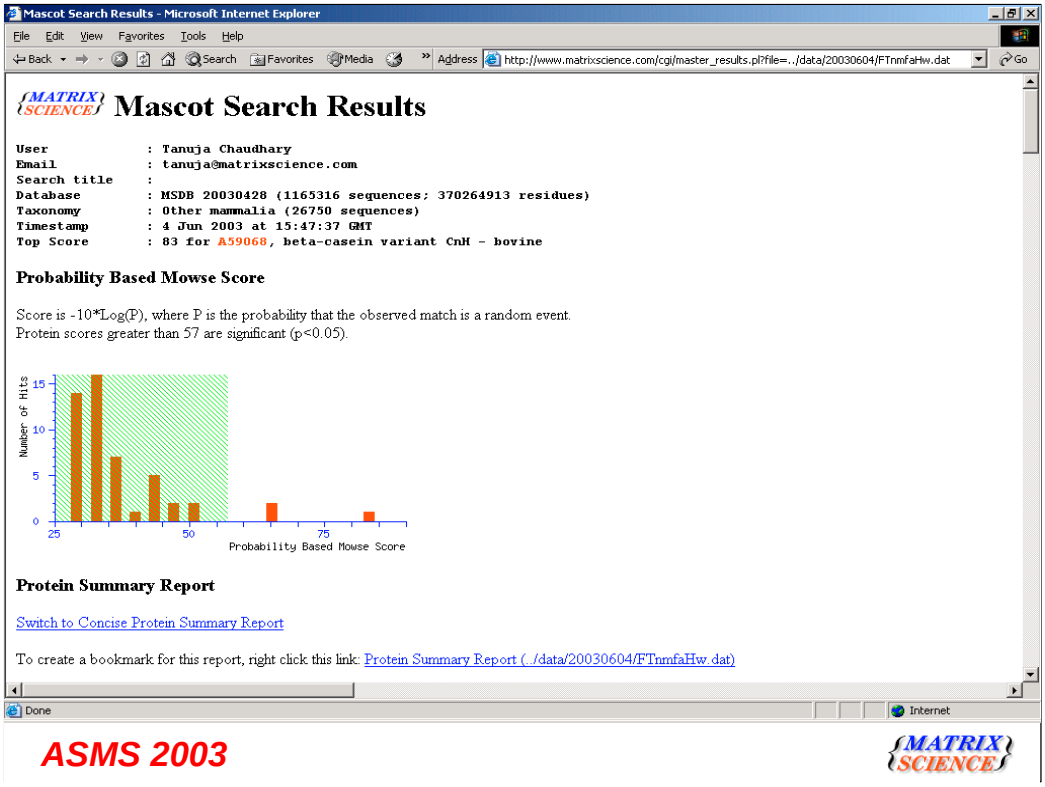
Using Mascot to characterize Protein modifications

ASMS 2003

MATRIX SCIENCE

Post-translational Modifications (PTMs)

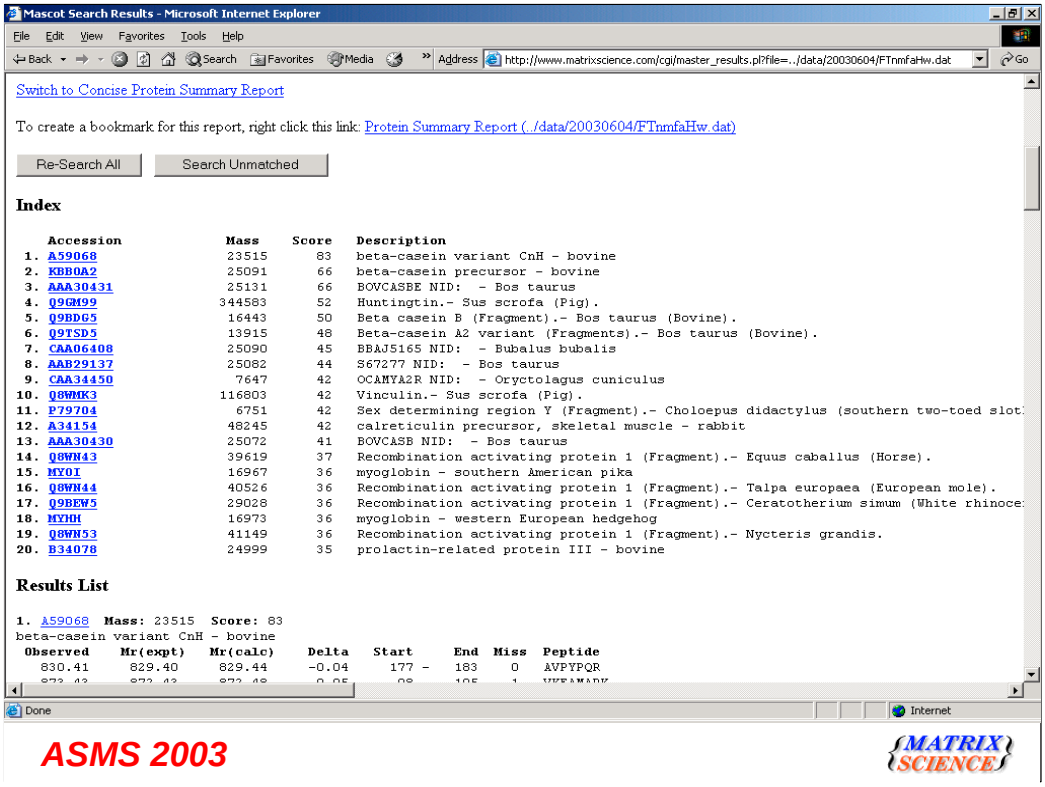
- Help understand complex biological systems
- Phosphorylation is one of the most important protein PTMs
- Mascot allows up to 9 variable modifications to be specified at a time
- Use variable modifications sparingly, follow it by error tolerant search



