Biofluids, Saliva and Clinical Chemistry

Ken Parker

2-12-2016

What's in Saliva

- Lots of Peptides
- Lots of Proteins
- Lots of Small Molecules
- Lots of Cellular Debris
- Salts
- Particles of Food
- Blood
- Bacteria, Viruses

For Clinical Analytes

- Extract the molecules you want by
  - Differential centrifugation
  - Solid Phase Extraction
  - Selective Precipitation
  - Immuno-affinity enrichment

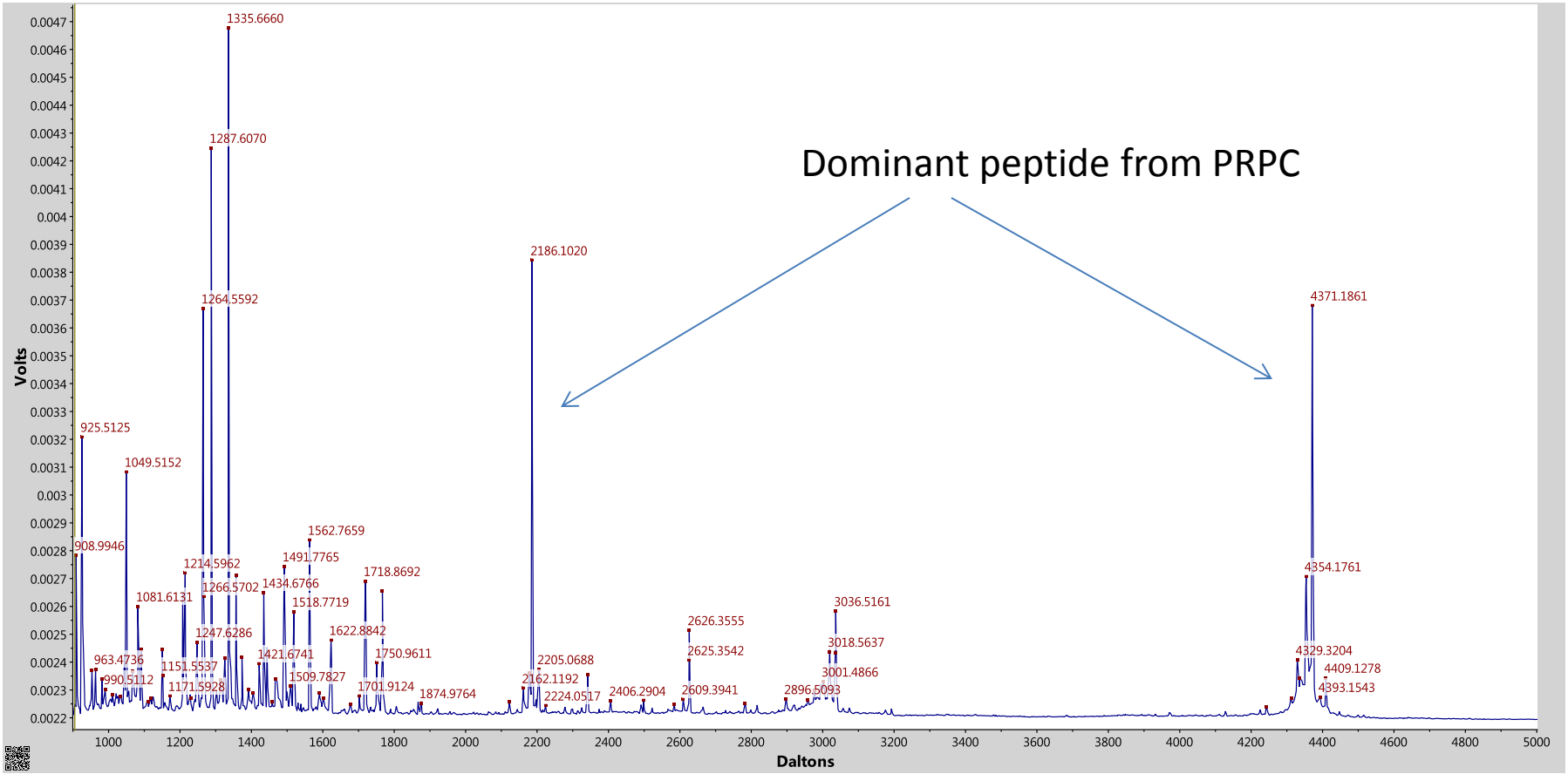
Why am I talking about this?

- Saliva is easy to measure (and relatively safe)
- We hope to have a functional instrument during the course
- We can measure saliva patterns
- Remarkable combination of stability and heterogeneity in the patterns
- These patterns may mean very little

Direct examination of Saliva by MALDI

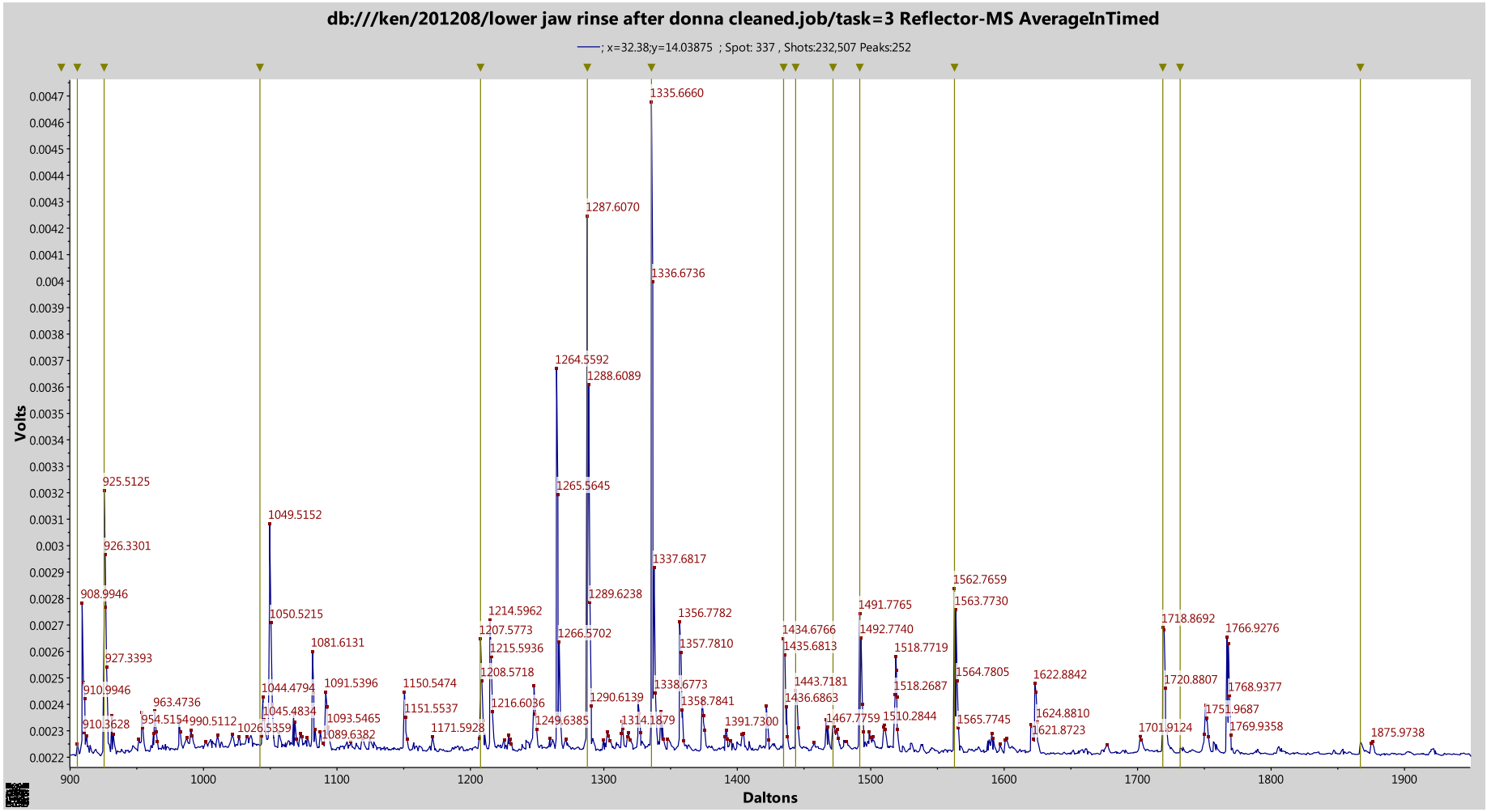
- Just dilute into matrix, spot and shoot
  - Might trouble to let particles settle
- What are the dominant players (detectable by MALDI)?
  - **Histatin-3** (His3)and **salivary acidic proline-rich phosphoprotein ½** (PRPC)
- Both proteins are typically digested into peptides by proteases
  - Some proteases may be bacterial in origin
- Many peptides from each are usually apparent
- Proportions of them may reflect differential secretion by submandibular gland Vs. parotid gland
  - Sensitive to oral cleanliness
  - Eating
  - Circadian rhythms
  - Infection and disease

