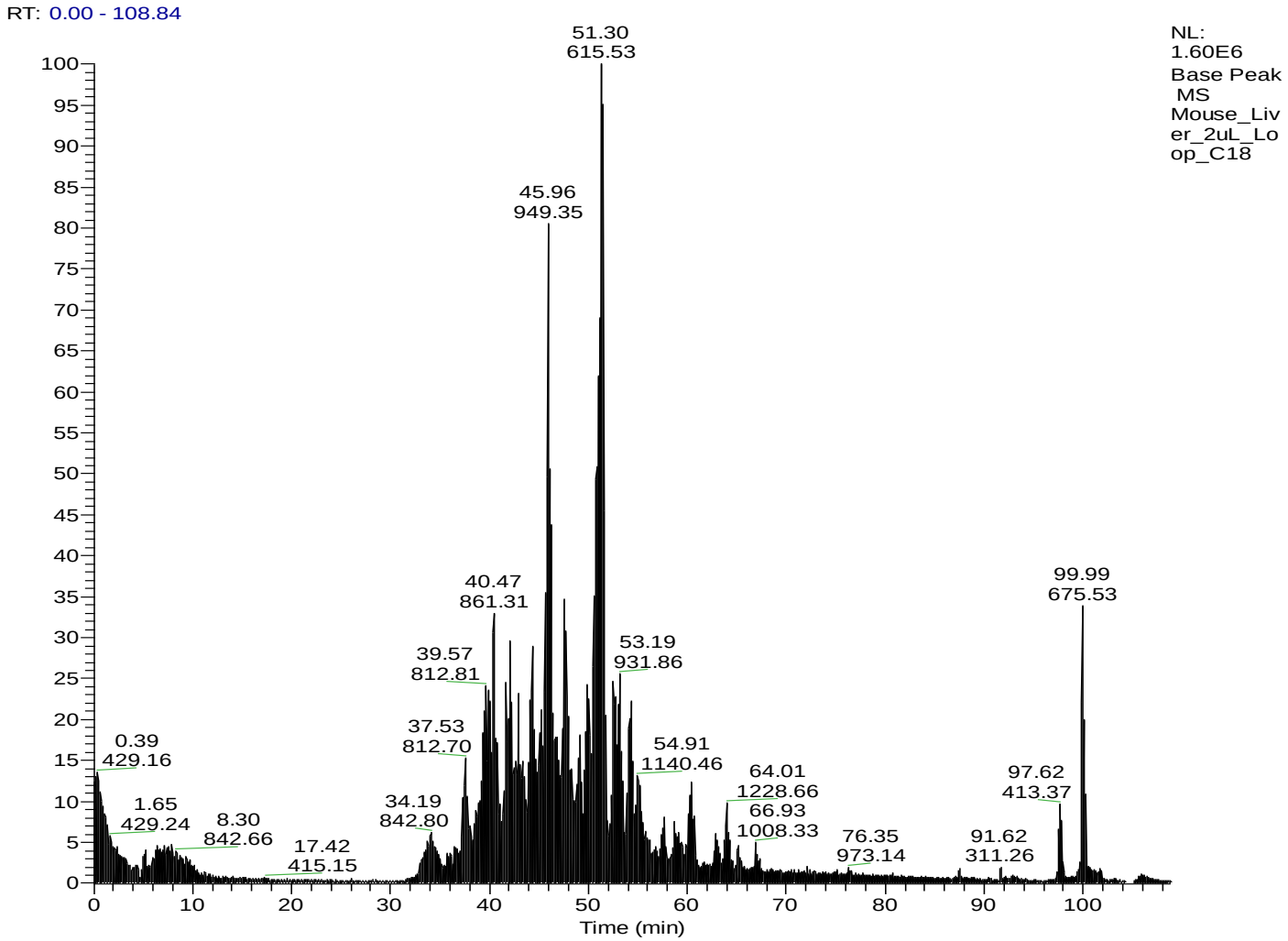
IMDL method

- Features
  - Two dimensional HPLC separation.
  - Peptide fractionated by pH step gradient followed by reverse phase chromatography (separation based on protein/peptide PIs and hydrophobicity).
  - Mass compatible buffers.
  - Only need one HPLC system, no need for additional switching valve or software.
  - No salt.
  - Large scale protein identification.
  - Works on any commercial LC/MS instrument.

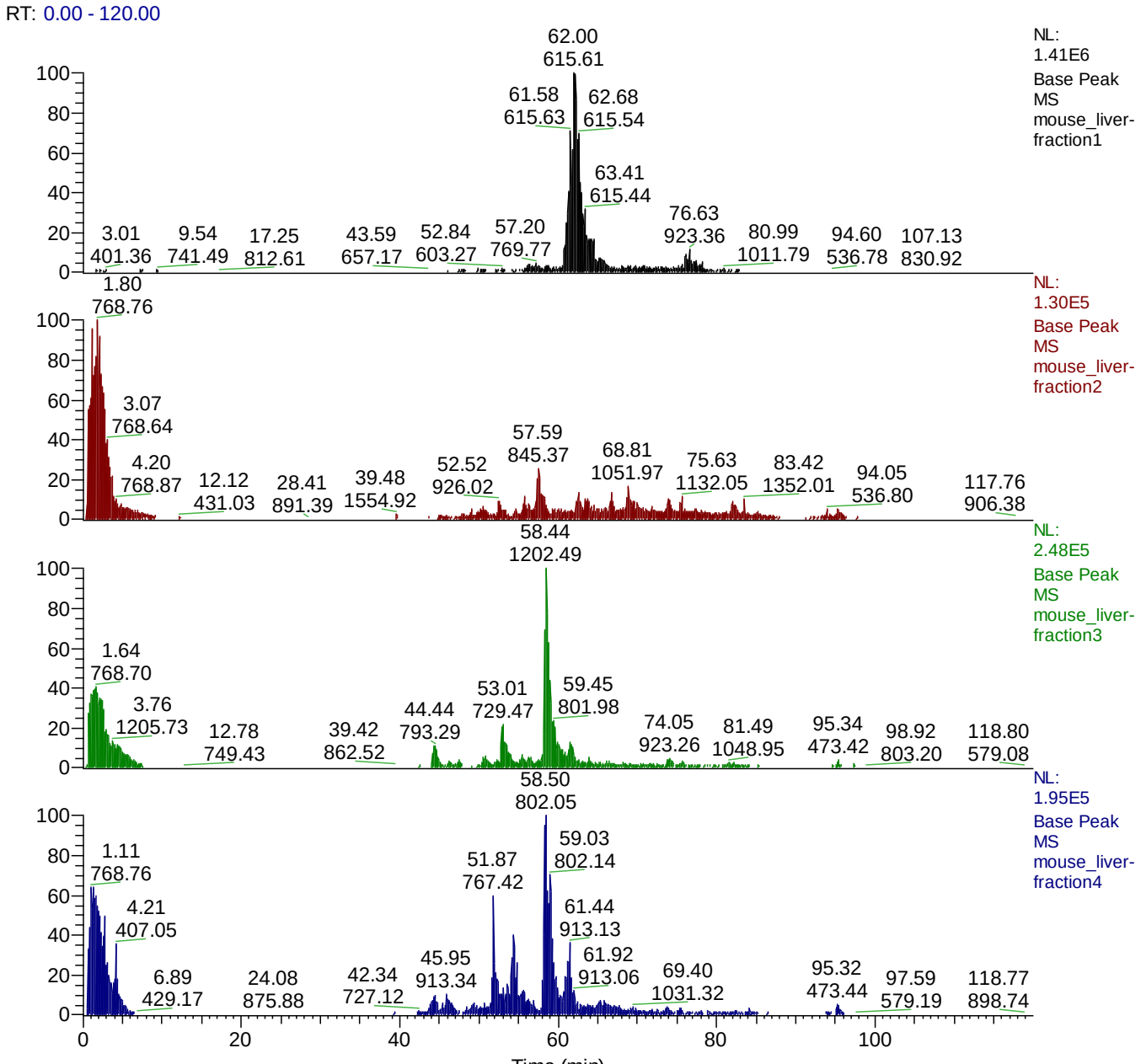
Mouse liver protein extraction

- SCX, 12 fractions
  - pH 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0
- Proteins were tryptic digested to peptides.
- Reverse phase mobile phase
  - A: 0.1% formic acid in water
  - B: 0.1% formic acid in acetonitrile