ASMS 2003

MATRIX SCIENCE

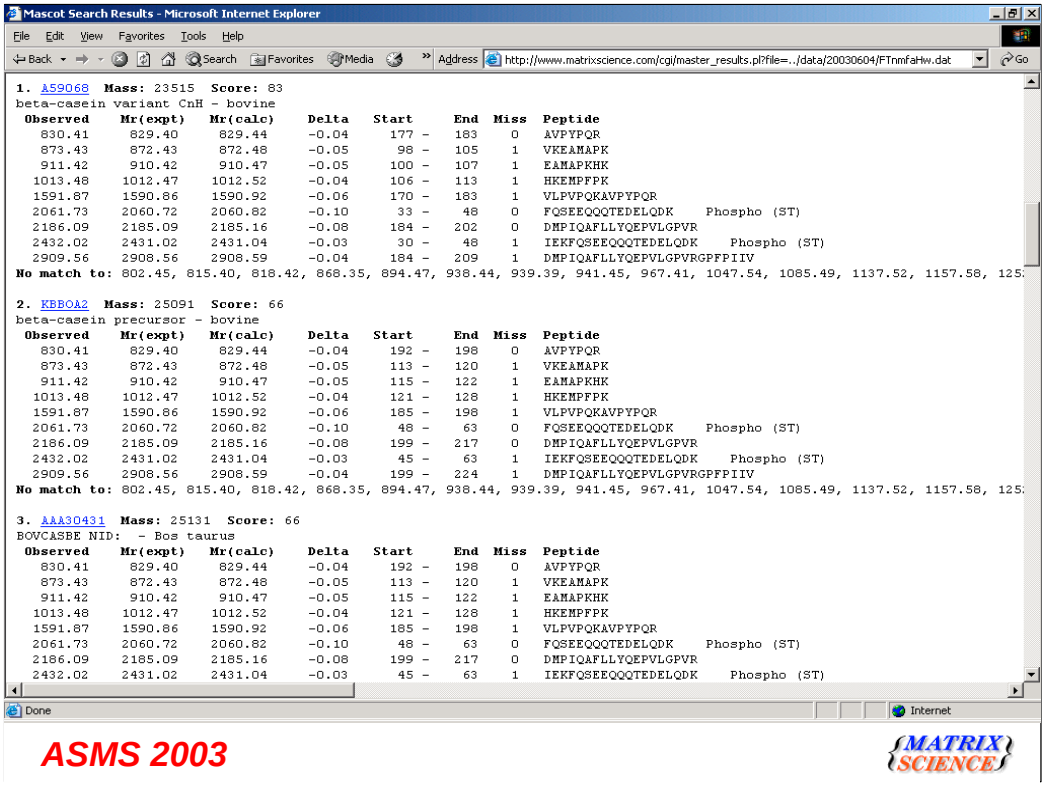
This slide shows that two proteins have scores above the threshold. One of them is beta-Casein variant CnH.



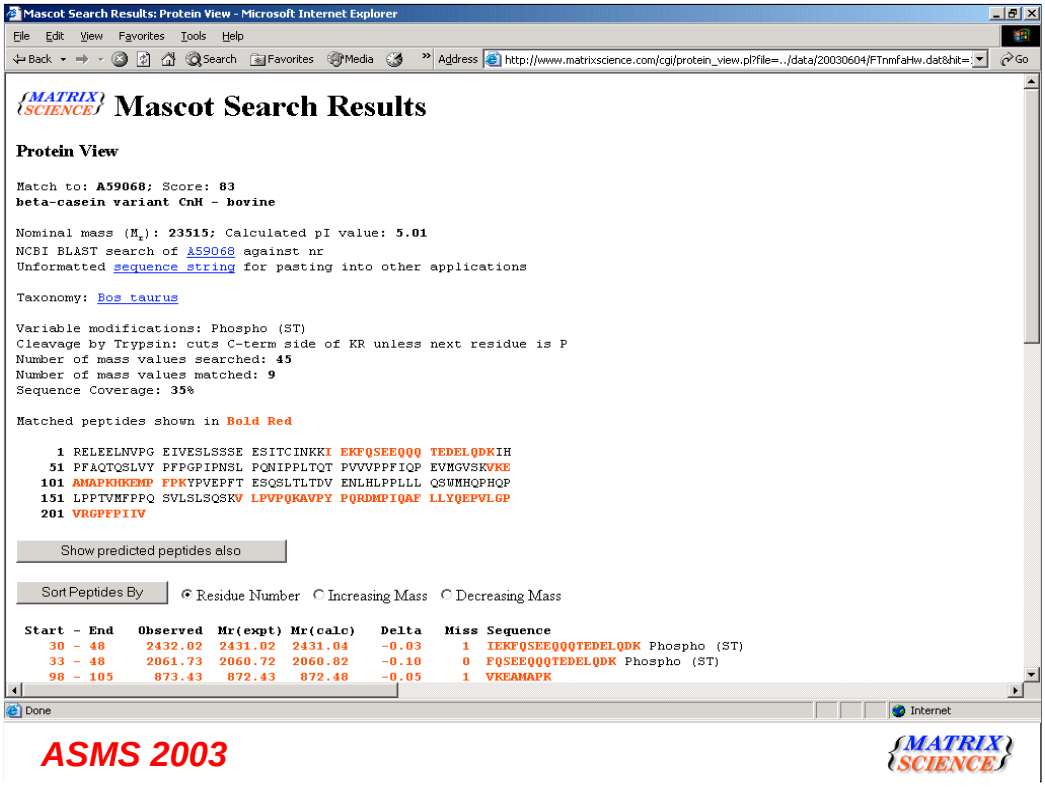
ASMS 2003

MATRIX SCIENCE

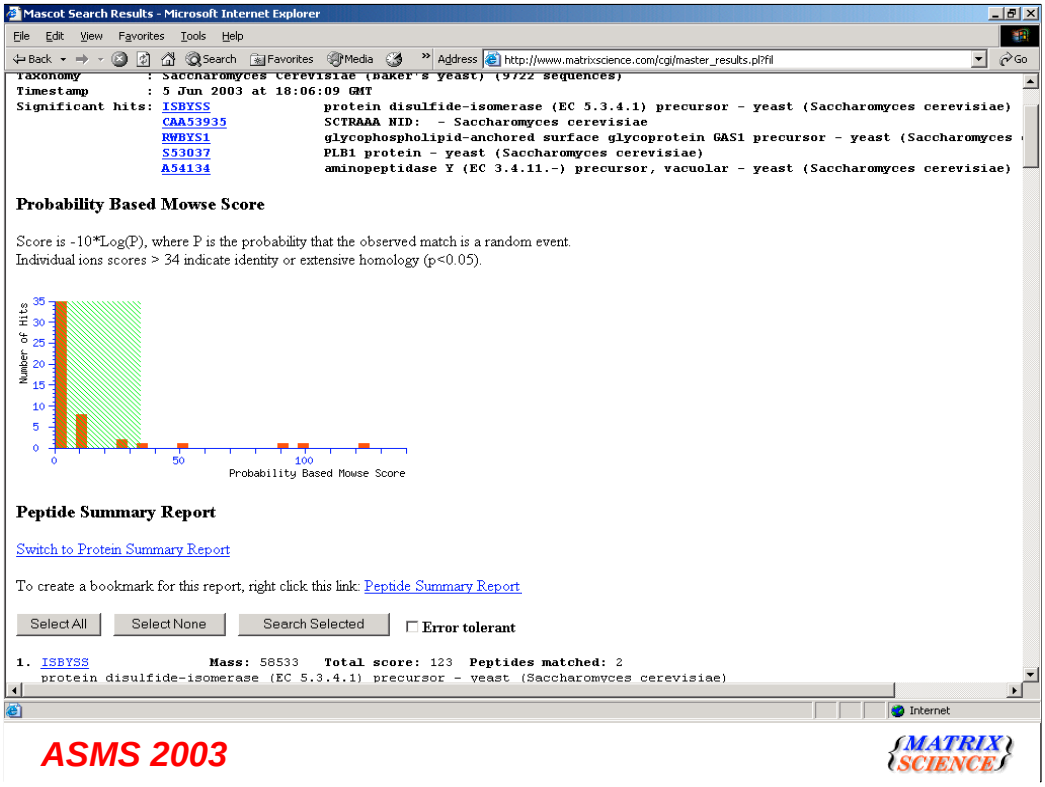
The other protein is beta-casein precursor.