Histatin 3- enriched pattern after teeth cleaned

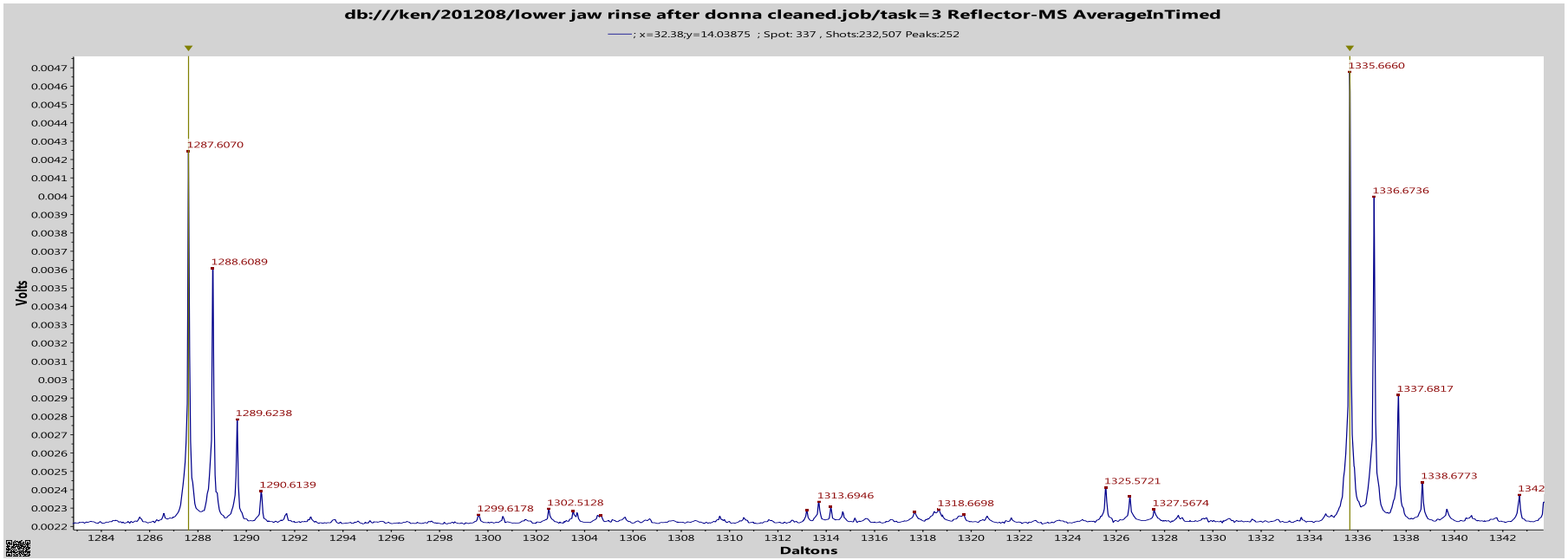


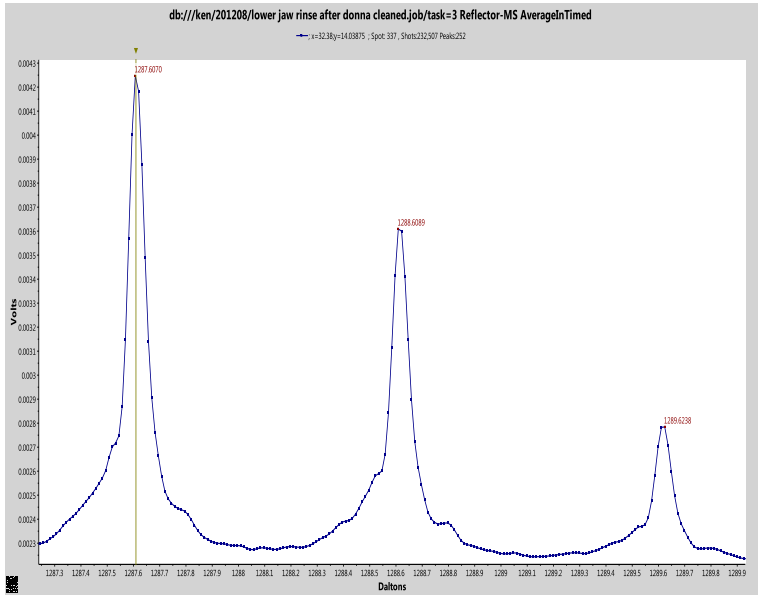
GRPQGPPQQGGHQQGPPPPPPGKPQGPPPQGGRPQGPPQGQSPQ

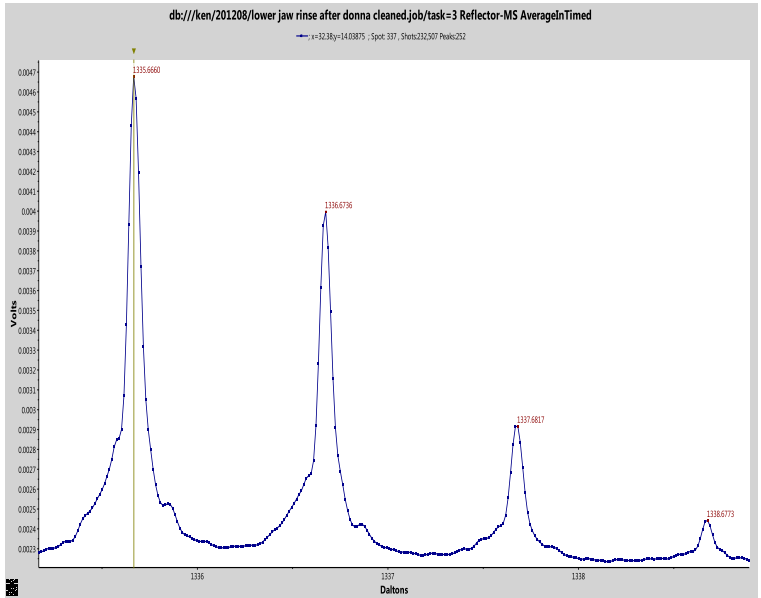
Vertical lines show matches to Histatin-3 peptides



Inset around two strong peaks







Histatin-3 natural saliva peptides
Colors highlight shared regions

I	MassThe	MassExpe	ppm	<	Sequence	>
11	925.512	925.513	0.5	A	KRHHGYK	R
11	925.512	925.513	0.5	R	HHGYKRK	_
51	953.518	953.524	5.9	K	RHHGYKR	K
68	962.496	962.488	-8.2	R	KFHEKHH	S
50	1044.475	1044.479	4.3	Y	RSNYLYDN	
48	1067.525	1067.532	7.0	F	HEKHHSHR	E
24	1081.613	1081.613	-0.1	K	RHHGYKRK	_
24	1081.613	1081.613	-0.1	A	KRHHGYKR	K
34	1150.551	1150.547	-2.8	H	EKHHSHRGY	R
21	1207.572	1207.577	4.3	A	DSHAKRHHGY	K
4	1264.560	1264.559	-0.5	R	GYRSNYLYDN	
3	1287.610	1287.607	-1.9	F	HEKHHSHRGY	R
2	1335.667	1335.666	-0.8	A	DSHAKRHHGYK	R
16	1434.678	1434.677	-0.9	K	FHEKHHSHRGY	R
28	1443.711	1443.718	5.2	F	HEKHHSHRGYR	S
12	1491.768	1491.777	5.6	A	DSHAKRHHGYKR	K
7	1562.773	1562.766	-4.5	R	KFHEKHHSHRGY	R
63	1590.779	1590.783	2.4	K	FHEKHHSHRGYR	S
39	1619.863	1619.863	0.1	A	DSHAKRHHGYKRK	F
10	1718.874	1718.869	-2.8	R	KFHEKHHSHRGYR	_
10	1718.874	1718.869	-2.8	K	RKFHEKHHSHRGY	R
79	1874.975	1874.976	0.7	K	RKFHEKHHSHRGYR	S
20	3035.523	3035.497	-8.4	_	DSHAKRHHGYKRKFHEKHHSHRGY	_

Other Biofluids by MALDI

- Blood has a very simple proteome if you don't look too hard
  - Most abundant: Albumin, IgG (very complex in detail), Transferrin
  - Autodigested peptides in blood rarely derive from these proteins
  - Instead they derive from transthyretin, C3, haptoglobulin, fibrinogen, etc in different ways depending how the blood is handled.
  - Peptides tend to derive from selected regions, and be frayed at the ends, similar to the His3 peptides
- Natural peptides in tears also derive substantially from a small number of protein precursors

Major Saliva proteins determined by peptide mass fingerprinting following trypsin Reflector instrument 7.5 m flight tube.

N	symb	acc	leng	protein	pep	M	Mi	TriM	%CM	%IM	ppw
1	AMY1A	P04745	511	Alpha-amylase 1	107	24	24	1952453	59.4	33.3	1.6
2	CST1	P01037	141	Cystatin-SN	23	9	9	249026	65.9	7.1	2.3
3	IGHA1	P01876	353	Ig alpha-1 chain C region	43	7	8	177642	63.1	3.6	1.9
4	CST4	P01036	141	Cystatin-S	32	9	9	38741	51.7	2.6	4.1
5	LYZ	P61626	130	Lysozyme C	31	4	4	13924	43.8	0.6	3.0

Fig. 1A. Tryptic digest of whole serum (from Sigma), dominated by albumin.