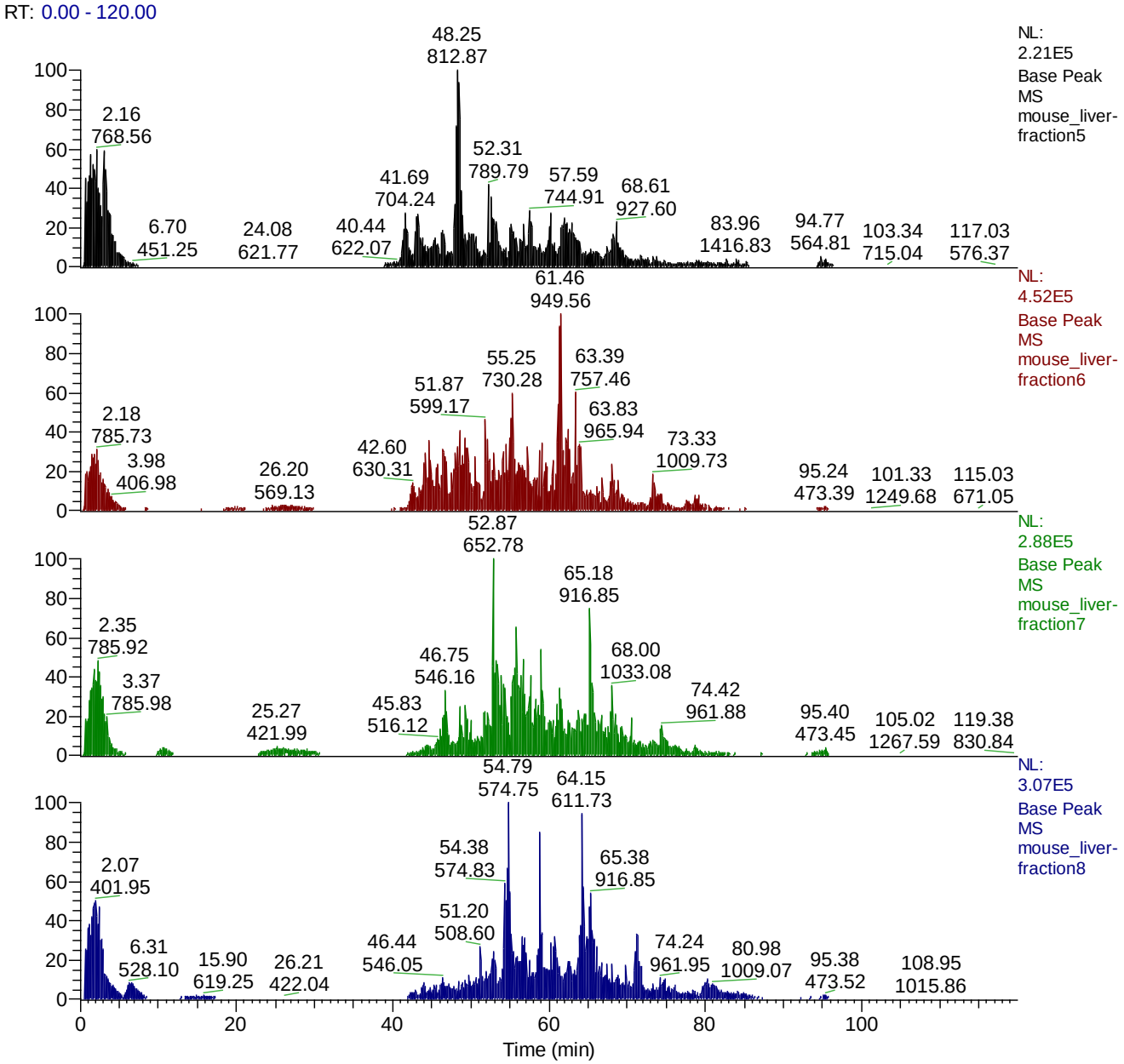
Mouse Liver 1D 100ug



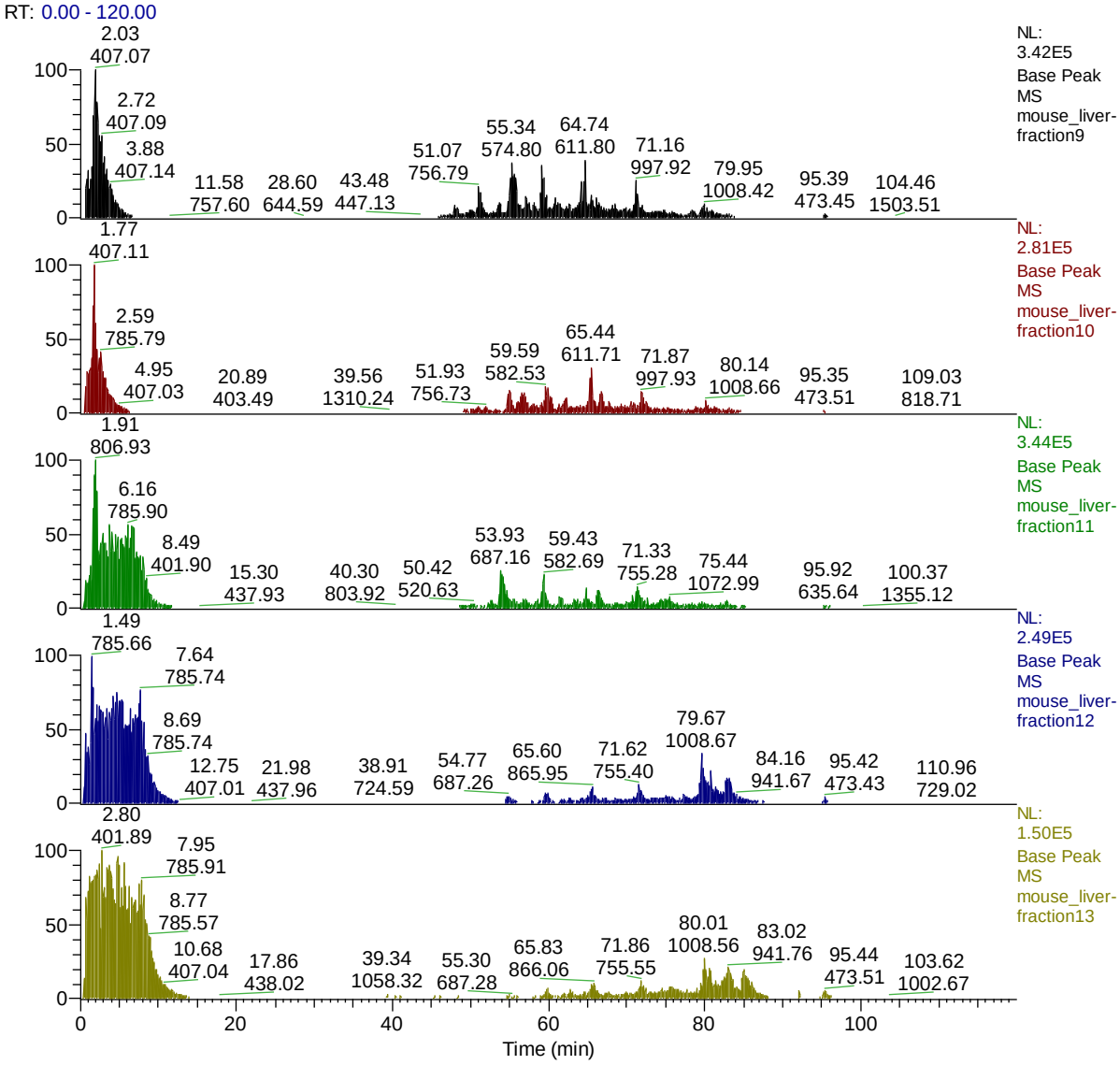
100ug sample, fraction 1-4



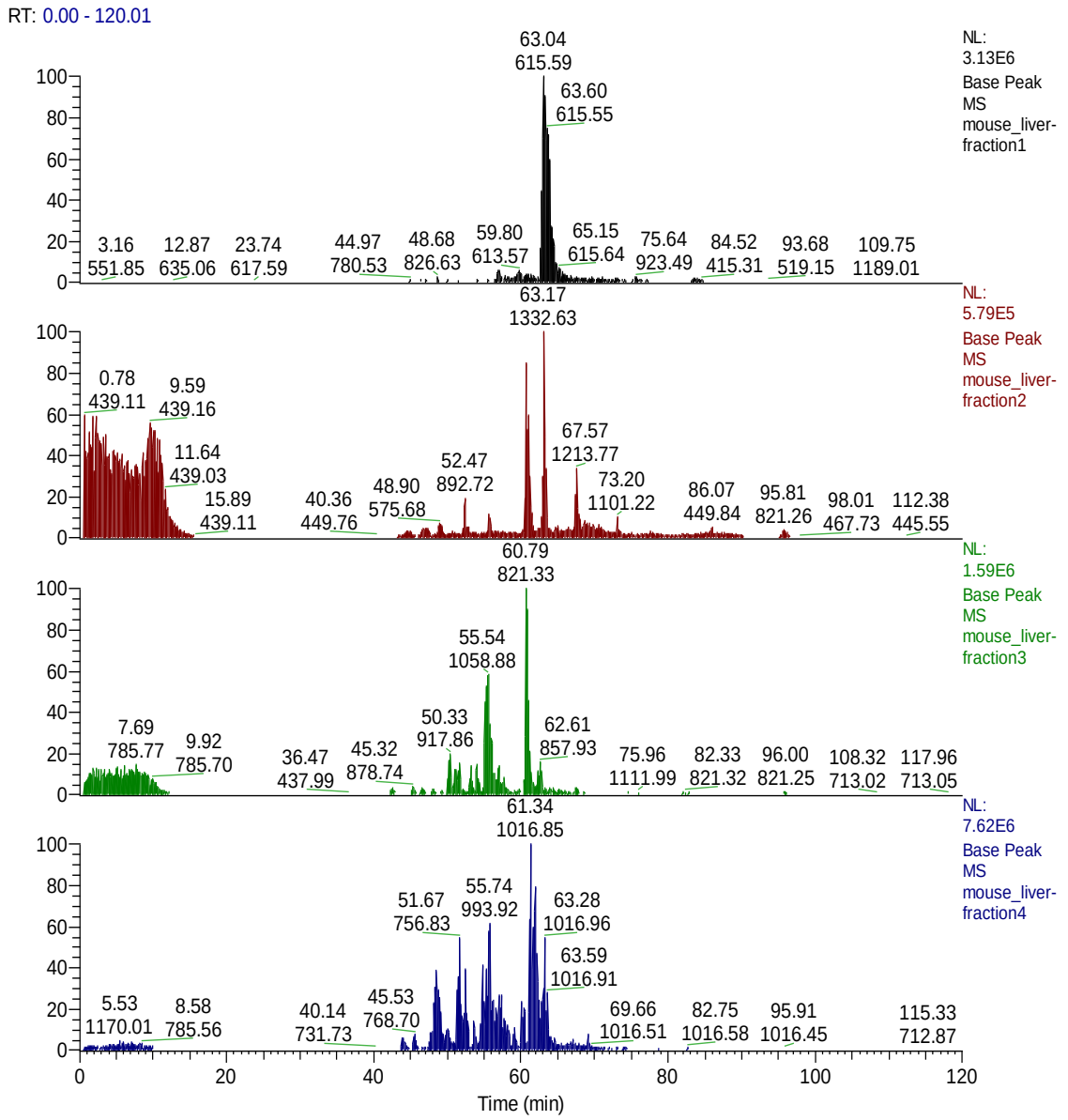
100ug sample, fraction 5-8



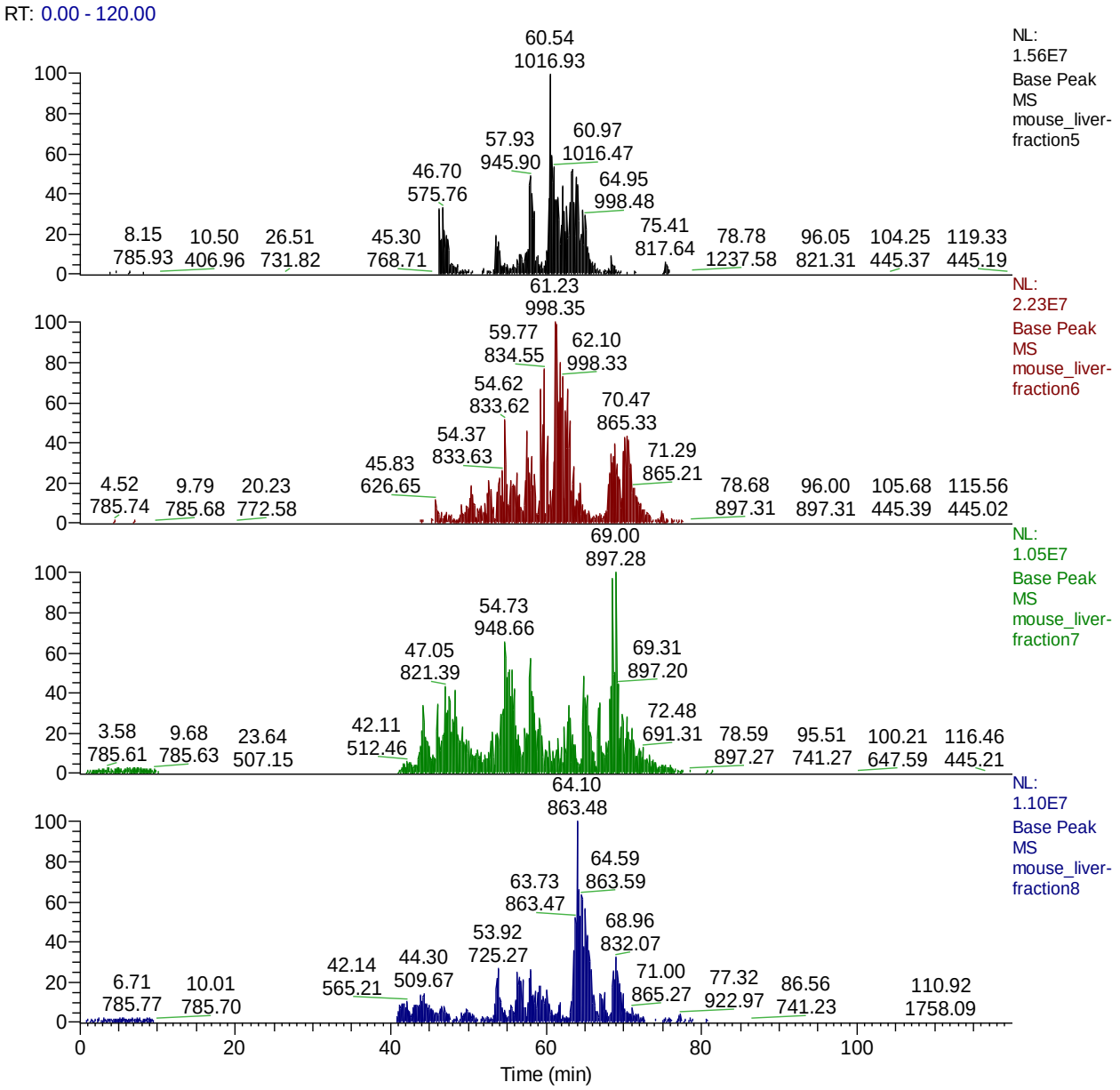
100ug sample, fraction 9-13



500ug sample, fraction 1-4



500ug sample, fraction 5-8



Result

- 100 ug
- 1D 1.9_2.2_3.75 <100 groups