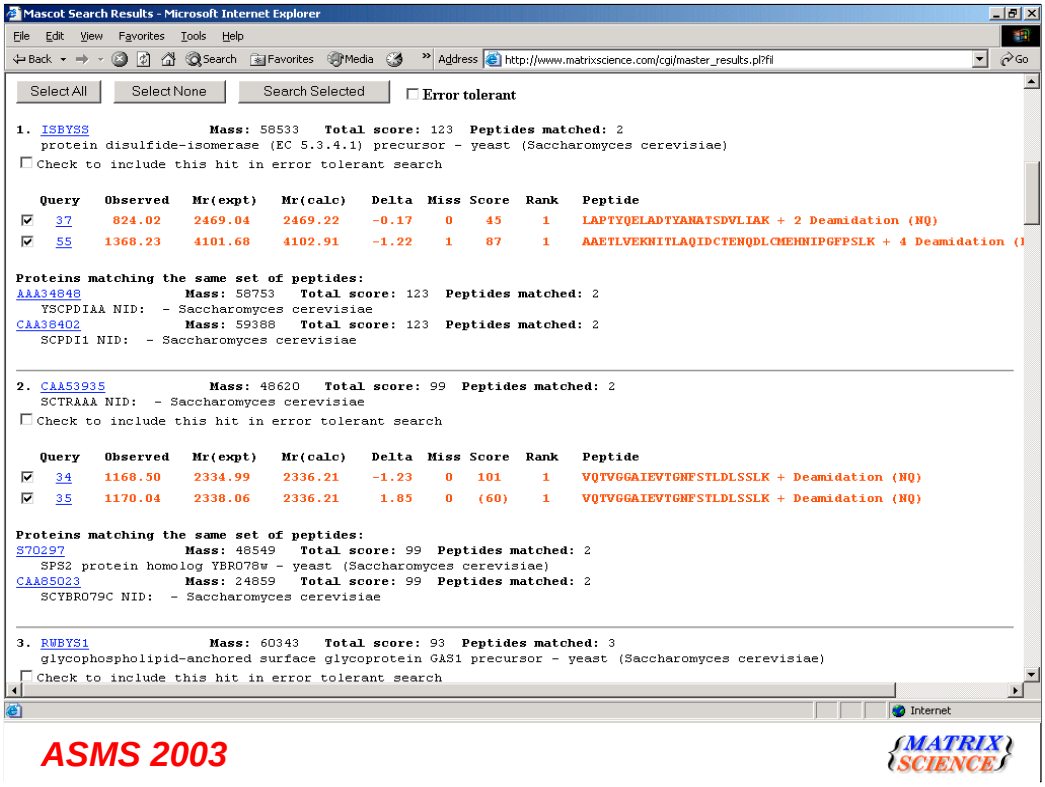
When we look at the detail, we see that Serine or Threonine is phosphorylated.



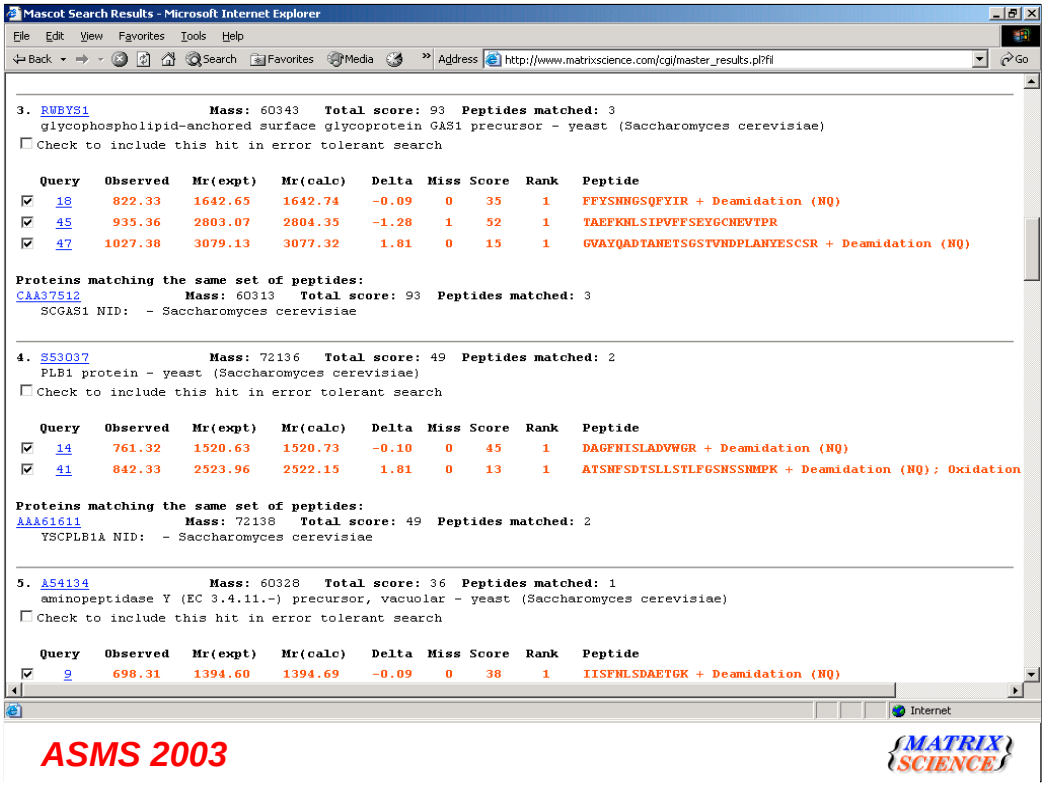
With Peptide Mass Fingerprint, it is not possible to map the modification sites. One needs tandem mass spectrometry to do that.



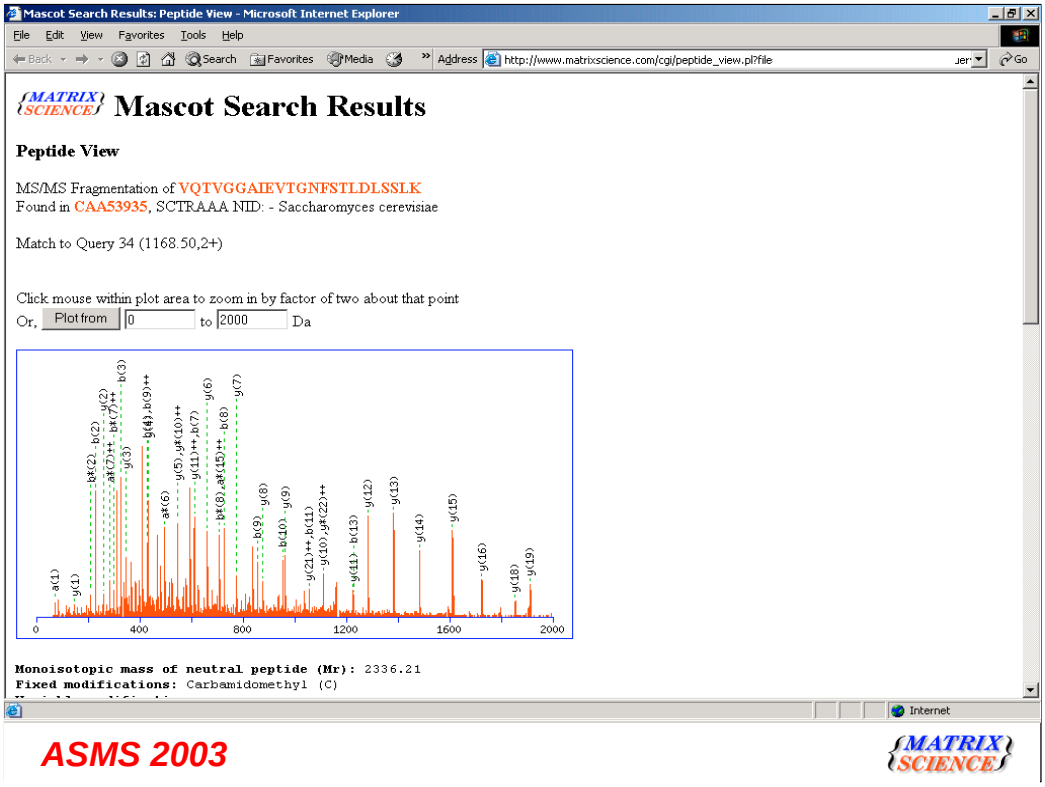
This slide shows at least 5 proteins with scores above the threshold.