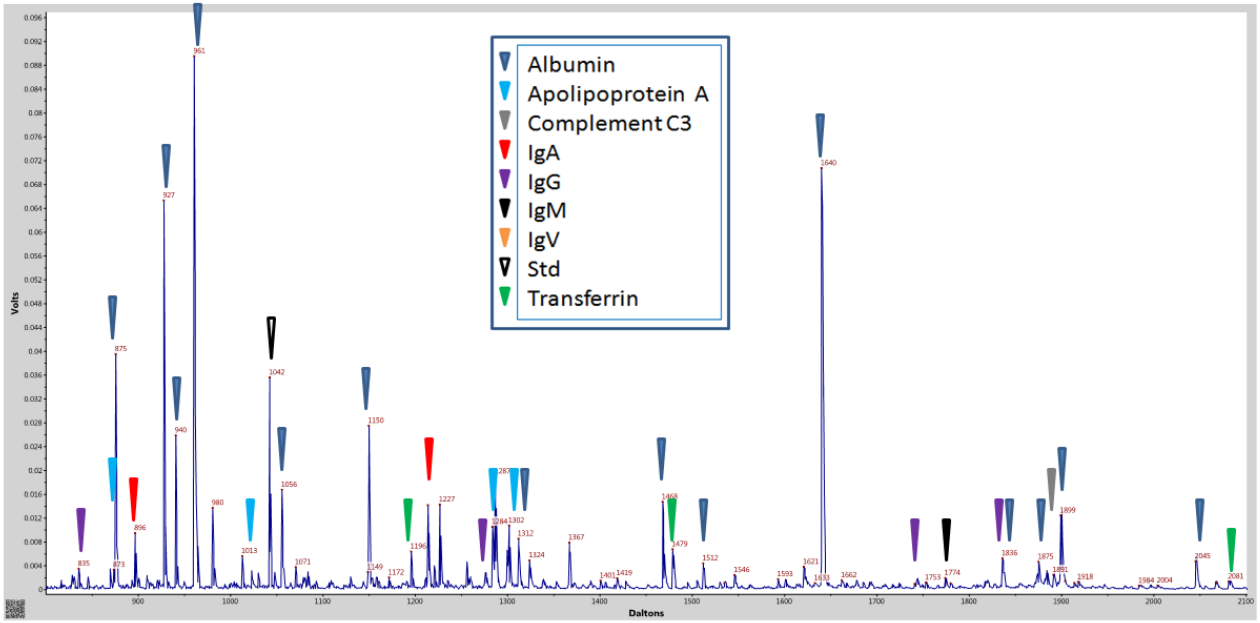
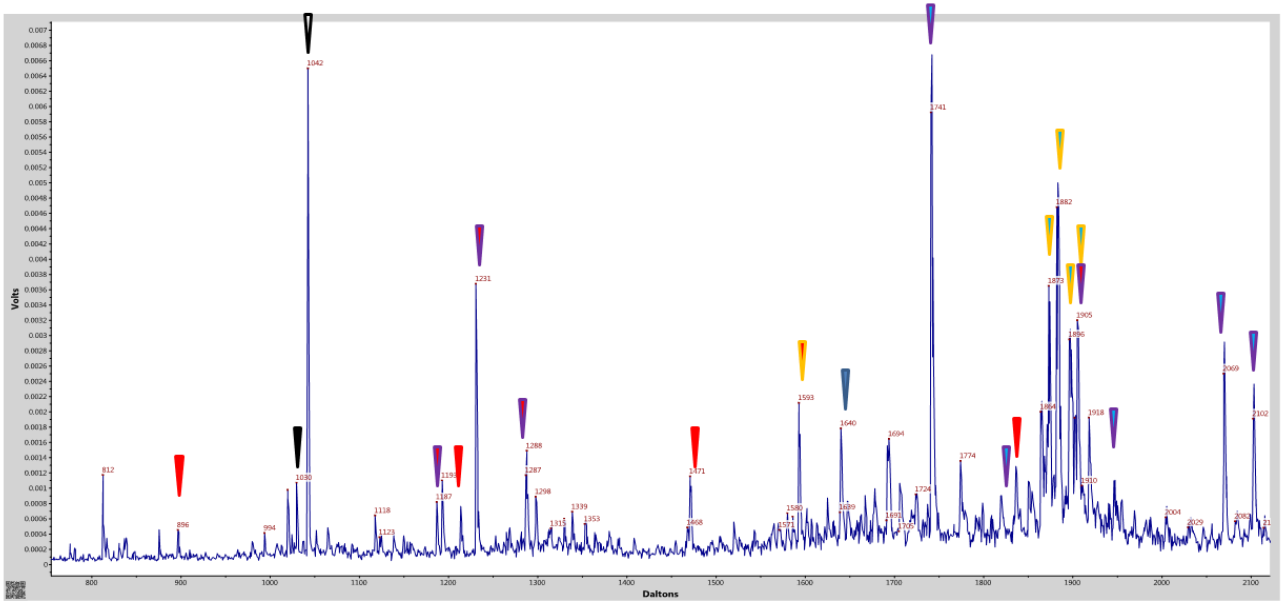


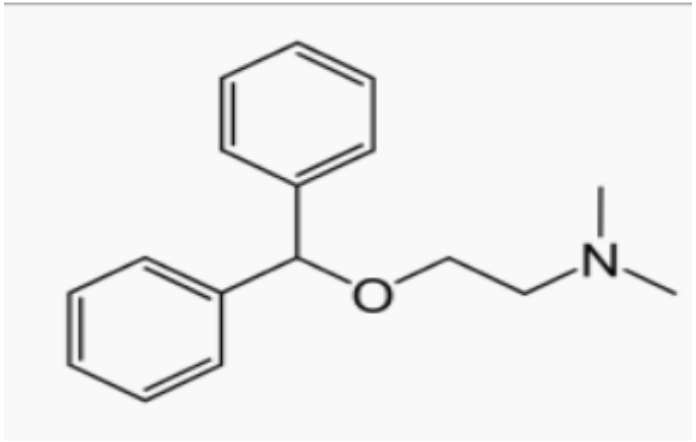
Fig 1**B. Trypic digest of proteins purified by Protein A from same digest.**



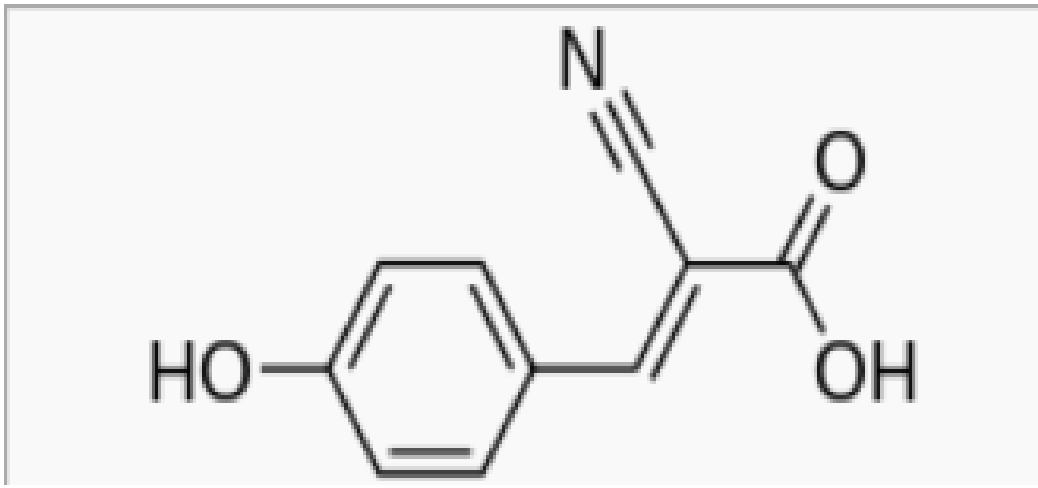
PRR4 peptides from tears (proposed)

Rank	aa	Mass	ppm	Sequence	Intensity	MSMS	before
63	110	1190.6	-5.7	RHPQEQPLW	30		
27	109	1261.6	-6.8	ARHPQEQPLW	84		
42	108	1358.7	-5.7	PARHPQEQPLW	47		
1	106	1629.8	0.0	DRPARHPQEQPLW	479	*	*
30	105	1785.9	-0.5	RDRPARHPQEQPLW	74		
93	102	2236.2	-4.3	FFRRDRPARHPQEQPLW	22		
16	101	2323.2	0.3	SFFRRDRPARHPQEQPLW	125		
3	100	2410.2	4.2	SSFFRRDRPARHPQEQPL*W*	479	*	
19	95	2938.5	0.0	SLQEASSFFRRDRPARHPQEQPLW	102	*	
145	105	1004.5	-13.9	RDRPARHP	13		
92	103	1307.7	-12.5	FRRDRPARHP	22		
5	102	1454.8	-0.5	FFRRDRPARHP	309		
13	101	1541.8	11.3	SFFRRDRPARHP	132		
14	100	1628.8	-4.7	SSFFRRDRPARHP	130		
98	96	2070.1	6.5	LQEASSFFRRDRPARHP	21		
108	95	2157.1	-12.8	SLQEASSFFRRDRPARHP	19		
4	100	1238.6	-9.2	SSFFRRDRPA	479	*	
48	98	1438.7	6.7	EASSFFRRDRPA	40		
124	96	1679.9	7.8	LQEASSFFRRDRPA	16		
2	95	1766.9	-3.0	SLQEASSFFRRDRPA	479	*	
47	88	1343.7	4.5	LPRFPSVSLQEA	43		
32	87	1430.8	0.4	SLPRFPSVSLQEA	68		
100	85	1671.9	-2.4	QLSLPRFPSVSLQEA	20		
61	85	1213.7	3.9	qLSLPRFPSVS	31	*	
29	85	1326.7	-7.8	qLSLPRFPSVSL	75	*	
5	85	1454.8	-9.3	qLSLPRFPSVSLQ	0		
65	85	1654.9	4.8	qLSLPRFPSVSLQEA	29		
44	85	1829.0	13.6	qLSLPRFPSVSLQEASS	43		

Diphenhydramine



α-Cyano-4-hydroxycinnamic acid



$(C_{17}H_{21}NO)H+$ ->256.1695

$(C_{20}H_{14}N_6O_2)H+$ ->379.0924

measured	thoeretical	ppm	atomic composition
146.0584	146.060039	-11.2	(C9 H7 N O)H+
172.0389	172.0393	-2.4	(C10 H5 N O2)H+
190.0507	190.049867	4.6	(C10 H7 N O3)H+
212.0333	212.0318	7.2	(C10 H7 N O3)Na+
228.0106	228.0058	21.1	(C10 H7 N O3)K+
256.1683	256.1695	-4.8	(C17 H21 NO)H+
278.1491	278.1515	-8.7	(C17 H21 NO)Na+
294.0765	294.1255	-166.7	(C17 H21 NO)K+
335.1017	335.1026	-2.8	(C19 H14 N2 O4)H+
379.0928	379.092458	0.9	(C20 H14 N2 O6)H+