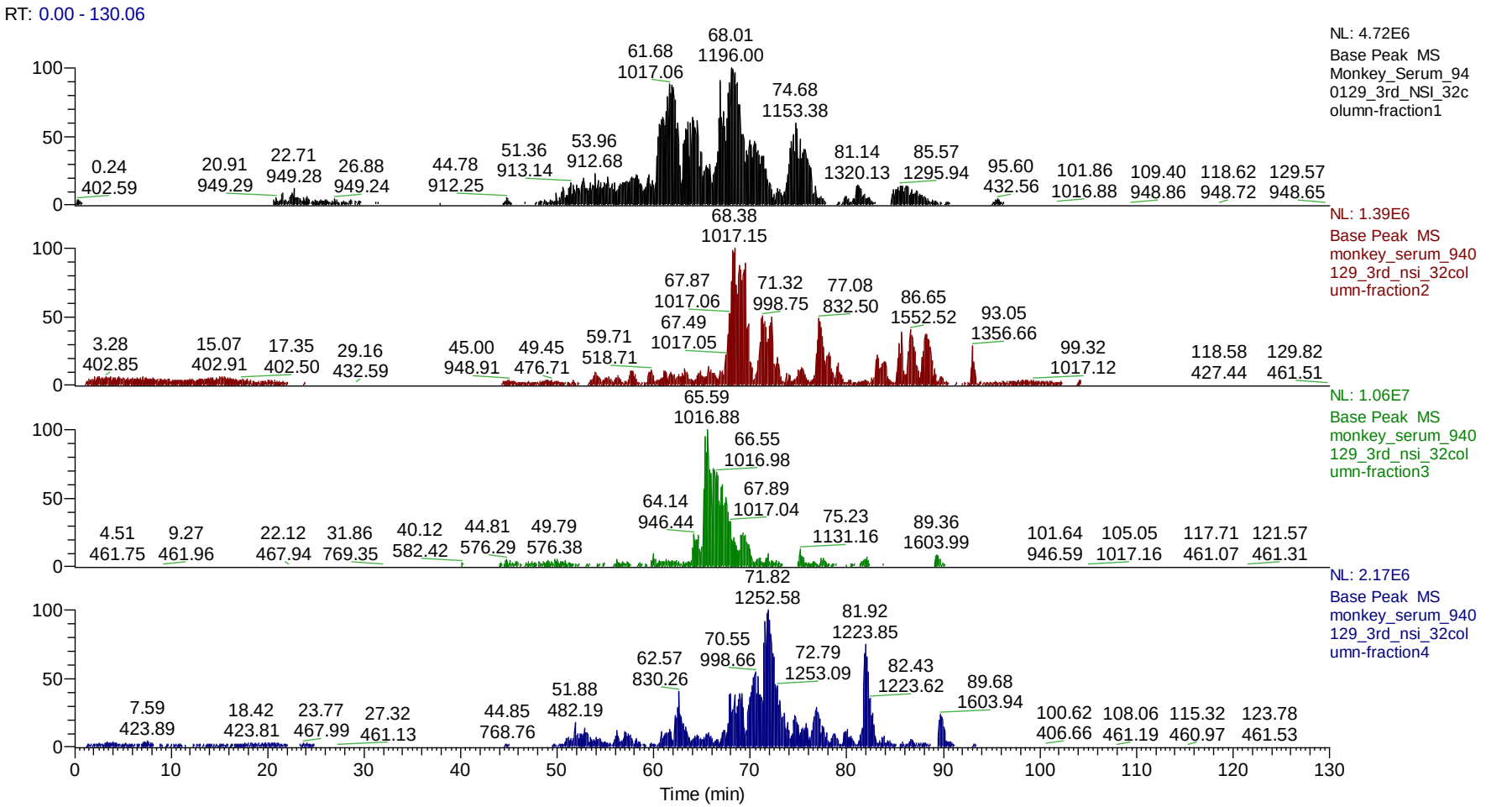
- 100 ug
- 2D 1.9_2.2_3.75 2099 groups

- 500 ug
- 2D 1.9_2.2_3.75 3364 groups

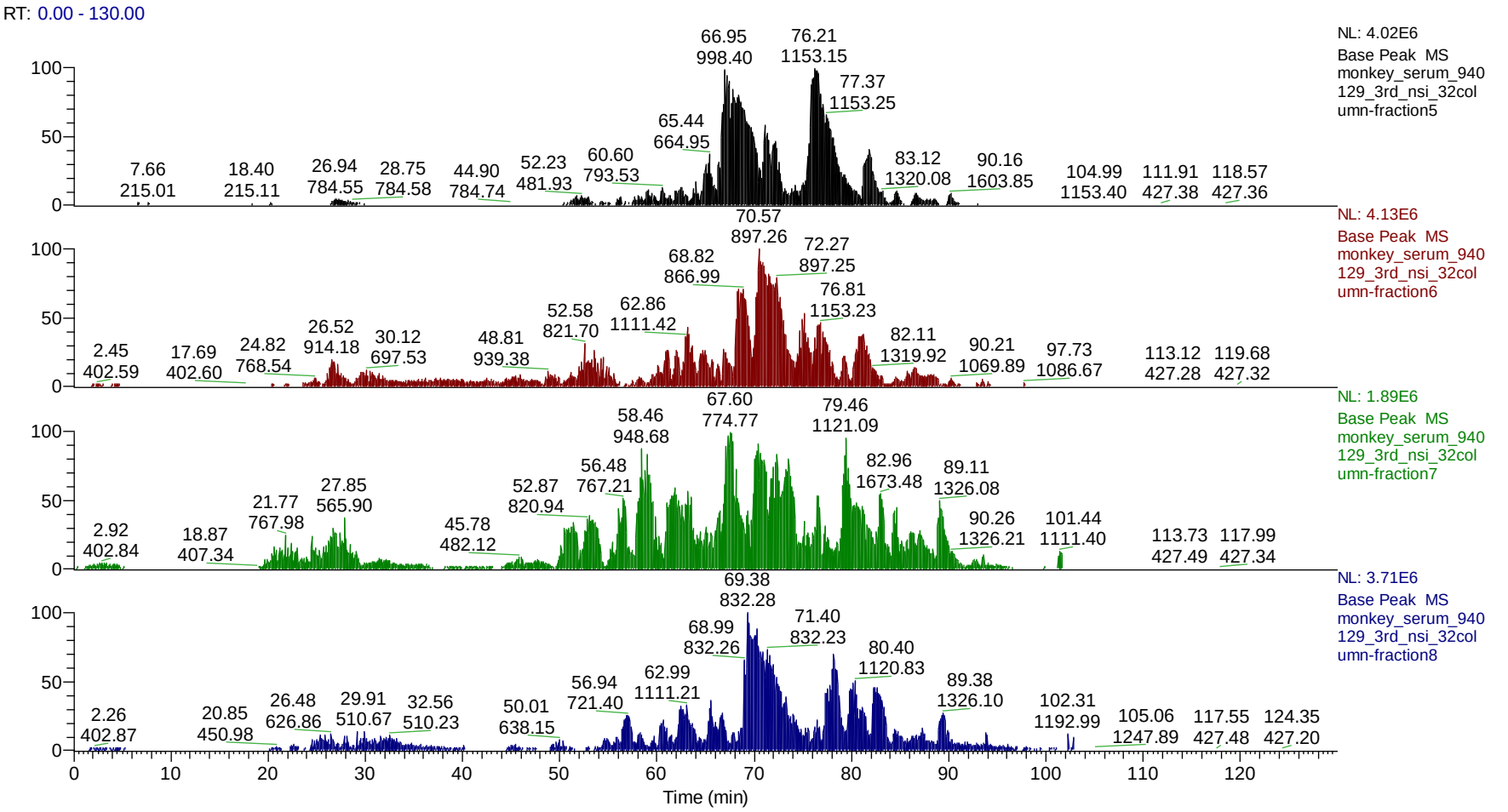
Separation of monkey serum

- 100 ul of serum was used for protein identification.
- No Albumin deletion.
- Proteins were tryptic digested to peptides.
- 10 PH buffer steps were used for the 2D separation.