Four peptides identified in this slide have scores higher than 34, the significance threshold. Deamidation seems to be a common modification for all of them.



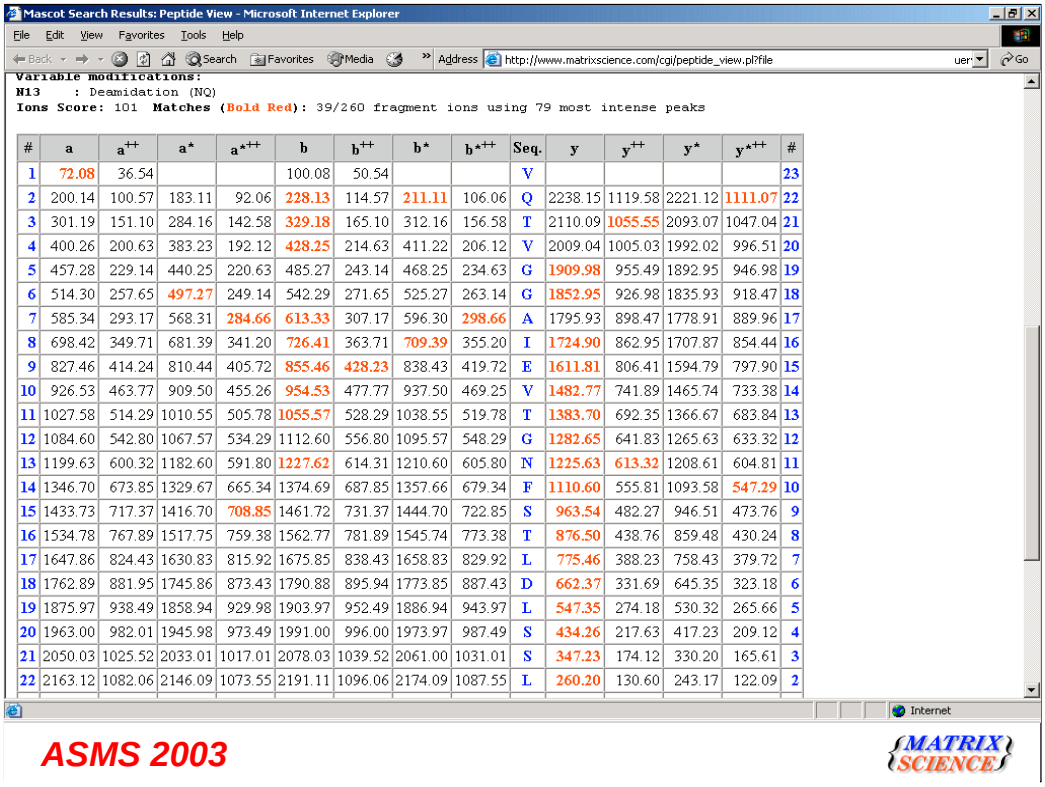
Here we see four other peptides that have scores higher than the significance threshold.



ASMS 2003

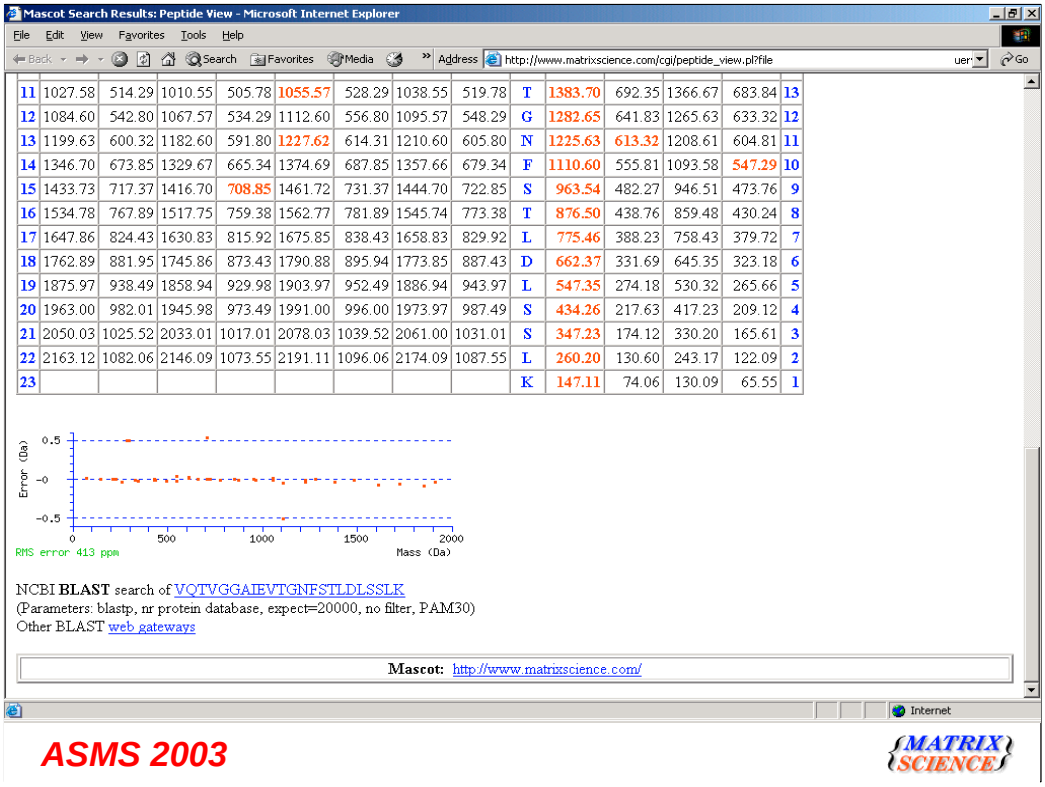
MATRIX SCIENCE

Let's focus on Query 34, where the ion score was 101. We see a rich MS/MS spectrum.