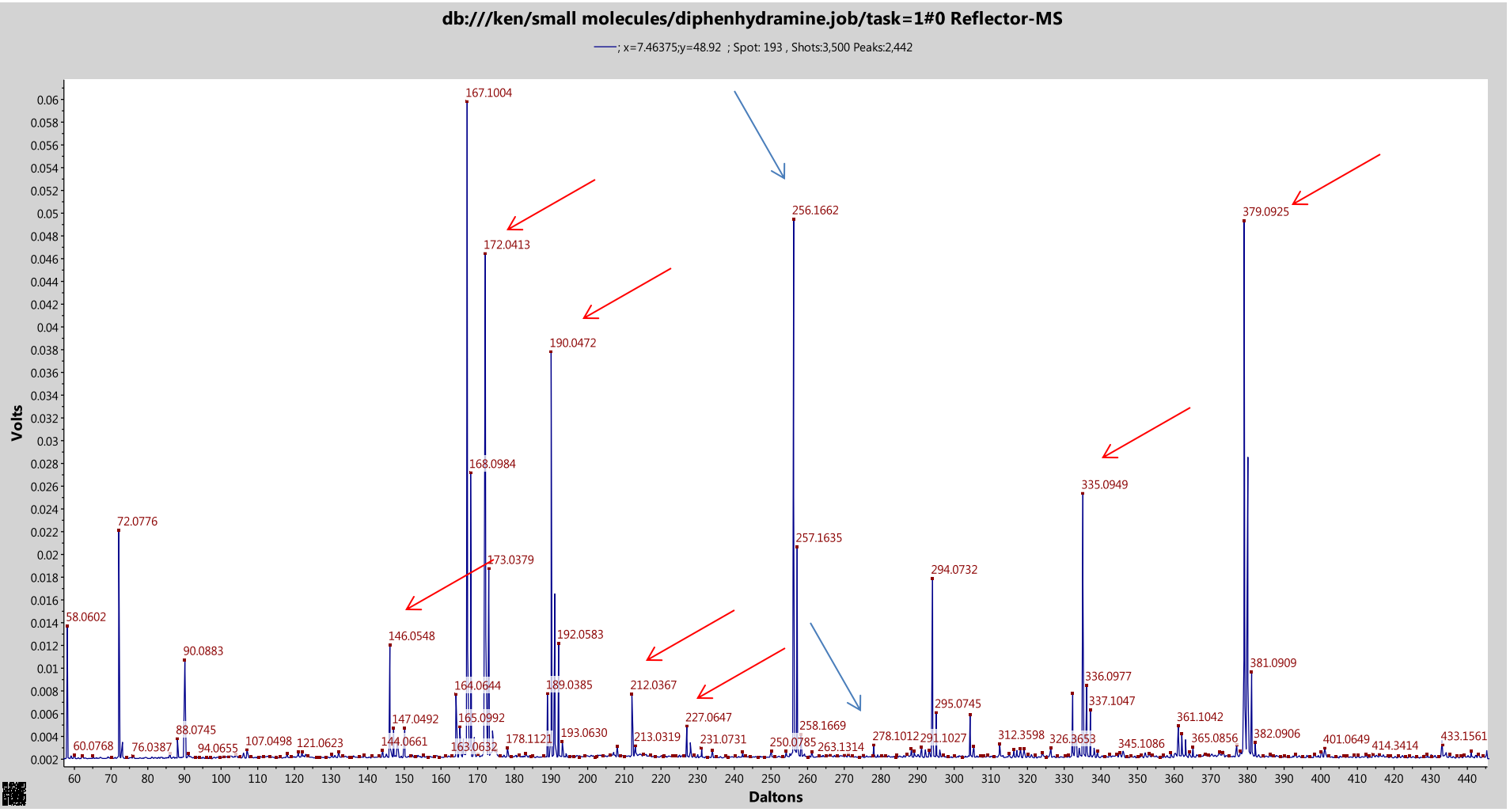
Resolution ~ 8400

Note:
H+= 1.007276
H= 1.007825

This mass difference
=2.1 ppm at mz 256.1685

Reflector spectrum of diphenyhydramine in HCCA

Red arrows: known matrix –derived peaks
Blue arrows: diphenhydramine related



Can see marijuana active ingredients over CHCA background if desired

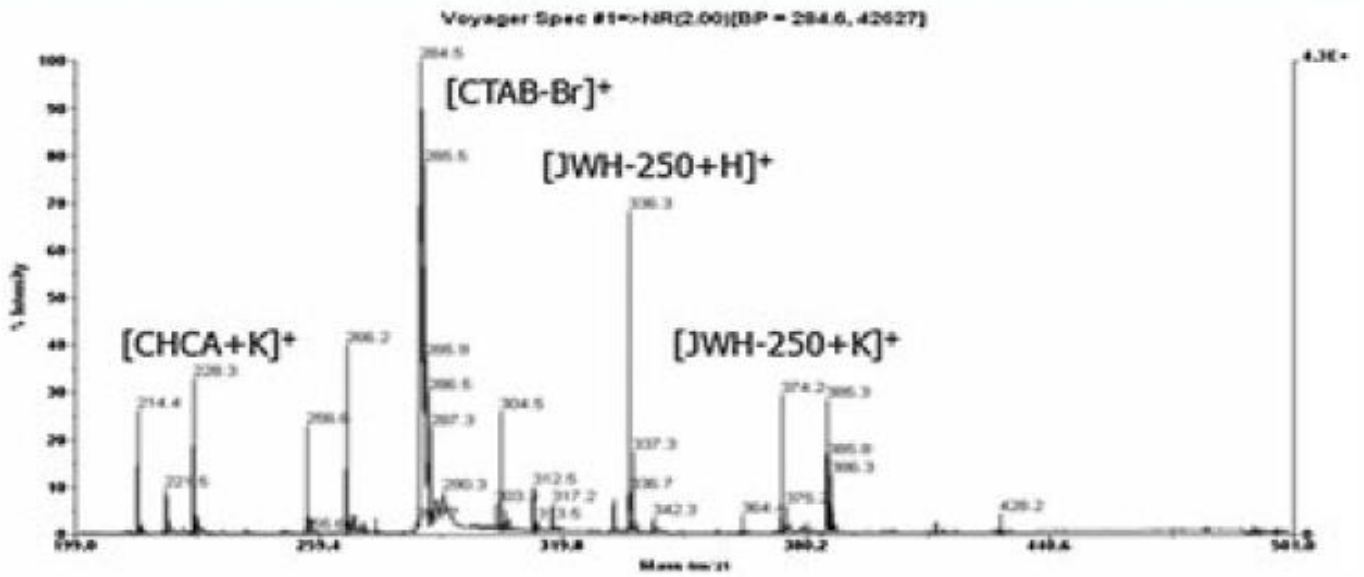
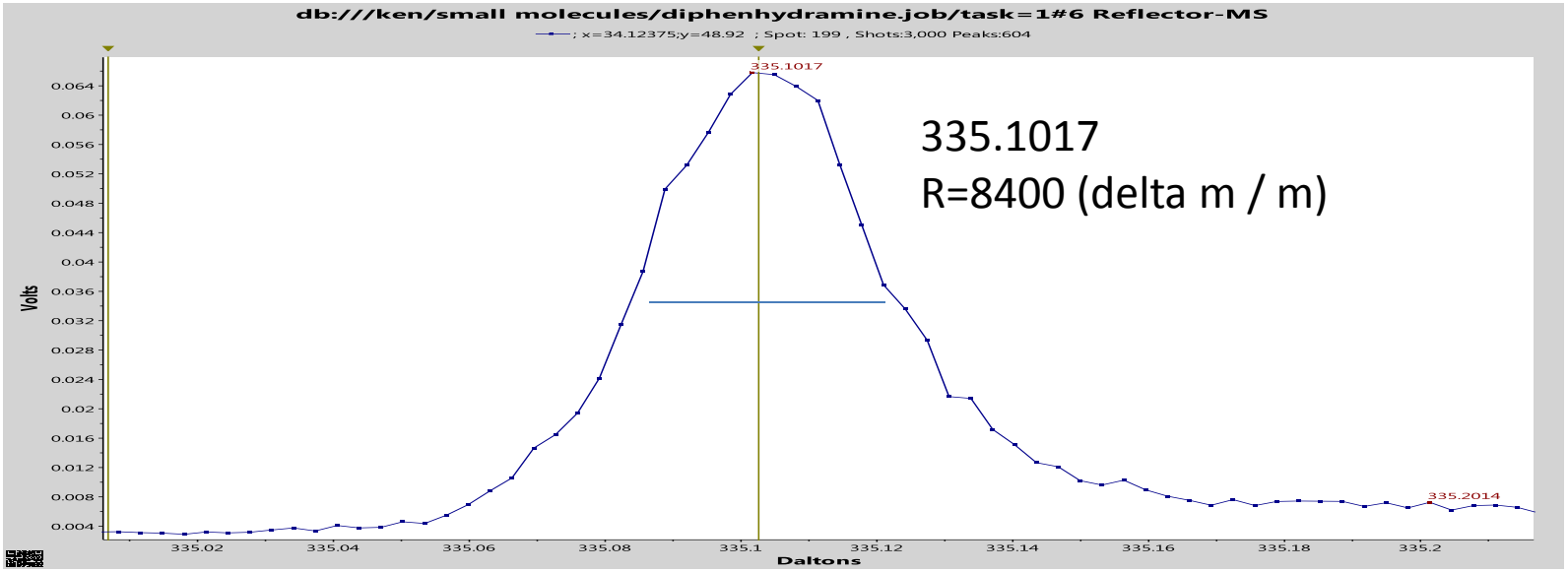
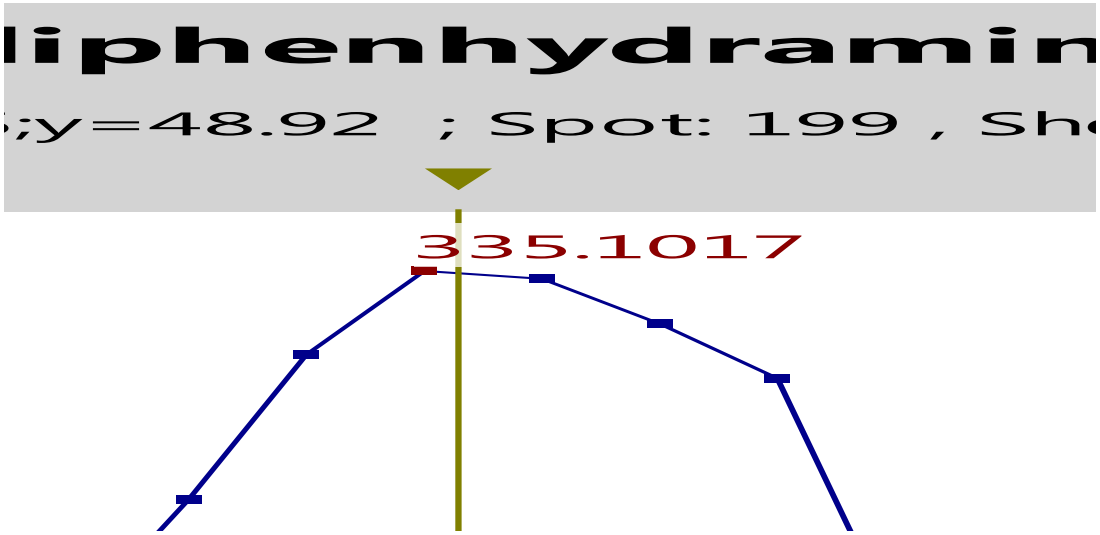


Figure 4. MALDI-TOF mass spectra of the commercial product named Amazonas. In the spectrum, it is possible to identify the ions corresponding to the synthetic cannabinoid JWH-250 well distinguishable from those related to the matrix. The *y* axis represents the relative intensity (%), and the *x* axis reports the *m/z* values.