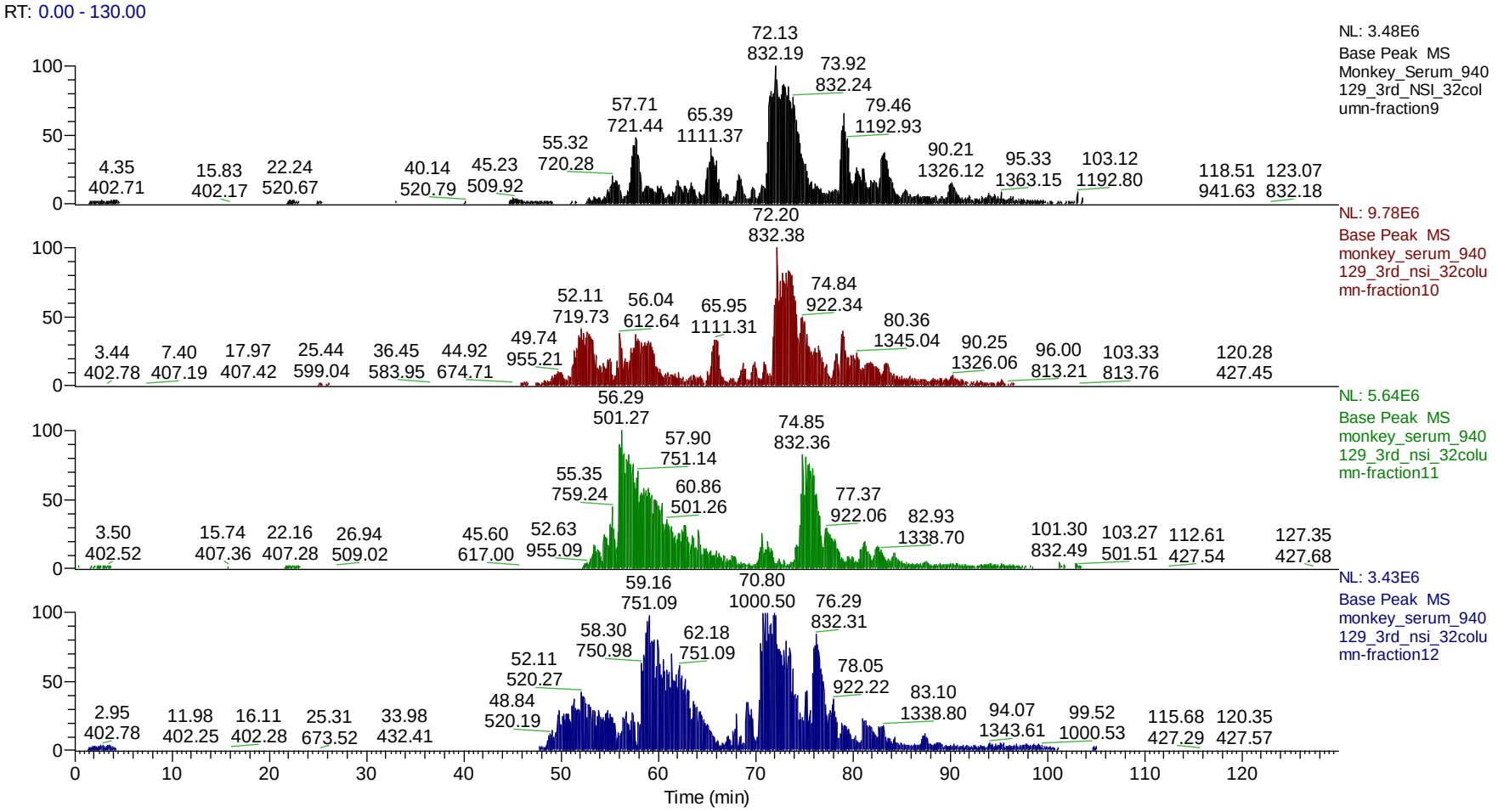
Monkey Serum 10 Fractions F1-F4



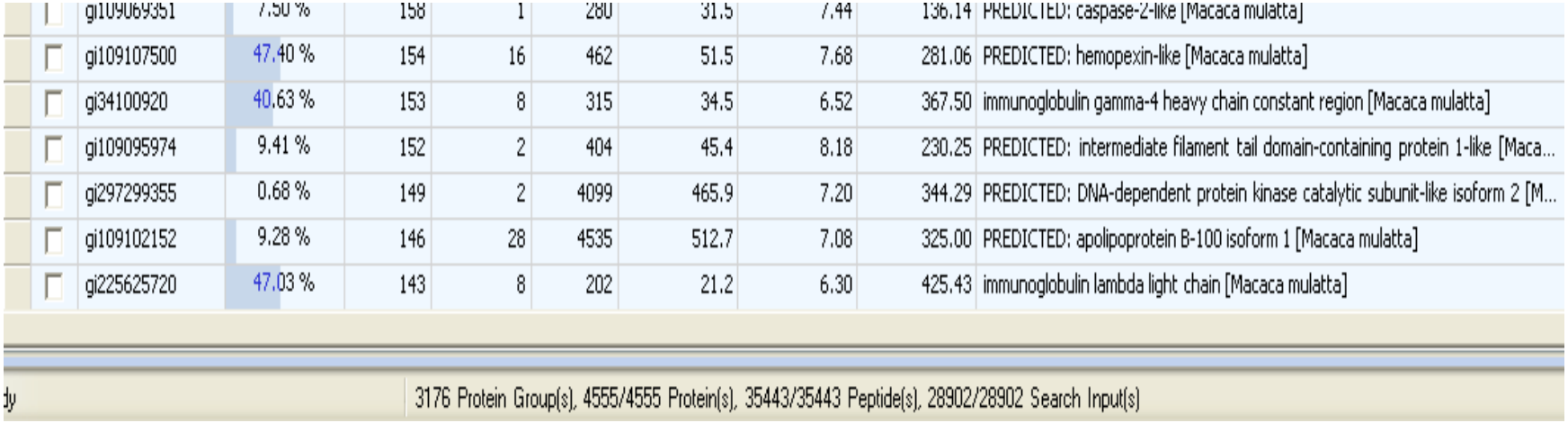
Monkey Serum 10 Fractions F5-F8



Monkey Serum 10 Fractions F9-F12



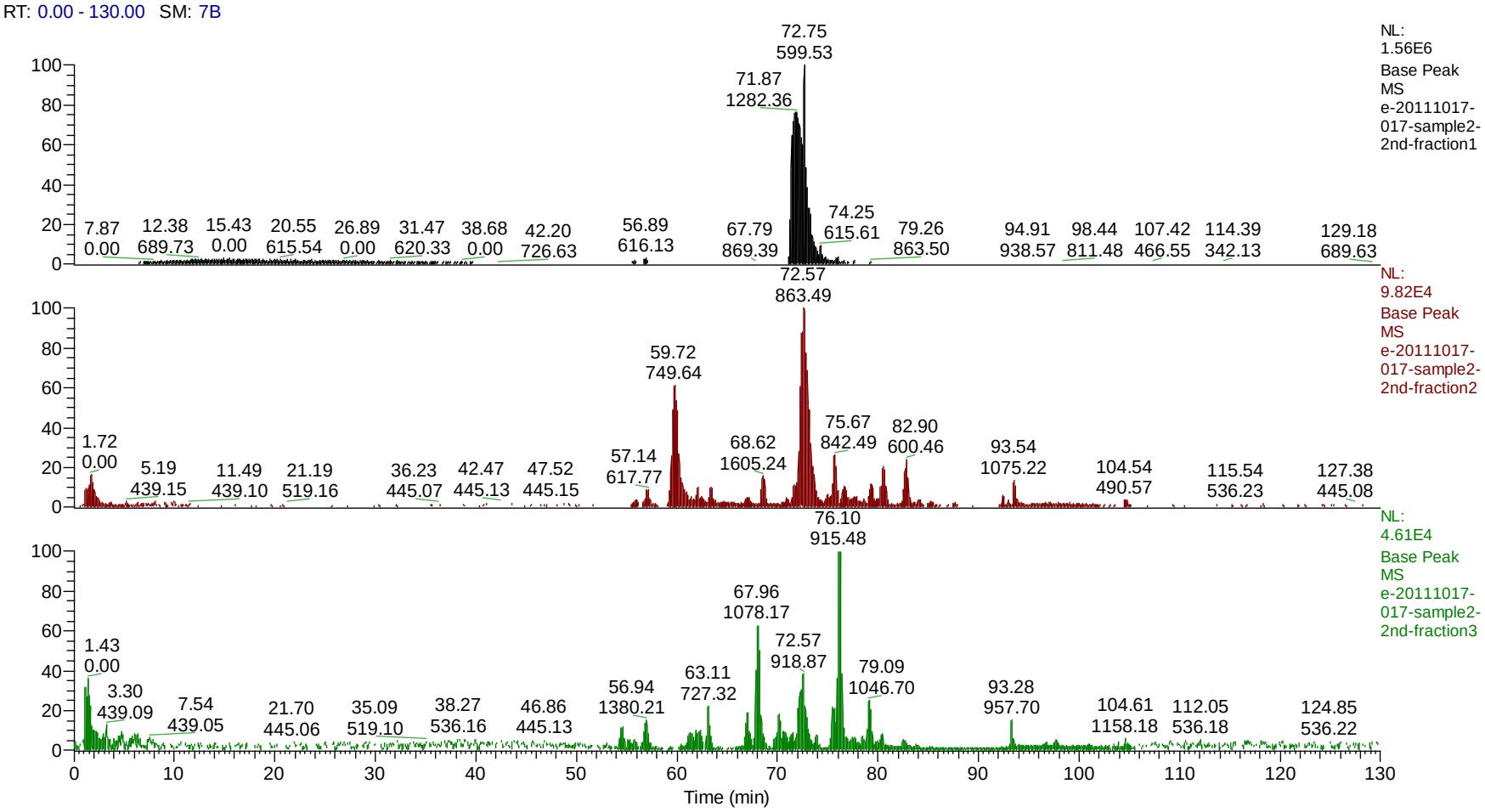
Filter by Xcor>1.9, 2.3 , 3.75
Protein identified: 3176 groups