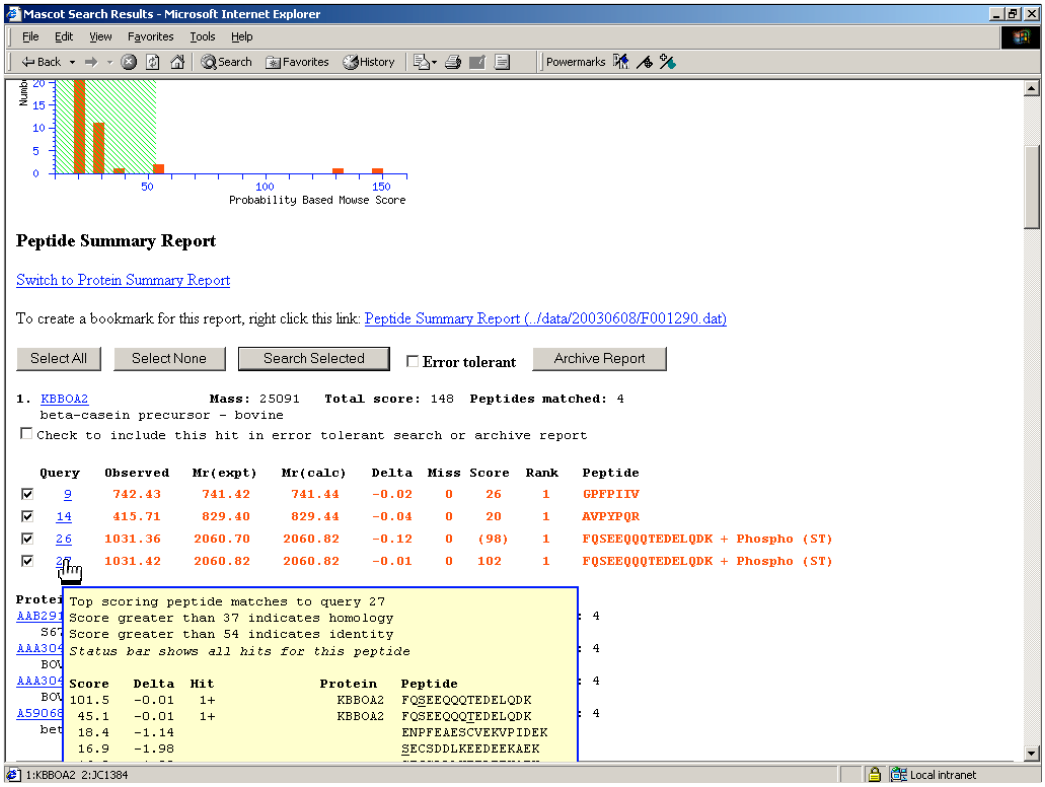


#	a	a^{++}	a*	a^{*++}	b	b^{++}	b*	b^{*++}	Seq.	y	y^{++}	y*	y^{*++}	#
1	72.08	36.54			100.08	50.54			V					23
2	200.14	100.57	183.11	92.06	228.13	114.57	211.11	106.06	Q	2238.15	1119.58	2221.12	1111.07	22
3	301.19	151.10	284.16	142.58	329.18	165.10	312.16	156.58	T	2110.09	1055.55	2093.07	1047.04	21
4	400.26	200.63	383.23	192.12	428.25	214.63	411.22	206.12	V	2009.04	1005.03	1992.02	996.51	20
5	457.28	229.14	440.25	220.63	485.27	243.14	468.25	234.63	G	1909.98	955.49	1892.95	946.98	19
6	514.30	257.65	497.27	249.14	542.29	271.65	525.27	263.14	G	1852.95	926.98	1835.93	918.47	18
7	585.34	293.17	568.31	284.66	613.33	307.17	596.30	298.66	A	1795.93	898.47	1778.91	889.96	17
8	698.42	349.71	681.39	341.20	726.41	363.71	709.39	355.20	I	1724.90	862.95	1707.87	854.44	16
9	827.46	414.24	810.44	405.72	855.46	428.23	838.43	419.72	E	1611.81	806.41	1594.79	797.90	15
10	926.53	463.77	909.50	455.26	954.53	477.77	937.50	469.25	V	1482.77	741.89	1465.74	733.38	14
11	1027.58	514.29	1010.55	505.78	1055.57	528.29	1038.55	519.78	T	1383.70	692.35	1366.67	683.84	13
12	1084.60	542.80	1067.57	534.29	1112.60	556.80	1095.57	548.29	G	1282.65	641.83	1265.63	633.32	12
13	1199.63	600.32	1182.60	591.80	1227.62	614.31	1210.60	605.80	N	1225.63	613.32	1208.61	604.81	11
14	1346.70	673.85	1329.67	665.34	1374.69	687.85	1357.66	679.34	F	1110.60	555.81	1093.58	547.29	10
15	1433.73	717.37	1416.70	708.85	1461.72	731.37	1444.70	722.85	S	963.54	482.27	946.51	473.76	9
16	1534.78	767.89	1517.75	759.38	1562.77	781.89	1545.74	773.38	T	876.50	438.76	859.48	430.24	8
17	1647.86	824.43	1630.83	815.92	1675.85	838.43	1658.83	829.92	L	775.46	388.23	758.43	379.72	7
18	1762.89	881.95	1745.86	873.43	1790.88	895.94	1773.85	887.43	D	662.37	331.69	645.35	323.18	6
19	1875.97	938.49	1858.94	929.98	1903.97	952.49	1886.94	943.97	L	547.35	274.18	530.32	265.66	5
20	1963.00	982.01	1945.98	973.49	1991.00	996.00	1973.97	987.49	S	434.26	217.63	417.23	209.12	4
21	2050.03	1025.52	2033.01	1017.01	2078.03	1039.52	2061.00	1031.01	S	347.23	174.12	330.20	165.61	3
22	2163.12	1082.06	2146.09	1073.55	2191.11	1096.06	2174.09	1087.55	L	260.20	130.60	243.17	122.09	2

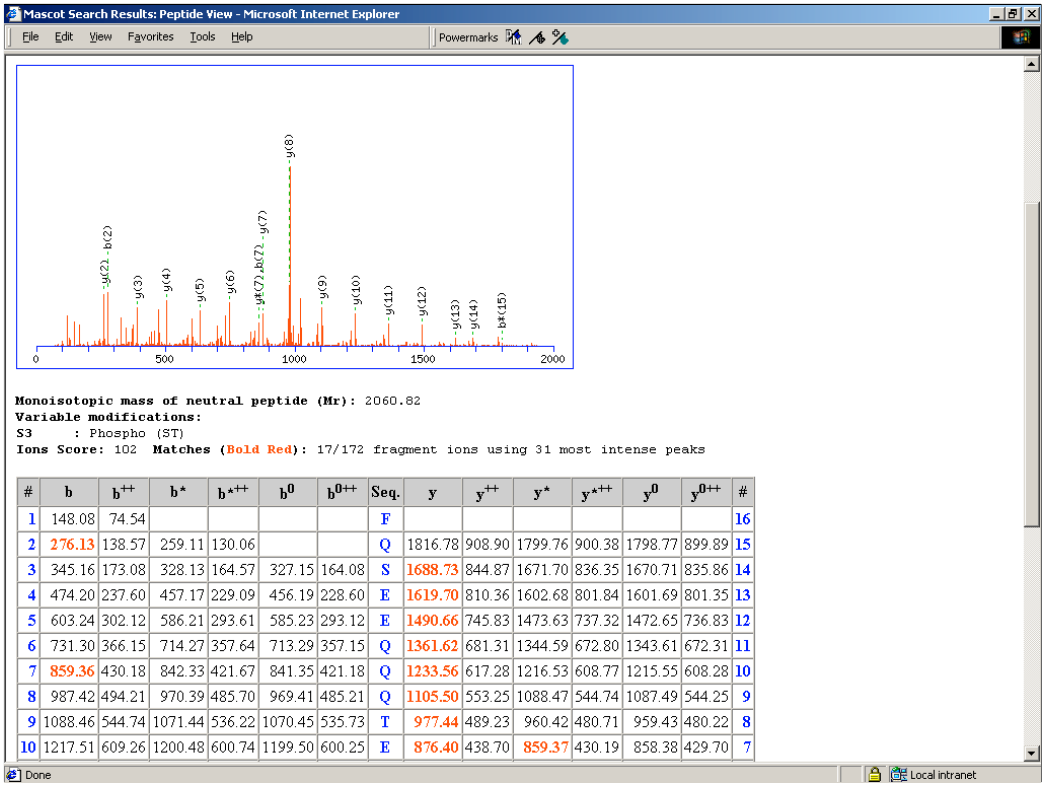
Most of the Y ions have been identified.