From: Direct screening of herbal blends for new synthetic cannabinoids by MALDI-TOF MS.
Rossella Gottardo, Anna Chiarini, Ilaria Dal Prà, Catia Seri,c Claudia Rimondo, Giovanni Serpelloni, Ubaldo Armatob and Franco Tagliaroa
J. Mass. Spectrom. 2012, 47, 141–146





1 khz,
bin = 1 nsec

Label here is on the most intense bin, not on the peak centroid.
The horizontal line is theoretical for (C19 H14 N2 O4)H+

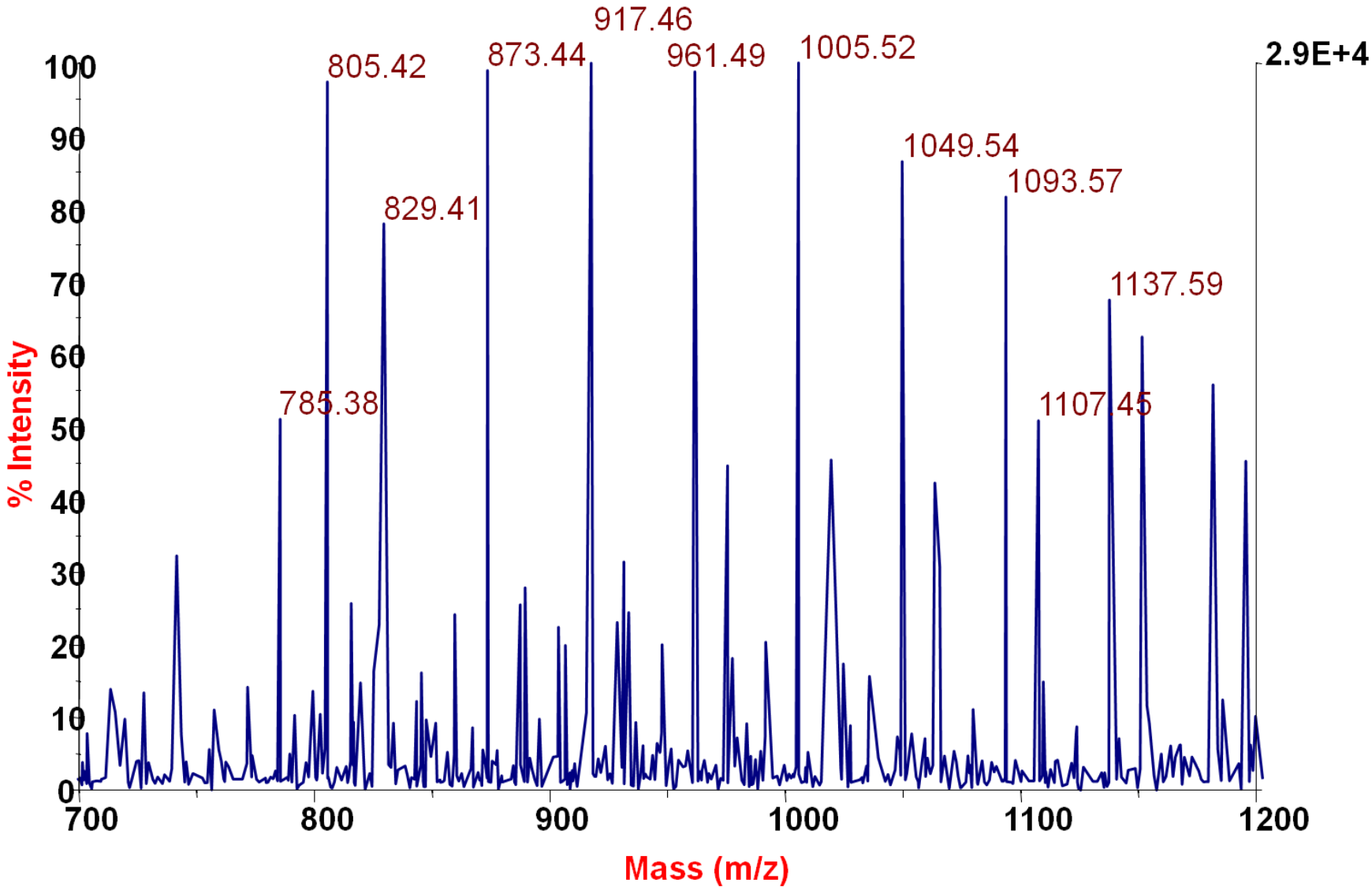
Artifacts to look out for

- Are the peaks labeled correctly?
- Are there unexpected peaks in the sample?
  - Could be in the water
  - Could derive from
    - Microcentrifuge tube
    - Pipet tips
  - Do solvents evaporate at an expected rate?
- Are there adducts?
  - +16 could be oxygen
  - +16 could also mean one is Na+ (+22), the other is K+ (+38).
  - We often find Mn+ and Cu+ adducts, especially to HCCA dimer.
- Polyethylene glycol (+44) series are common [C2H5O]
- Silicone (+74) series also common [Si(CH3)2 O]
- Beware of magic numbers
  - Some multimers of matrix and alkali metal atoms are more favored than others, depending on subtle conditions

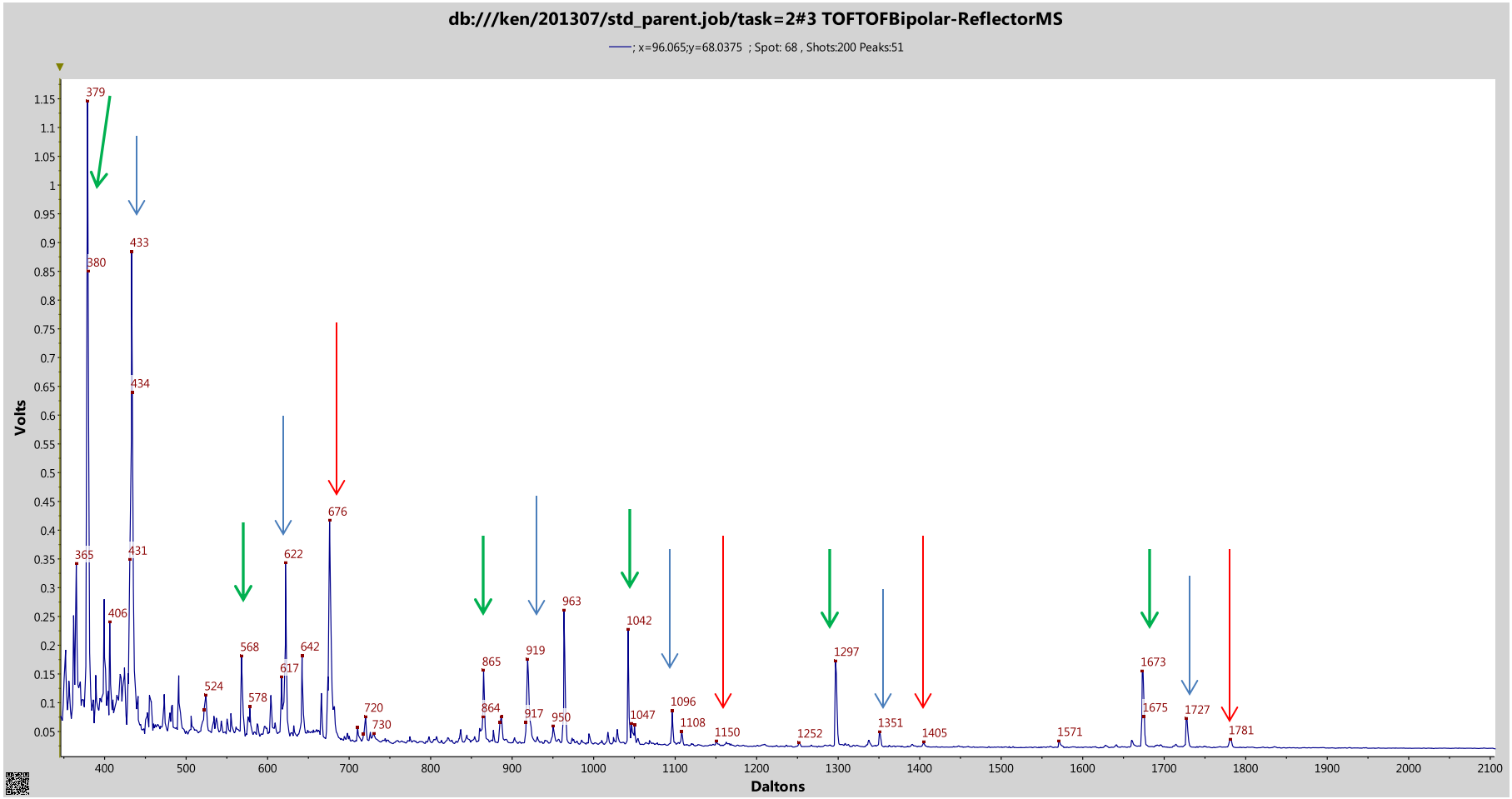
Suppression of alpha-cyano-4-hydroxycinnamic acid matrix clusters and reduction of chemical noise in **MALDI**-TOF **mass spectrometry**.

Smirnov IP, Zhu X, Taylor T, Huang Y, Ross P, Papayanopoulos IA, Martin SA, Pappin DJ. Anal Chem. 2004 76:2958-65.