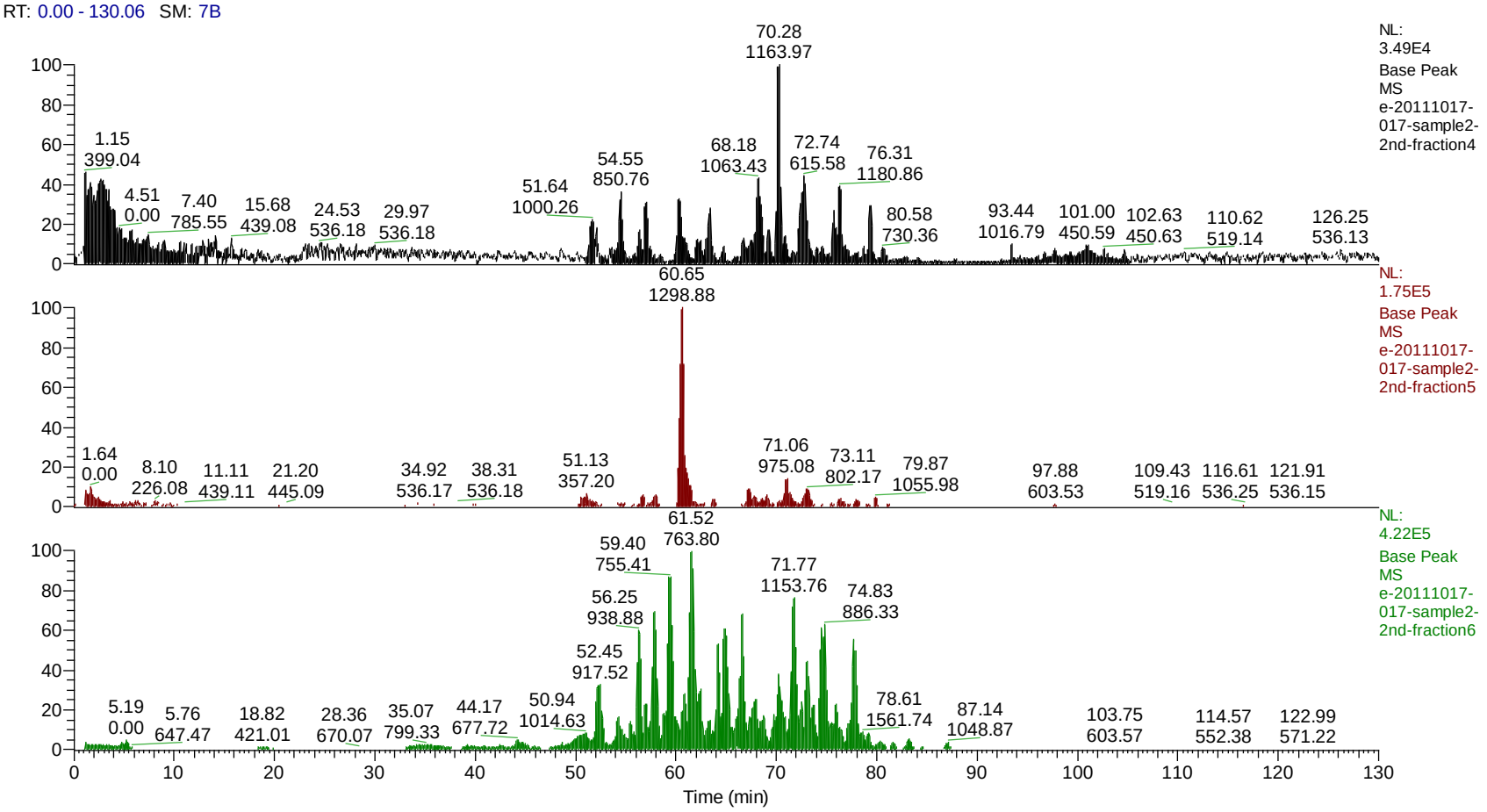
gi109069351	7.50 %	158	1	280	31.5	7.44	136.14	PREDICTED: caspase-2-like [Macaca mulatta]
gi109107500	47.40 %	154	16	462	51.5	7.68	281.06	PREDICTED: hemopexin-like [Macaca mulatta]
gi34100920	40.63 %	153	8	315	34.5	6.52	367.50	immunoglobulin gamma-4 heavy chain constant region [Macaca mulatta]
gi109095974	9.41 %	152	2	404	45.4	8.18	230.25	PREDICTED: intermediate filament tail domain-containing protein 1-like [Maca...
gi297299355	0.68 %	149	2	4099	465.9	7.20	344.29	PREDICTED: DNA-dependent protein kinase catalytic subunit-like isoform 2 [M...
gi109102152	9.28 %	146	28	4535	512.7	7.08	325.00	PREDICTED: apolipoprotein B-100 isoform 1 [Macaca mulatta]
gi225625720	47.03 %	143	8	202	21.2	6.30	425.43	immunoglobulin lambda light chain [Macaca mulatta]

3176 Protein Group(s), 4555/4555 Protein(s), 35443/35443 Peptide(s), 28902/28902 Search Input(s)

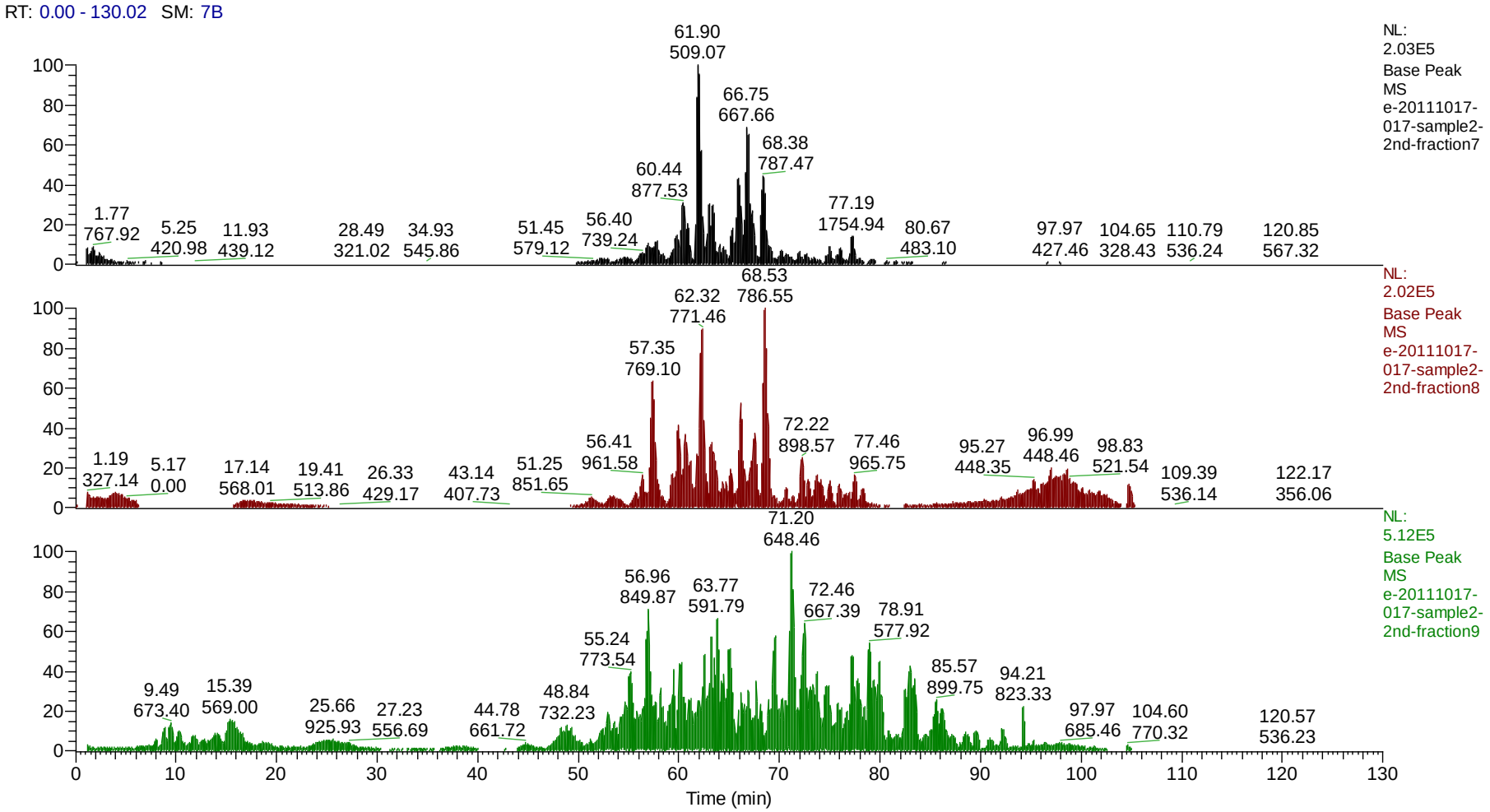
Microorganism (Fraction 1-3)