From the error graph, one can see that the calibration is quite good. I would believe that this peptide is deamidated.

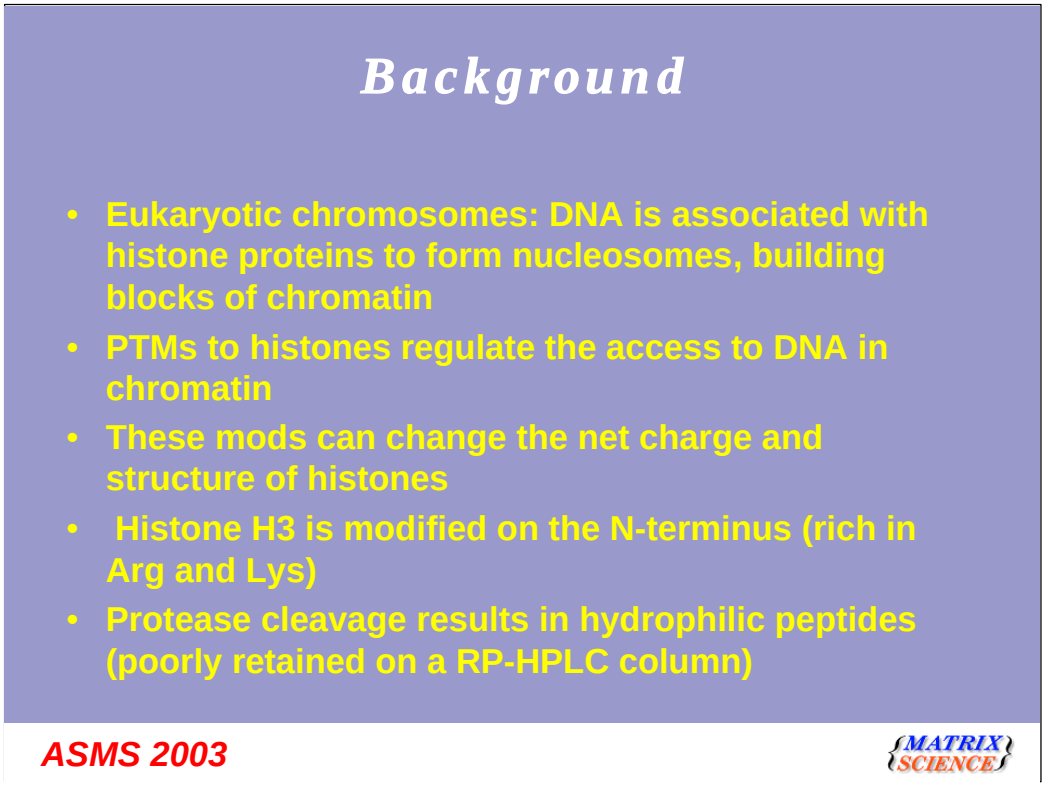


This example shows a phosphorylated peptide. There are two possible phosphorylation sites, but the score for phospho-serine (102) is much, much greater than that for phospho-threonine. I such cases, there is no doubt about the location of the modification



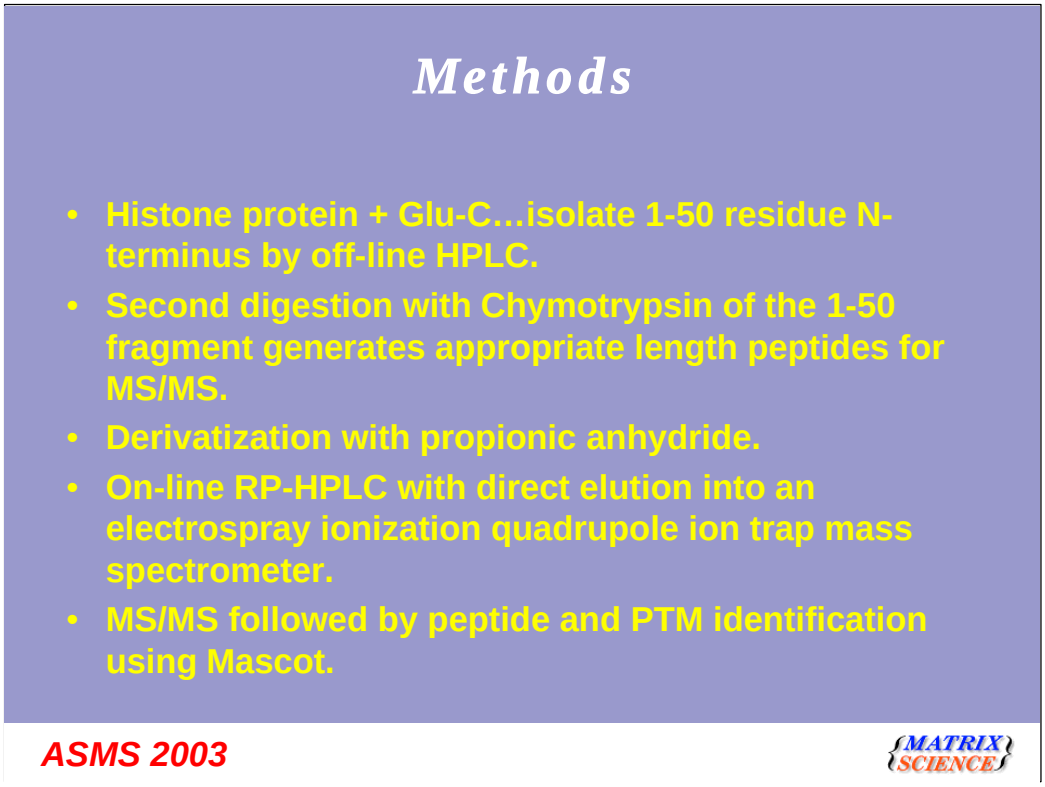
#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	148.08	74.54					F							16
2	276.13	138.57	259.11	130.06			Q	1816.78	908.90	1799.76	900.38	1798.77	899.89	15
3	345.16	173.08	328.13	164.57	327.15	164.08	S	1688.73	844.87	1671.70	836.35	1670.71	835.86	14
4	474.20	237.60	457.17	229.09	456.19	228.60	E	1619.70	810.36	1602.68	801.84	1601.69	801.35	13
5	603.24	302.12	586.21	293.61	585.23	293.12	E	1490.66	745.83	1473.63	737.32	1472.65	736.83	12
6	731.30	366.15	714.27	357.64	713.29	357.15	Q	1361.62	681.31	1344.59	672.80	1343.61	672.31	11
7	859.36	430.18	842.33	421.67	841.35	421.18	Q	1233.56	617.28	1216.53	608.77	1215.55	608.28	10
8	987.42	494.21	970.39	485.70	969.41	485.21	Q	1105.50	553.25	1088.47	544.74	1087.49	544.25	9
9	1088.46	544.74	1071.44	536.22	1070.45	535.73	T	977.44	489.23	960.42	480.71	959.43	480.22	8
10	1217.51	609.26	1200.48	600.74	1199.50	600.25	E	876.40	438.70	859.37	430.19	858.38	429.70	7