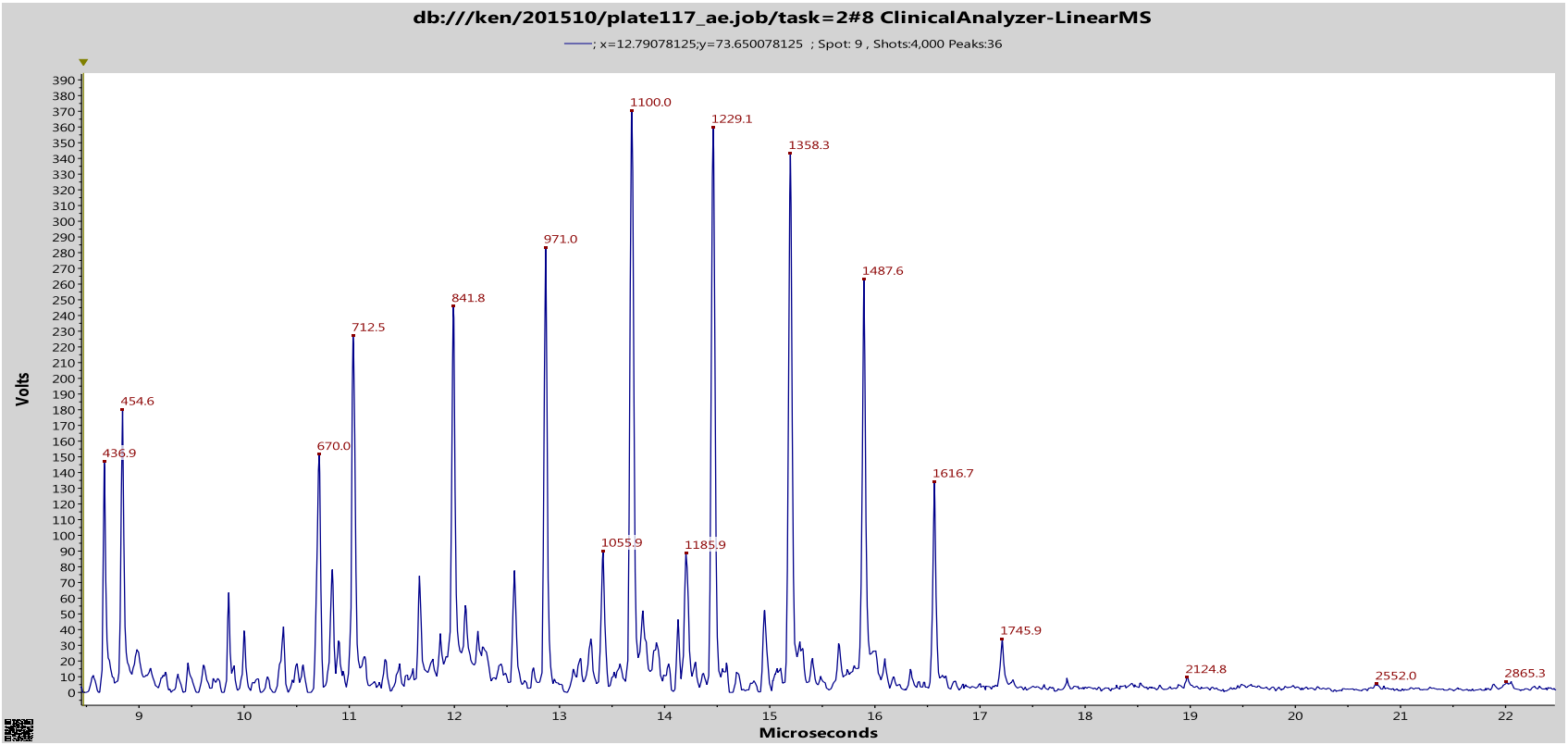
Protein Digest Badly Contaminated with Polymer



Std peptides with Mn contamination
Base Peak green
Single Mn in blue
Two Mn in red



Natural biobolymer of glutamic acid on perhaps a folic acid derivative



[Observation of tetrahydrofolylpolyglutamic acid in bacteria cells by matrix-assisted laser desorption/ionization mass spectrometry.]

Arnold RJ, **Reilly JP**.

Anal Biochem. 2000;281:45-54.

SISCAPA enrichment of EDQYHYLLDR from Protein C inhibitor (PCI)

Clusterin peptide contaminants sharing YHYL

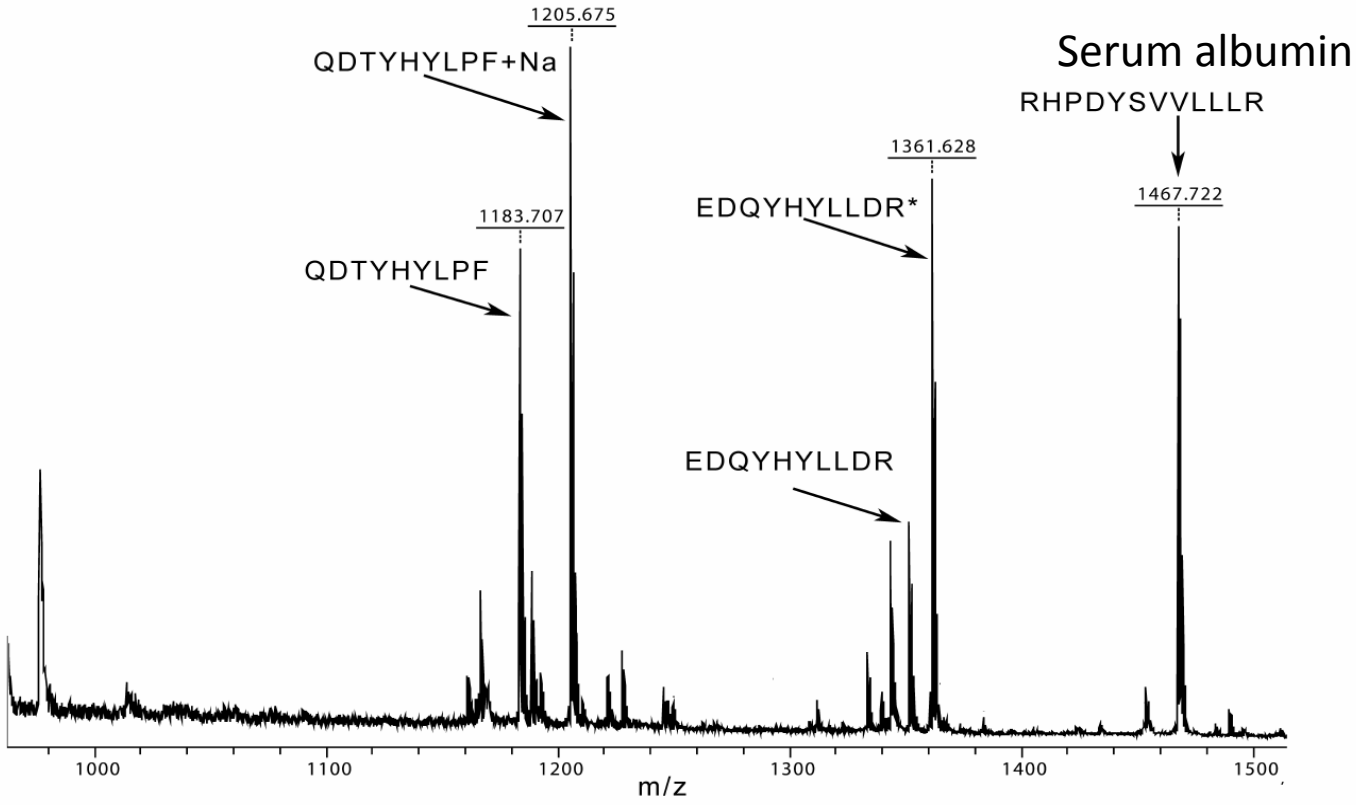


Figure 7. MALDI-TOF spectrum of peptides enriched from a spiked human plasma digest by antipeptide antibodies (EDQYHYLLDR specificity) on magnetic beads (SISCAPA workflow).

N. Leigh Anderson, Morteza Razavi, Terry W. Pearson, Gary Kruppa, Rainer Paape, and Detlef Suckau. Precision of Heavy–Light Peptide Ratios Measured by MALDI-TOF Mass Spectrometry J Proteome Res. 2012;11:1868-78.

Biomarker Discovery

- Matrix methods
- Line up all of the mass measurments using bins
- Get a matrix of mz vs intensity vs sample
- Can now perform hierarchical clustering and PCA
- Easy to do
- Also, easy to fool oneself that the correlations are meaningful
- One could argue that measuring glycosylation of hemoglobin in a special case of a matrix method

Serum Cancer Biomarker Highly Selected References

1.)Use of proteomic patterns in serum to identify ovarian cancer.
Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, Mills GB, Simone C, Fishman DA, Kohn EC, **Liotta** LA.
Lancet. 2002 ;359:572-7.

Set off a flurry of research using MALDI (or SELDI) to quantify small peptides as tumor markets. Much of this research was later discredited, due to flaws in methodology, in particular to problems in peak picking, and regarding instrument stability.

2.) Differential exoprotease activities confer tumor-specific serum peptidome patterns.
Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, Fleisher M, Lilja H, Brogi E, Boyd J, Sanchez-Carbayo M, Holland EC, Cordon-Cardo C, Scher HI, Tempst P.
J Clin Invest. 2006 Jan;116(1):271-84. PMID: 16395409

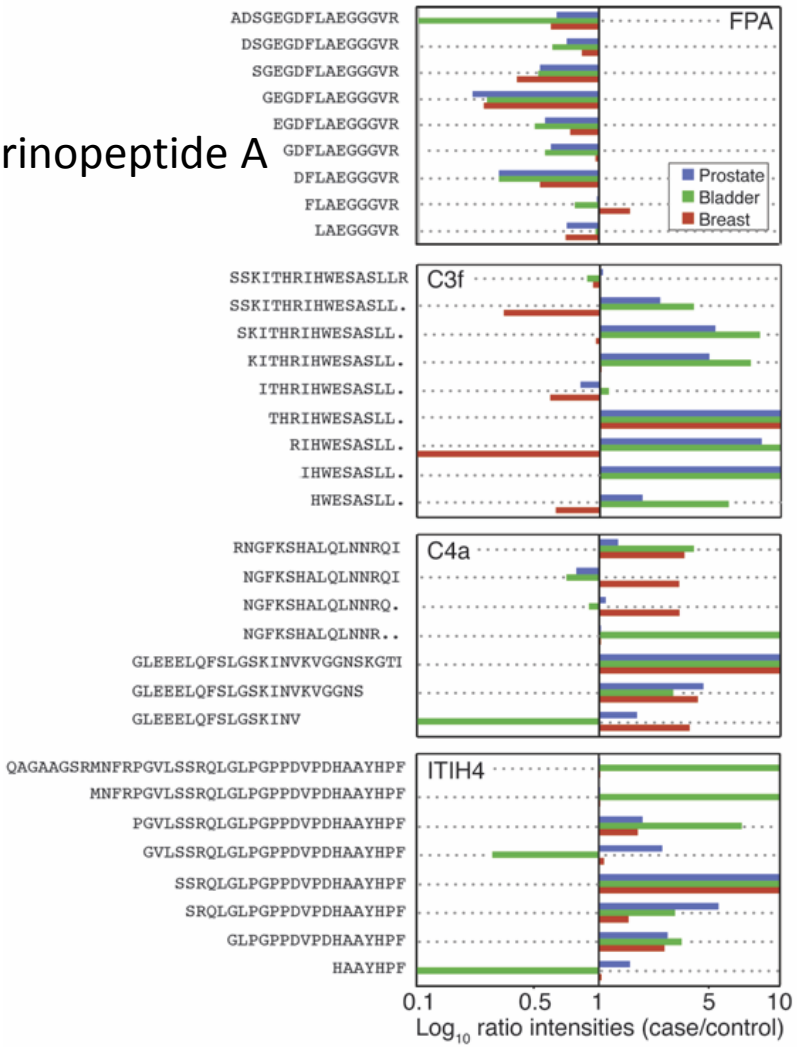
Seems to claim that particular patterns of overlapping peptides from common serum proteins are useful tumor biomarkers
(see next slide)