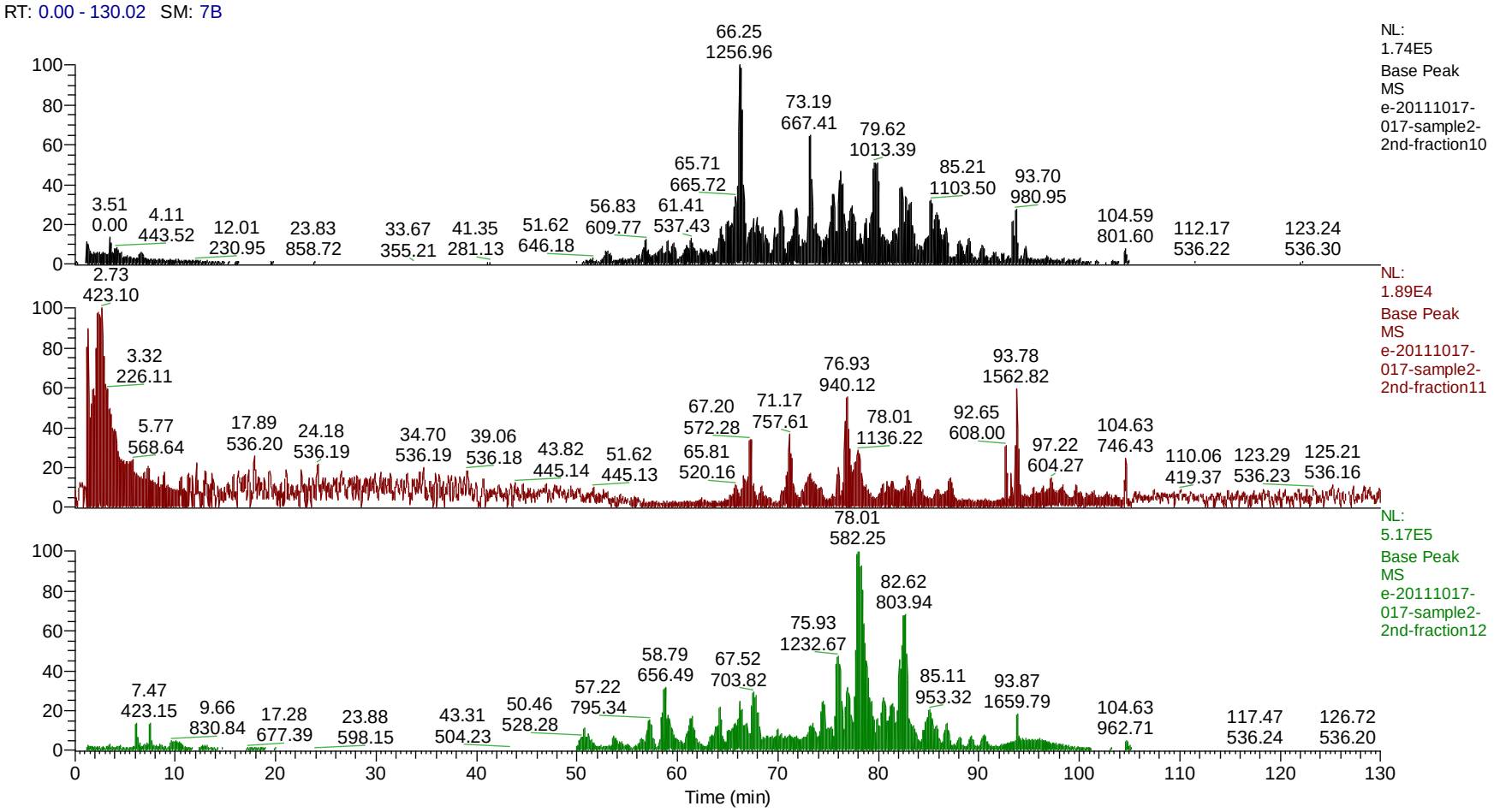
Microorganism (Fraction 4-6)



Microorganism (Fraction 7-9)



Microorganism (Fraction 10-12)



Filter by Xcor>1.9, 2.3, 3.75
Protein identified: 5148 groups

gi166860137	4.53 %	1	1	287	31.3	4.93	2.20	Dyp-type peroxidase family [Pseudomonas putida GB-1]
gi166859743	7.14 %	1	1	224	23.5	6.43	2.20	glutamate synthase alpha subunit domain protein [Pseudomonas putida GB-1]
gi148510744	6.09 %	1	1	197	21.4	10.07	2.20	phosphonate metabolism protein/1,5-bisphosphokinase (PRPP-forming) PhnN...
gi40019186	5.51 %	1	1	254	27.8	5.99	2.20	putative protein-disulfide isomerase [Pseudomonas putida]
gi148511283	4.40 %	1	1	273	31.6	6.44	2.20	MCP methyltransferase, CheR-type [Pseudomonas putida F1]
gi26987341	7.39 %	1	1	176	19.7	8.72	2.20	lipoprotein signal peptidase [Pseudomonas putida KT2440]

eady

5148 Protein Group(s), 10751/10751 Protein(s), 46943/46943 Peptide(s), 39938/39938 Search Input(s)

Protein identification from pig reproductive organs

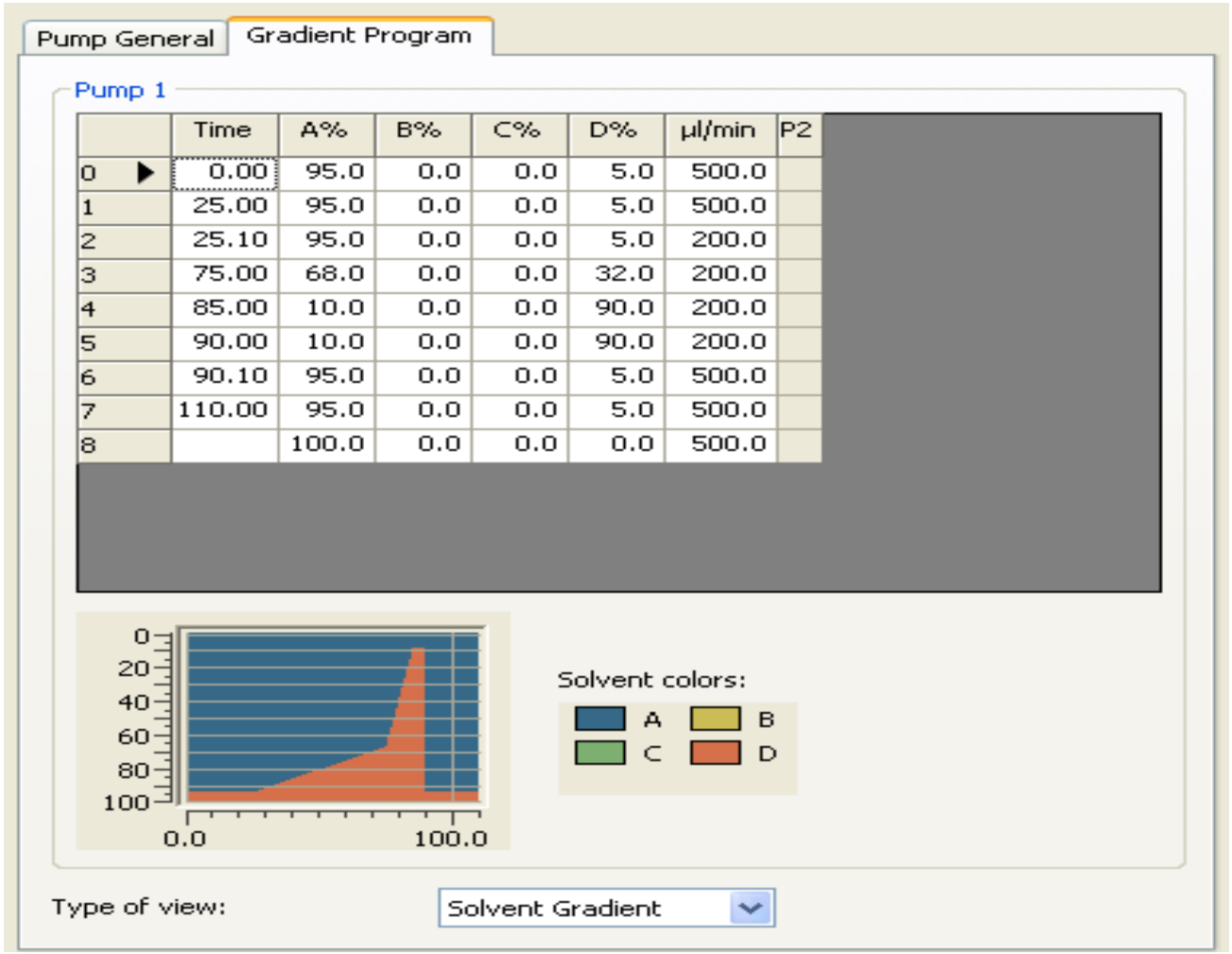
Worklist

Sequencing method with 9 pH steps.