The spectrum shows a strong run of y ions that have lost phosphate as a neutral loss

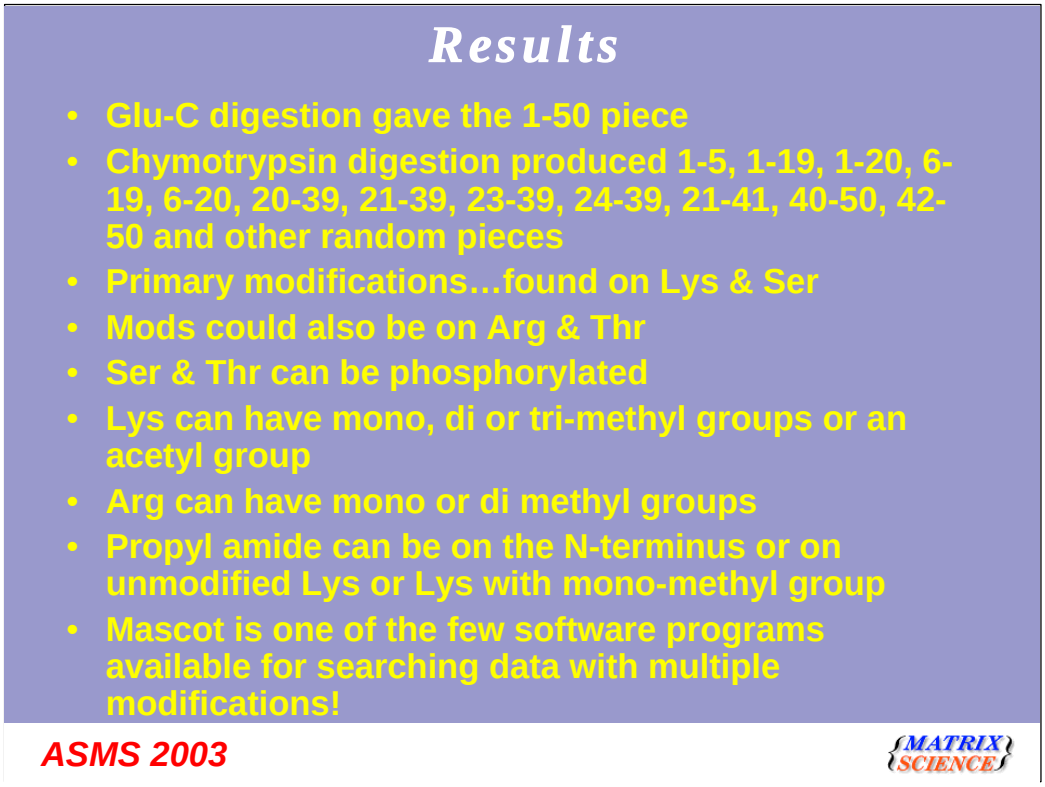


In eukaryotic chromosomes, large amounts of DNA are compacted by association with histone proteins to form nucleosomes, the building blocks of chromatin. Access to DNA in chromatin is regulated by post-translational modifications (PTMs) (methylation, phosphorylation, etc.) to histones. Such modifications can change the histone's net charge and structure thus playing a pivotal role in the control of chromatin structure and function. The majority of histone H3 modifications occur on the 1-50 residue N-terminal tail, which is rich in Arginine and Lysine residues. This tail produces peptides upon protease cleavage that are very hydrophilic and poorly retained on a RP-HPLC column.

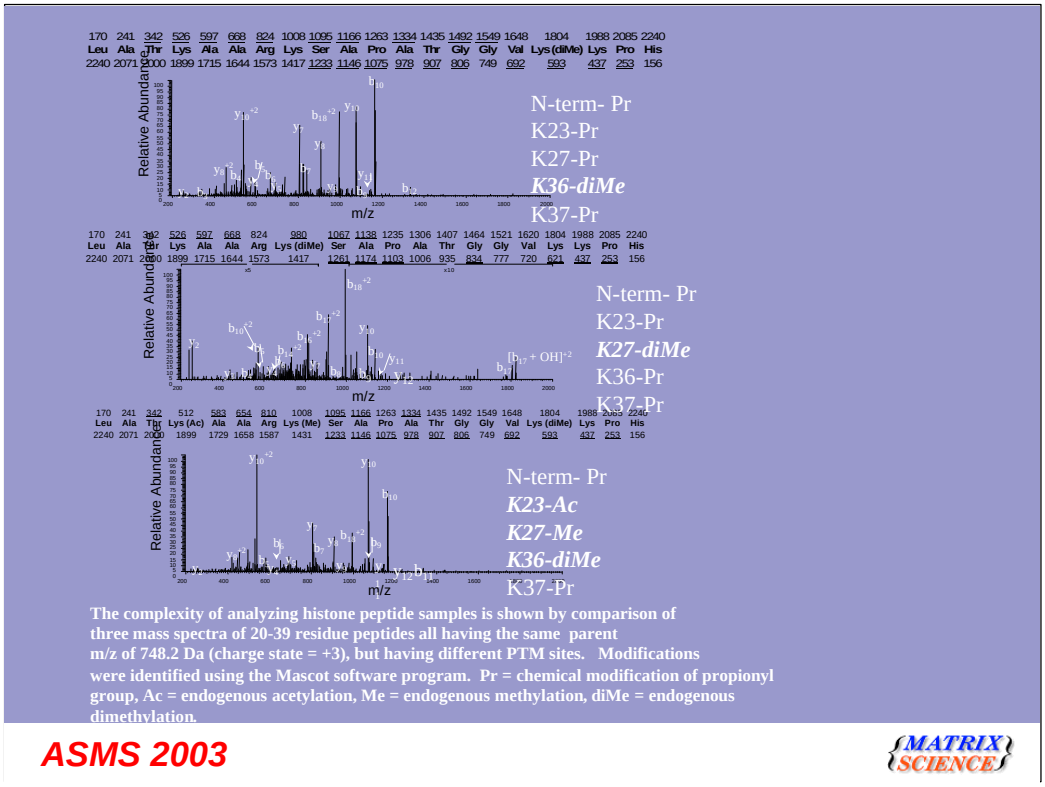
The collisionally-activated dissociation (CAD) mass spectra of highly multiply charged peptides are very difficult to interpret. Their research efforts are placed on developing methods to better analyze histones by HPLC and mass spectrometry. Treatment of peptides with propionic anhydride converts free amine groups on unmodified Lysines or Lysine residues containing 1 methyl group modification and N-termini to propyl amides. The consequence of the above strategy is a decrease in the net charge of the peptides, as well as increased hydrophobicity thus facilitating their analysis by increasing retention times on an HPLC column and simplifying their CAD mass spectra.