3.) Correcting common errors in identifying cancer-specific serum peptide signatures.
Villanueva J, Philip J, Chaparro CA, Li Y, Toledo-Crow R, DeNoyer L, Fleisher M, Robbins RJ, **Tempst P**. J Proteome Res. 2005 4:1060-72.PMID: 16083255

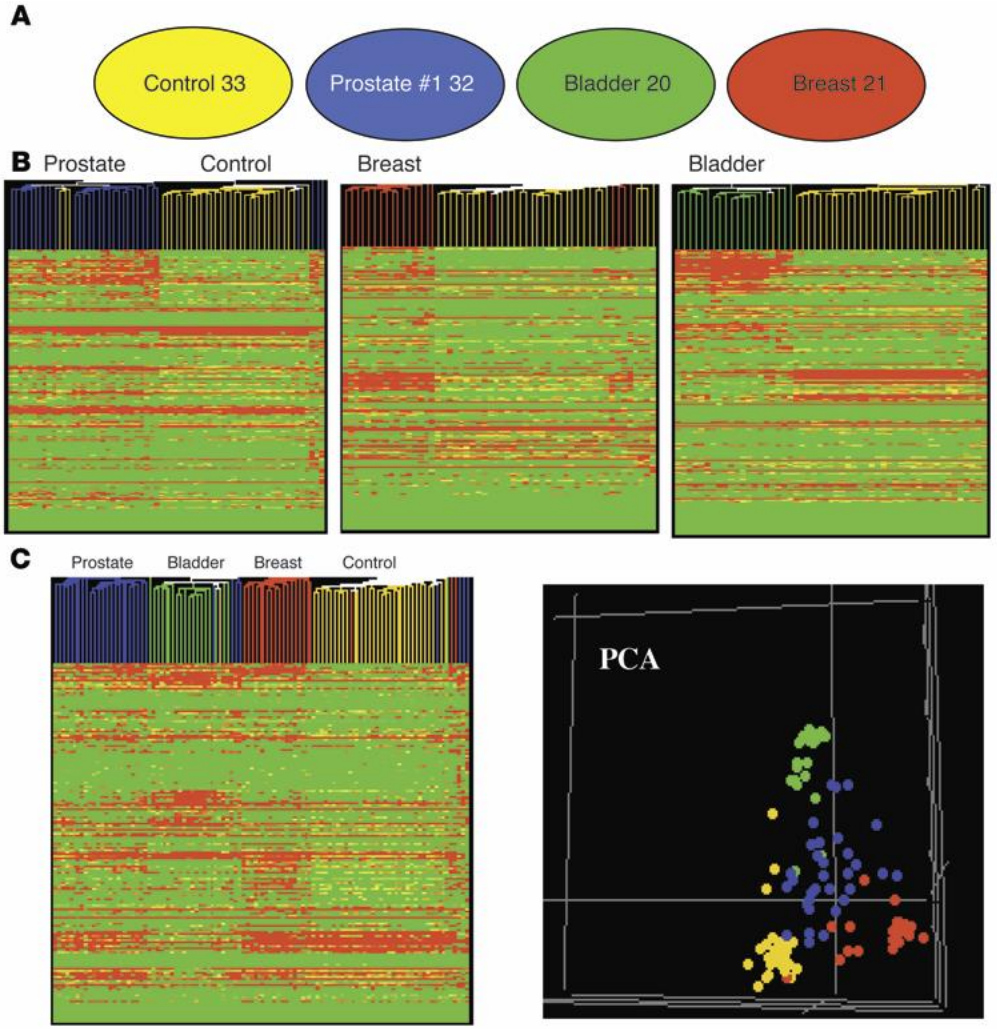
Describes many of the problems with sample reproducibility as consequence of uneven sample handling

Villanueva et al 2006

Unsupervised analysis of 651 peptide ion signals from MS-based serum profiling differentiates 3 types of cancer and controls.



Fibrinopeptide A



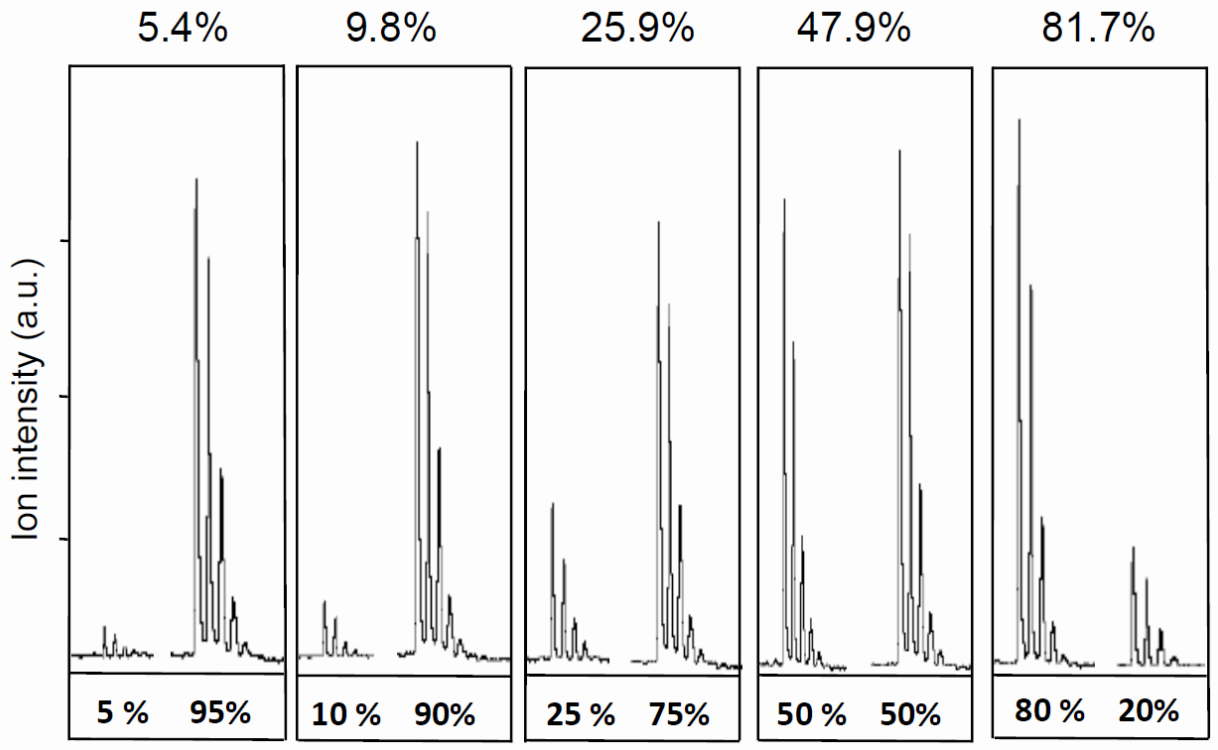
Nazarian A, Lawlor K, Yi SS, Philip J, Ghosh M, Yaneva M, **Villanueva J, Saghatelian A, Assel M, Vickers AJ, Eastham JA, Scher HI, Carver BS, Lilja H, Tempst P.**

Inhibition of circulating dipeptidyl peptidase 4 activity in patients with metastatic prostate cancer.

Mol Cell Proteomics. 2014 13:3082-96. PMID: 25056937

Quantify relative intensity of precursor product peptides vs time

Supplementary Figure 2

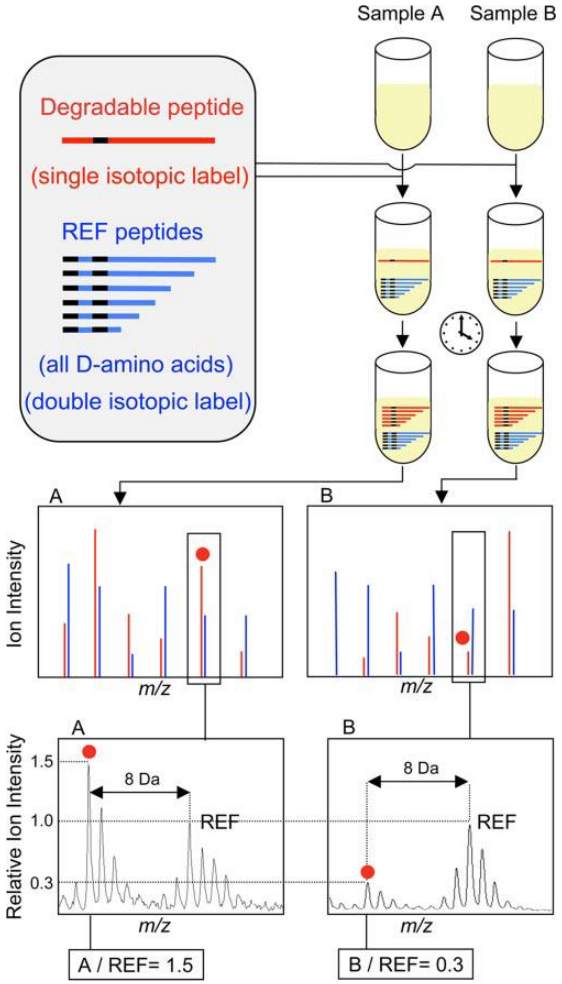


Authors determined tumor correlated with an inhibitor of a natural serum protease. The protease itself was unchanged in concentration

A Sequence-specific Exopeptidase Activity Test (SSEAT) for "Functional" Biomarker Discovery

Josep Villanueva, Arpi Nazarian, Kevin Lawlor, San San Yi, Richard J. Robbins and Paul Tempst

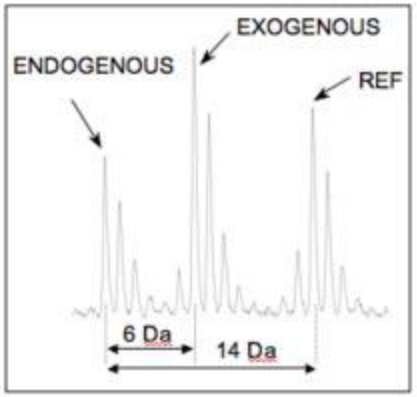
Mol Cell Proteomics. 2008 7:509-18. PMID: 17986438