42	E-20120320-021-MB1-fraction1	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	c:5
43	E-20120320-021-MB1-fraction2	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	c:22
44	E-20120320-021-MB1-fraction3	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:23
45	E-20120320-021-MB1-fraction4	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	c:24
46	E-20120320-021-MB1-fraction5	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:25
47	E-20120320-021-MB1-fraction6	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:26
48	E-20120320-021-MB1-fraction7	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:28
49	E-20120320-021-MB1-fraction8	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:30
50	E-20120320-021-MB1-fraction9	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:32
51	E-20120320-021-MB1-fraction10-blank-end	E:\data\20120323	D:\methods\MudPIT\2D_Peptides_ESI_110min_20120322	100.0	C:32

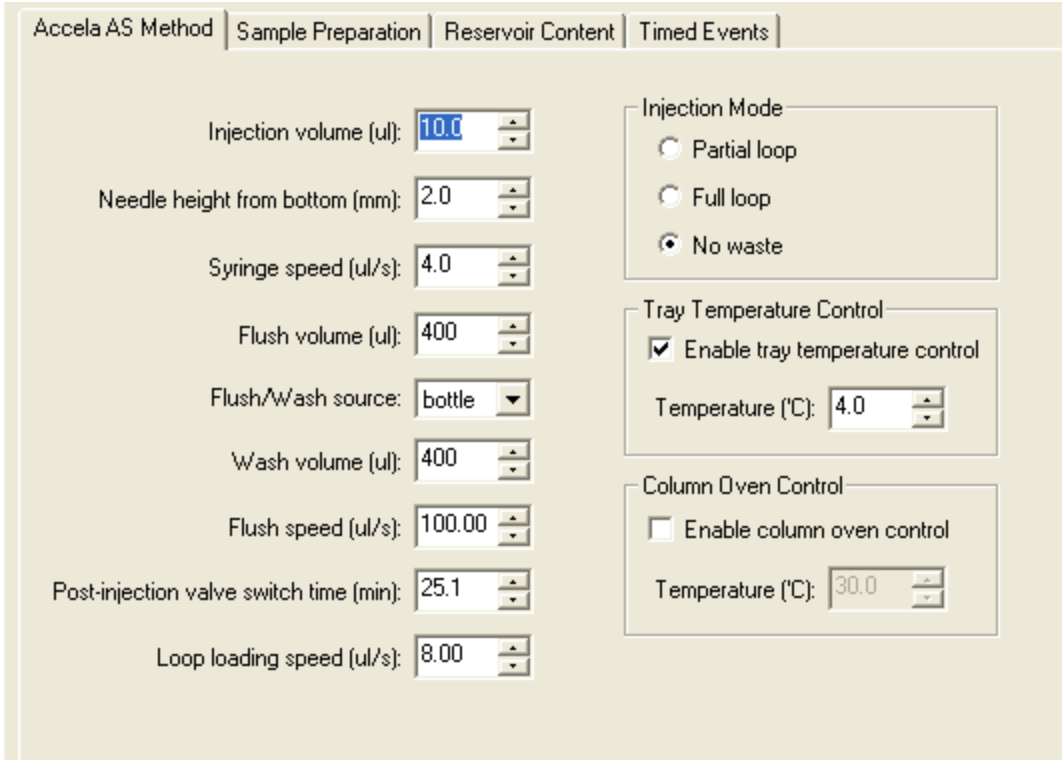
Gradient profile

Higher flow rate used during buffer loading or column equilibration step
Flow split at 1:100 ratio.

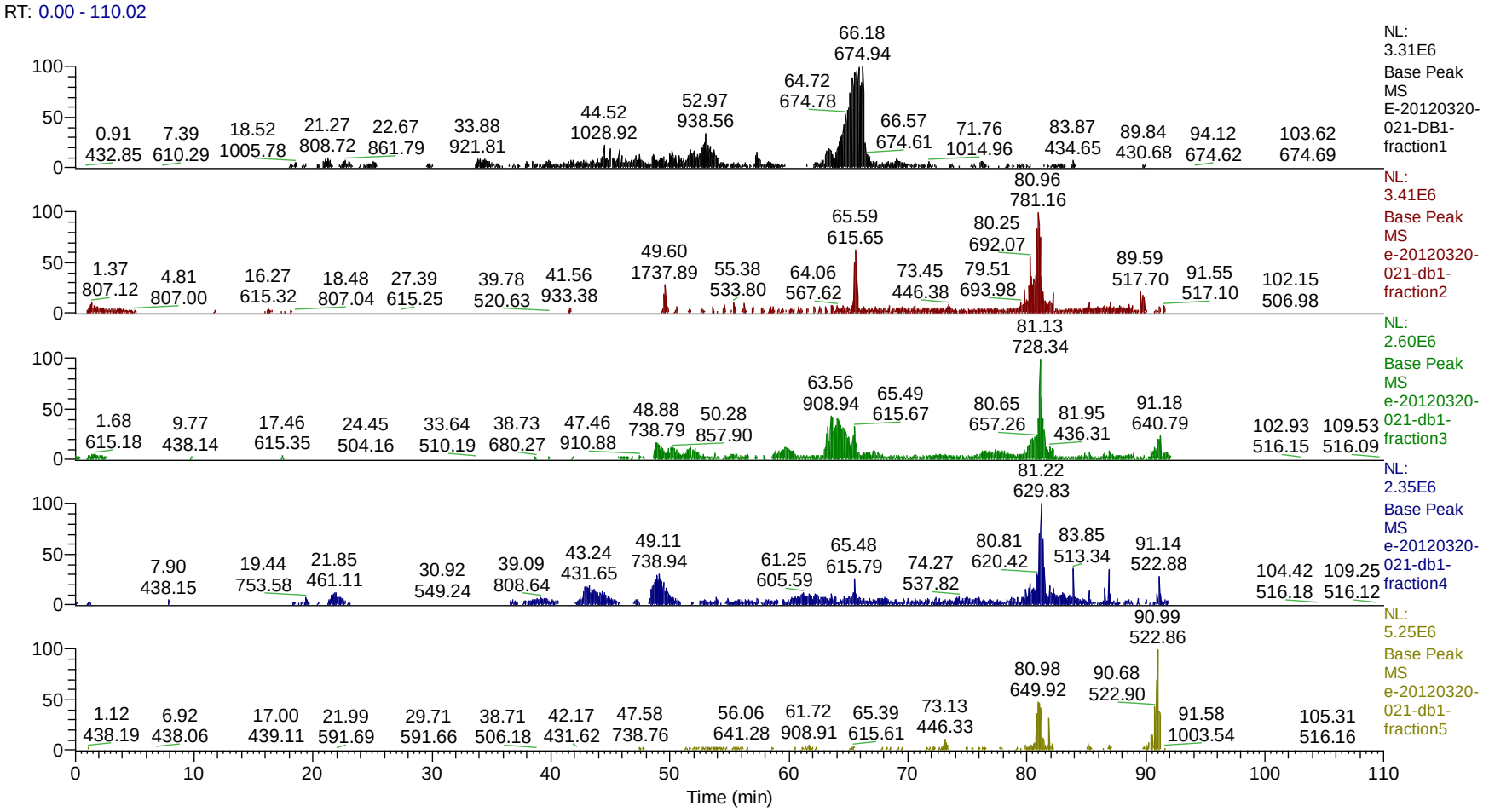


Autosampler

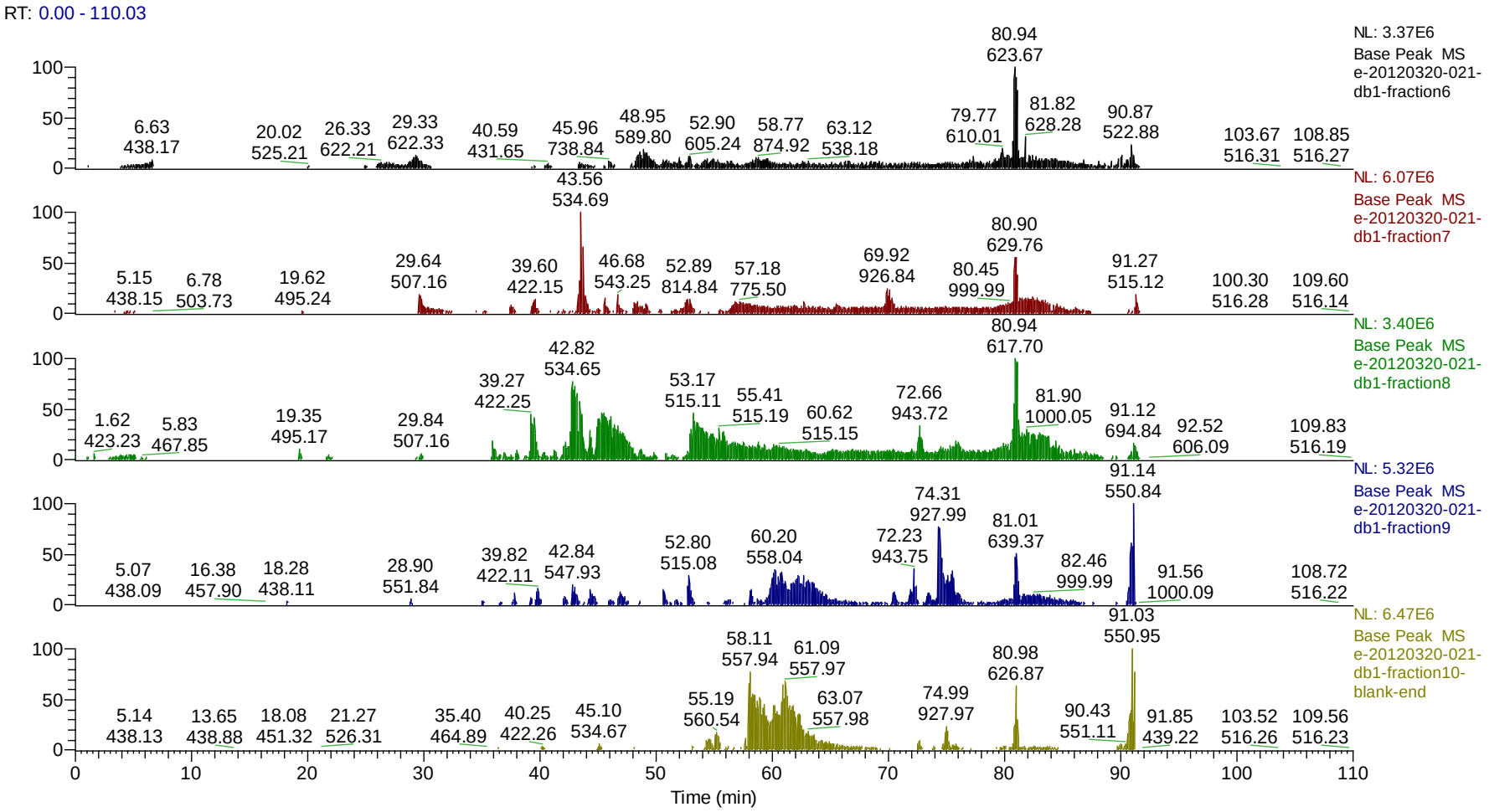
Injection valve switched off at 25.1 min to reduce gradient delay volume.