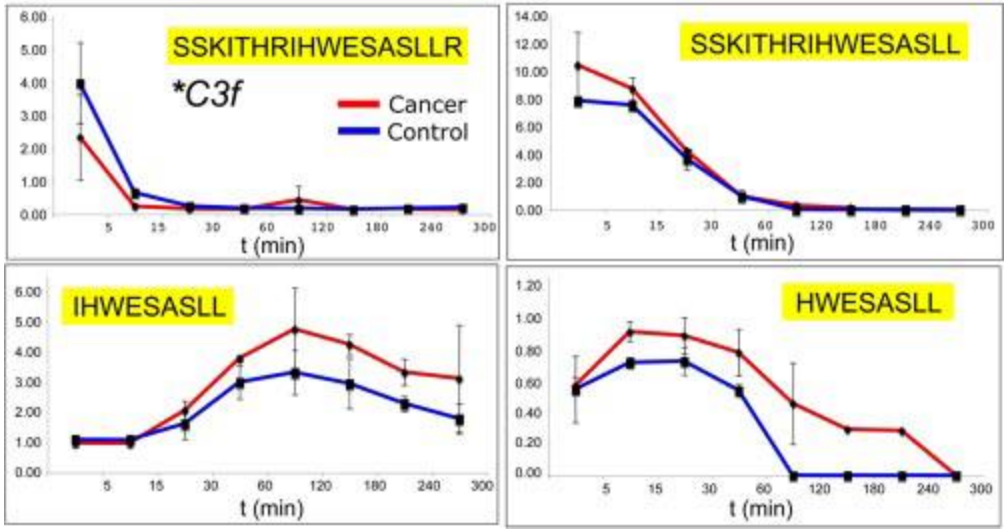
a

2021.10	SSKITHRIHWESASLLR
1864.99	SSKITHRIHWESASLL.
1777.97	SKITHRIHWESASLL.
1690.93	KITHRIHWESASLL.
1562.84	ITHRIHWESASLL.
1449.76	THRIHWESASLL.
1348.71	HRIHWESASLL.
1211.65	RIHWESASLL.
1055.55	IHWESASLL.
942.47	HWESASLL.

b



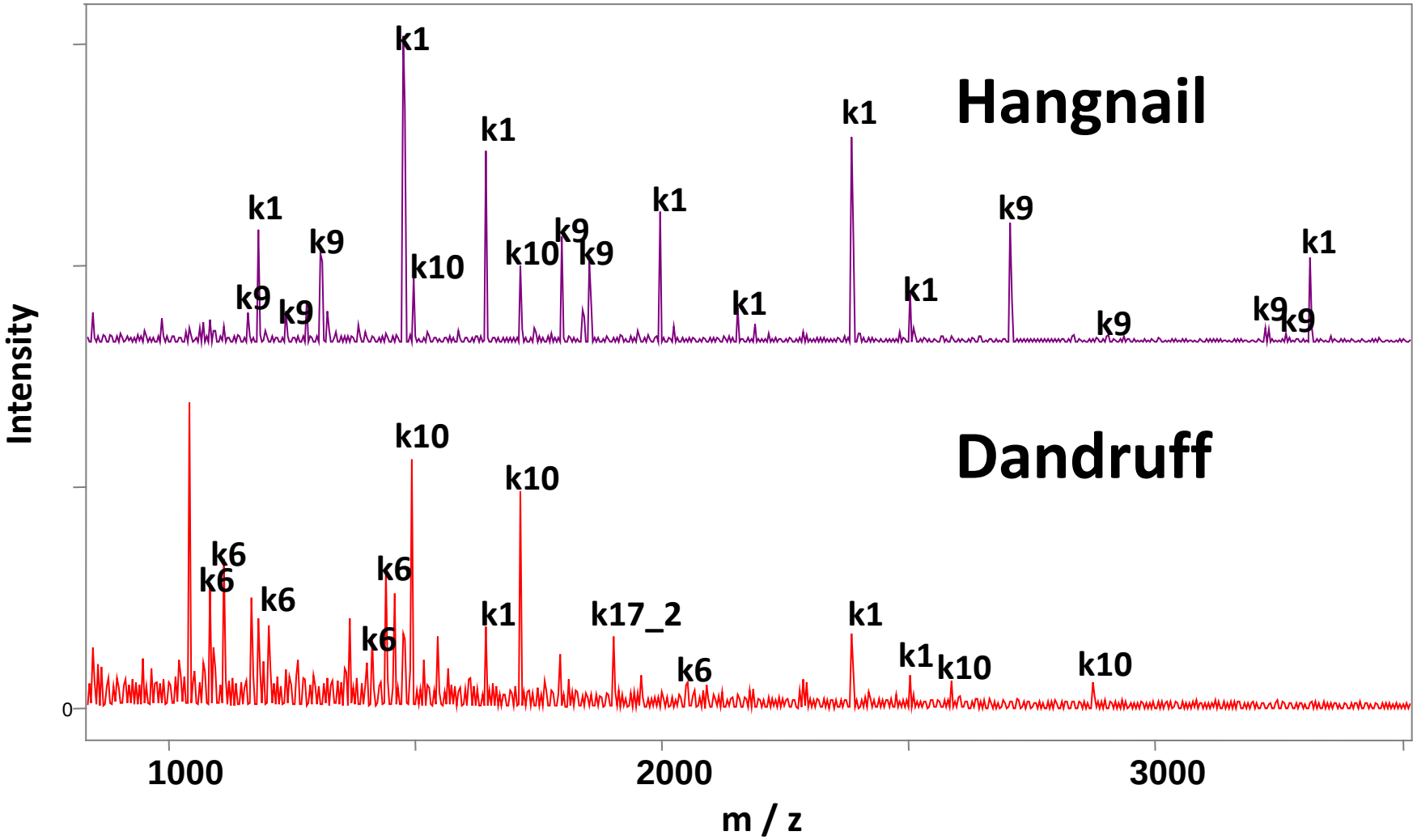
c From Complement C3



Which Masses are Possible?

- All unmodified peptides contain 5 elements:
  - **Carbon mass= 12.000000**
  - **Hydrogen mass= 1.007825**
  - **Nitrogen mass= 14.003074**
  - **Oxygen mass= 15.994915**
  - **Sulfur mass= 31.972071**
- Hydrogen and Nitrogen have a “mass excess”
- Oxygen and Sulfur have a “mass deficit”

Hangnail vs. Dandruff



What about isotopes?

- At high mass accuracy, all isotopes are well resolved up to masses of > 5000 (attainable by state-of-the-art MALDI)
- The isotope distribution also is informative
- C13, at ~ 1% natural abundance contributes the bulk of the heavy atoms
- Sulfur, which is rare, has two isotopes, 95% 32, 5% 34.
  - so sulfur-rich compounds have more heavy isotopes
- Certain atoms rarely found in peptides, but abundant in many contaminants, like Cl, Fe, and Si have distinctly different isotope distributions and mass deficits

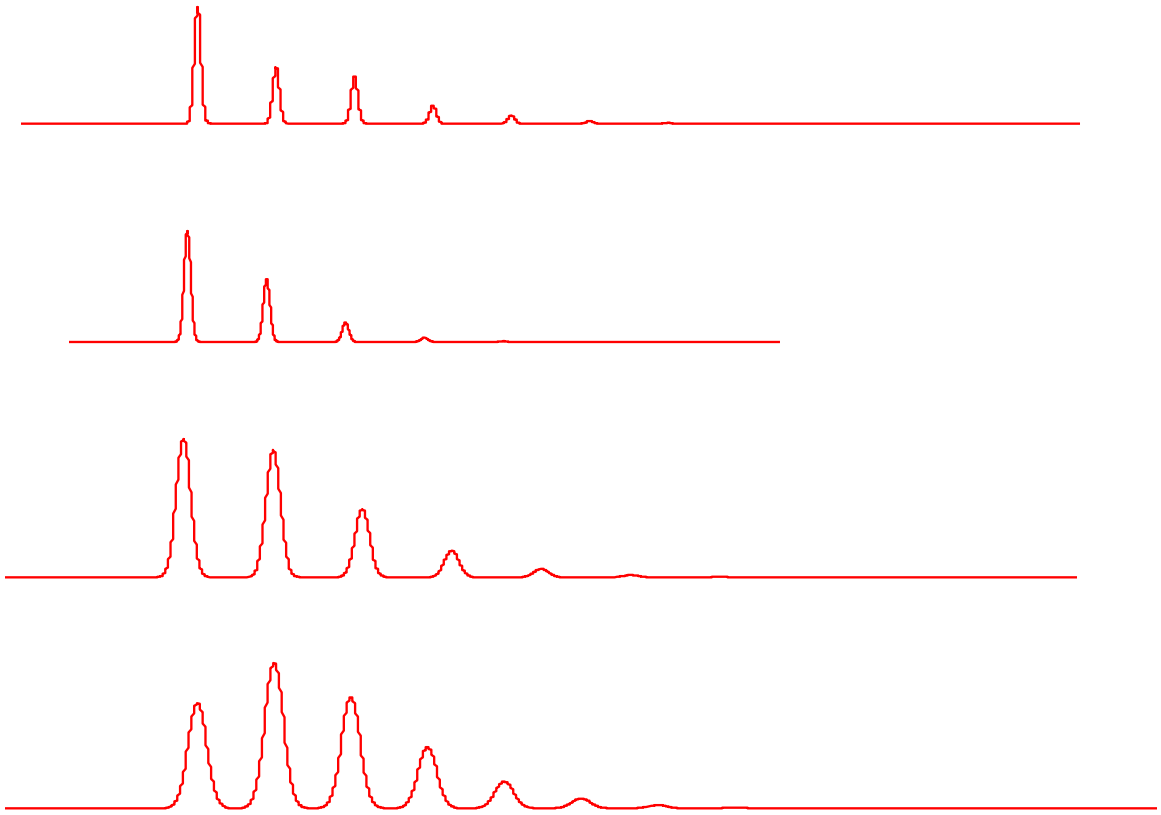
Typical isotope profiles

AC_6EY 1000.22

GK_6R 1000.71

$A_{12}S_{12}$ 1914.84

$A_{18}S_{18}$ 2863.25



Calculated by IsoPro at http://members.aol.com/msmssoft/