F1-F5


RT: 0.00 - 110.02
NL: 3.31E6 Base Peak MS E-20120320-021-DB1-fraction1
NL: 3.41E6 Base Peak MS e-20120320-021-db1-fraction2
NL: 2.60E6 Base Peak MS e-20120320-021-db1-fraction3
NL: 2.35E6 Base Peak MS e-20120320-021-db1-fraction4
NL: 5.25E6 Base Peak MS e-20120320-021-db1-fraction5
Time (min)

F6-F10


RT: 0.00 - 110.03
NL: 3.37E6
Base Peak MS
e-20120320-021-
db1-fraction6
80.94
623.67
6.63
438.17
20.02
525.21
26.33
622.21
29.33
622.33
40.59
431.65
45.96
738.84
48.95
589.80
52.90
605.24
58.77
874.92
63.12
538.18
79.77
610.01
81.82
628.28
90.87
522.88
103.67
516.31
108.85
516.27
NL: 6.07E6
Base Peak MS
e-20120320-021-
db1-fraction7
43.56
534.69
80.90
629.76
5.15
438.15
6.78
503.73
19.62
495.24
29.64
507.16
39.60
422.15
46.68
543.25
52.89
814.84
57.18
775.50
69.92
926.84
80.45
999.99
91.27
515.12
100.30
516.28
109.60
516.14
NL: 3.40E6
Base Peak MS
e-20120320-021-
db1-fraction8
80.94
617.70
42.82
534.65
39.27
422.25
1.62
423.23
5.83
467.85
19.35
495.17
29.84
507.16
53.17
515.11
55.41
515.19
60.62
515.15
72.66
943.72
81.90
1000.05
91.12
694.84
92.52
606.09
109.83
516.19
NL: 5.32E6
Base Peak MS
e-20120320-021-
db1-fraction9
91.14
550.84
74.31
927.99
81.01
639.37
72.23
943.75
5.07
438.09
16.38
457.90
18.28
438.11
28.90
551.84
39.82
422.11
42.84
547.93
52.80
515.08
60.20
558.04
82.46
999.99
91.56
1000.09
108.72
516.22
NL: 6.47E6
Base Peak MS
e-20120320-021-
db1-fraction10-
blank-end
91.03
550.95
58.11
557.94
61.09
557.97
80.98
626.87
55.19
560.54
63.07
557.98
74.99
927.97
5.14
438.13
13.65
438.88
18.08
451.32
21.27
526.31
35.40
464.89
40.25
422.26
45.10
534.67
90.43
551.11
91.85
439.22
103.52
516.26
109.56
516.23
100
50
0
0
10
20
30
40
50
60
70
80
90
100
110
Time (min)

Xcor
7018 groups

Filter and Protein Grouping Settings

Charge State versus Score-Value

Peptide Score:	SequestNode (XCorr)
Minimal XCorr-Value for Charge State = 1:	1.9
Minimal XCorr-Value for Charge State = 2:	2.2
Minimal XCorr-Value for Charge State = 3:	3.75

DB1.msf

Proteins | Peptides | Search Input | Result Filters | Peptide Confidence | Search Summary

Accession	Coverage	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
gi335306030	70.34 %	2732	54	863	95.2	6.84	4385.43	PREDICTED: A-kinase anchor protein 4-like [Sus scrofa]
gi72535165	72.26 %	1743	7	137	15.0	8.88	3144.86	seminal plasma sperm motility inhibitor precursor [Sus scrofa]
gi311273025	40.42 %	1504	59	2177	239.6	6.05	2746.15	PREDICTED: fibronectin isoform 3 [Sus scrofa]
gi7545446	62.04 %	1346	6	137	15.2	9.06	2474.56	seminal plasma sperm motility inhibitor/spermadhesin AQN-3-like protein [Sus...
gi301173563	36.79 %	1036	30	810	93.2	7.97	1912.61	outer dense fiber protein 2 [Sus scrofa]
gi7645971	8.42 %	1012	1	202	21.9	5.02	598.63	BiP [Sus scrofa]
gi237798	72.41 %	778	7	116	12.8	8.68	1485.34	AQN-3=carbohydrate-binding protein [swine, Peptide, 116 aa]
gi3318759	70.69 %	717	6	116	12.6	8.44	2095.89	Chain B, The Crystal Structures Of Two Members Of The Spermadhesin Family...
gi335281298	73.48 %	615	19	445	49.8	4.89	950.31	PREDICTED: tubulin beta-2C chain-like [Sus scrofa]
gi3745822	63.93 %	586	16	427	47.9	5.31	874.49	Chain B, Tubulin Alpha-Beta Dimer, Electron Diffraction
gi350580622	64.19 %	564	16	444	49.6	4.88	819.76	PREDICTED: tubulin beta-4 chain-like [Sus scrofa]
gi343478189	58.65 %	512	15	445	49.9	4.89	788.76	tubulin beta-2B chain [Sus scrofa]

Ready | 7018 Protein Group(s), 9771/9771 Protein(s), 52241/52241 Peptide(s), 72651/72651 Search Input(s)

Xcor & PSMS>=2
4524 groups

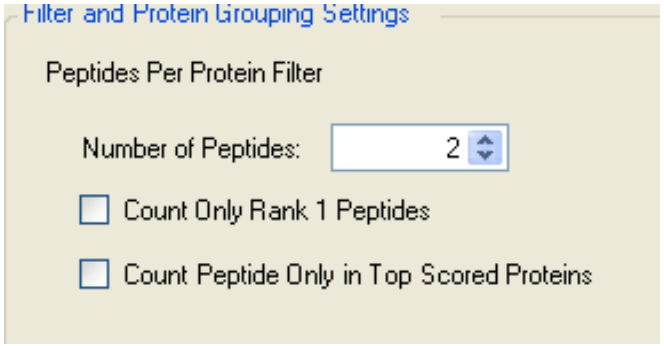
Filter and Protein Grouping Settings

Charge State versus Score-Value

Peptide Score:	SequestNode (XCorr)
Minimal XCorr-Value for Charge State = 1:	1.9
Minimal XCorr-Value for Charge State = 2:	2.2
Minimal XCorr-Value for Charge State = 3:	3.75

	Accession	Coverage	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
2	gi335306030	70.34	2732	54	863	95.2	6.84	4385.43	PREDICTED: A-kinase anchor protein 4-like [Sus scrofa]
3	gi72535165	72.26	1743	7	137	15.0	8.88	3144.86	seminal plasma sperm motility inhibitor precursor [Sus scrofa]
4	gi311273025	40.42	1504	59	2177	239.6	6.05	2746.15	PREDICTED: fibronectin isoform 3 [Sus scrofa]
5	gi7545446	62.04	1346	6	137	15.2	9.06	2474.56	seminal plasma sperm motility inhibitor/spermadhesin AQN-3-like protein [Sus scrofa]
6	gi301173563	36.79	1036	30	810	93.2	7.97	1912.61	outer dense fiber protein 2 [Sus scrofa]
7	gi7645971	8.42	1012	1	202	21.9	5.02	598.63	BiP [Sus scrofa]
8	gi237798	72.41	778	7	116	12.8	8.68	1485.34	AQN-3=carbohydrate-binding protein [swine, Peptide, 116 aa]
9	gi3318759	70.69	717	6	116	12.6	8.44	2095.89	Chain B, The Crystal Structures Of Two Members Of The Spermadhesin Family Reveal The Folding Of The Cub Domain
10	gi335281298	73.48	615	19	445	49.8	4.89	950.31	PREDICTED: tubulin beta-2C chain-like [Sus scrofa]
4519	gi157830532	23.08	2	1	78	8.8	4.84	4.97	Chain A, Three-Dimensional Solution Structure Of Ca2+-Loaded Porcine Calbindin D
4520	gi353260842	2.48	2	1	524	55.1	5.63	20.40	mitochondrial IFN-beta promoter stimulator 1 [Sus scrofa domesticus]
4521	gi89230	36.67	2	1	30	3.2	4.91	9.75	lipocortin I - pig (fragment)
4522	gi1090093	13.48	2	1	141	15.5	8.82	7.69	17beta-estradiol dehydrogenase
4523	gi227130	40.00	2	2	85	10.0	8.48	4.74	chymodenin
4524	gi223522	38.89	2	2	54	5.4	9.00	14.99	opiomelanocortin N term G1,pro

Xcor & Peptides per Protein>=2
2597 groups



	Accession	Coverage	# PSMs	# Peptides	# AAs	MW [kDa]	calc. pI	Score	Description
☐	gi335306030	70.34 %	2732	54	863	95.2	6.84	4385.43	PREDICTED: A-kinase anchor protein 4-like [Sus scrofa]
☐	gi72535165	72.26 %	1743	7	137	15.0	8.88	3144.86	seminal plasma sperm motility inhibitor precursor [Sus scrofa]
☐	gi311273025	40.42 %	1504	59	2177	239.6	6.05	2746.15	PREDICTED: fibronectin isoform 3 [Sus scrofa]
☐	gi7545446	62.04 %	1346	6	137	15.2	9.06	2474.56	seminal plasma sperm motility inhibitor/spermadhesin AQN-3-like protein [Sus...
☐	gi301173563	36.79 %	1036	30	810	93.2	7.97	1912.61	outer dense fiber protein 2 [Sus scrofa]
☐	gi237798	72.41 %	778	7	116	12.8	8.68	1485.34	AQN-3=carbohydrate-binding protein [swine, Peptide, 116 aa]
☐	gi3318759	70.69 %	717	6	116	12.6	8.44	2095.89	Chain B, The Crystal Structures Of Two Members Of The Spermadhesin Family...

2597 Protein Group(s), 3411/9771 Protein(s), 52241/52241 Peptide(s), 72651/72651 Search Input(s)

Throughput of Protein identification

- Current:
  - ~2500 proteins/day for serum sample
  - ~5000 proteins /day for cell lysase
- Target:
  - ~5000 proteins/day for serum
  - ~10,000 proteins/day for cell lysase

Conclusions

- Proteins/peptides were separated and identified by 2D LC/MS/MS.
- Bi-Phasic method is easy, fast, reproducible.
- No salt was used during separation.
- Only one HPLC is needed for 2D separation.
- No additional valve, pump or software needed for the bi-phasic column approach.

Thank you