Their methods begin with an enzymatic digestion of the histone protein with Glu-C and isolation of the 1-50 residue amino terminus by off-line HPLC. A second digestion with Chymotrypsin is performed on the 1-50 fragment to produce peptides of suitable lengths for tandem mass spectrometry experiments. Peptides are then derivatized by the addition of the propionic anhydride reagent. The mixture of peptides are separated by on-line RP-HPLC before direct elution into an electrospray ionization quadrupole ion trap mass spectrometer. Tandem mass spectrometry experiments are then performed to fragment the ions and determine post-translational modification sites after searching with Mascot.



The digestion of the H3 protein with Glu-C cleaved the protein to produce the 1-50 piece further isolated by off-line HPLC. A Chymotrypsin digestion of the 1-50 piece primarily produced peptide residues of 1-5, 1-19, 1-20, 6-19, 6-20, 20-39, 21-39, 23-39, 24-39, 21-41, 40-50, 42-50 and other random pieces. The primary modifications found were on Lys and Ser residues. However, modifications could also be present on Arg and Thr residues. Ser and Thr can be modified by the addition of a phosphate group. Lys can be modified by the addition of mono, di or tri-methyl groups or by the addition of an acetyl group. Arg residues can be modified by the addition of one or two methyl groups. In addition, adding propionic anhydride creates propyl amide groups on the amino terminus of peptides and peptides containing unmodified Lys residues or Lys residues modified by one methyl group. Lys residues containing di, tri-methyl or acetyl groups are not modified by the propionic anhydride reagent. Thus, as can be easily seen, a vast amount of modifications and combination of modifications can be expected when analyzing histone peptides. Currently, Mascot is one of the few software programs available that is able to search data possibly containing multiple modifications. Modifications of N-terminal (Pr) and Lys modifications of propyl amide (Pr), propyl amide and methyl (Pr-Me), di-methyl (di-Me) and tri-methyl (tri-Me), Ser (phosphorylation) and Arg modification of (Me) were created and Mascot was used to search the database.



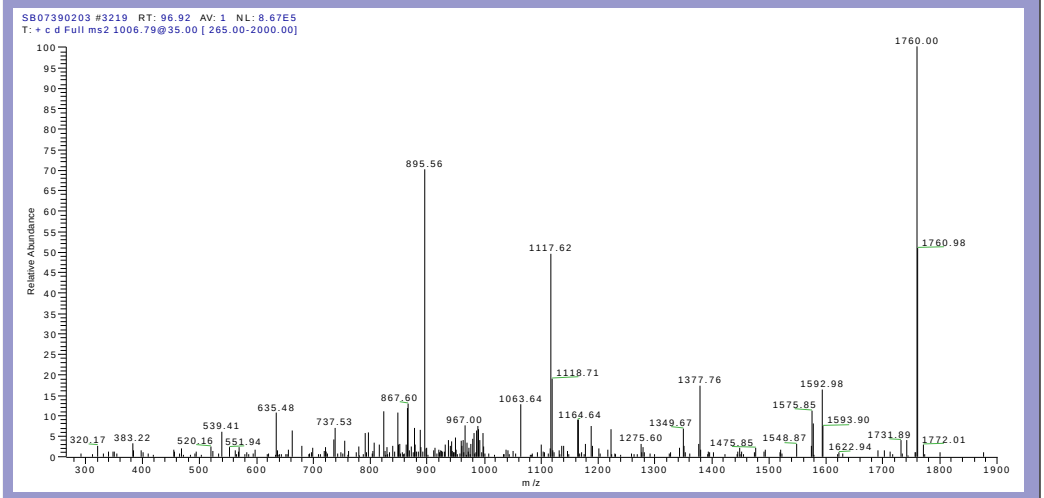
The complexity of analyzing histone peptide samples is shown by comparison of three mass spectra of 20-39 residue peptides all having the same parent m/z of 748.2 Da (charge state = +3), but having different PTM sites. Modifications were identified using the Mascot software program. Pr = chemical modification of propionyl group, Ac = endogenous acetylation, Me = endogenous methylation, diMe = endogenous dimethylation.

ASMS 2003

MATRIX SCIENCE

This slide shows three peptides with the same parent mass, but they differ in the PTM site(s). There are 4 Lysines that can be modified, there is the amino terminus and a Ser and an Arg residue. The possible combination of PTM sites is large. The bottom spectrum demonstrates how four different PTMs can be identified by Mascot. Mascot was used to identify combination of mods on histone peptides.

241	312	383	539	737	824	895	992	1063	1164	1221	1278	1377	1575	1759	1856	2011
Lys	**Ala**	**Ala**	**Arg**	**Lys (Me)**	**Ser**	**Ala**	**Pro**	**Ala**	**Thr**	**Gly**	**Gly**	**Val**	**Lys (Me)**	**Lys**	**Pro**	**His**
2011	1771	1700	1629	1473	1275	1188	1117	1020	949	848	791	734	635	437	253	156



ASMS 2003

MATRIX SCIENCE

Mascot correctly identified the endogenous Lys methylations on K27 and K36, as well as our chemical modifications of propionyl on K37 and K23 (amino terminus).

Mascot Search Results: Peptide View - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://noah/mascot/cgi/peptide_view.pl?File=../data/20021014/F001531.dat&query=39636&hit=18&index=gi%7c10198611&px=1

Links Customize Links Free Hotmail Windows

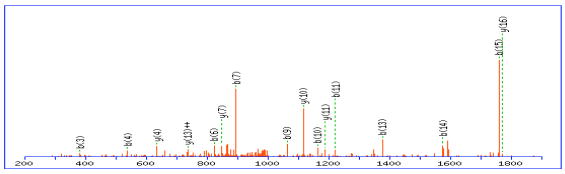
Mascot Search Results

Peptide View

MS/MS Fragmentation of KAARKSAPATGGVKKPH
Found in gi|10198611, dJ34B20.17 (H3 histone family, member B) [Homo sapiens]

Match to Query 3963 (1006.79,2+) sb07390203.3219.3219.2.dta
From data file \\Dfhserv\Users\Ben\Histones\H3 Mitotic Arrested\Fraction 22-23\sb07390203.RAW

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 200 to 1900 Da



Done Local intranet

Monoisotopic mass of neutral peptide (Mr): 2011.16
Fixed modifications: Pr (N-term)
Variable modifications:
K1 : Pr (K)
K5 : PrMe (K)
K14 : PrMe (K)
K15 : Pr (K)
Ions Score: 68 **Matches (Bold Red):** 17/164 fragment ions using 33 most intense peaks

#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	241.16	121.08	224.13	112.57			K							17
2	312.19	156.60	295.17	148.09			A	**1772.02**	886.51	1754.99	878.00	1754.01	877.51	16
3	**383.23**	192.12	366.20	183.61			A	1700.98	850.99	1683.95	842.48	1682.97	841.99	15
4	**539.33**	270.17	522.30	261.66			R	1629.94	815.48	1612.92	806.96	1611.93	806.47	14
5	**737.47**	369.24	720.44	360.72			K	1473.84	**737.43**	1456.82	728.91	1455.83	728.42	13
6	**824.50**	412.75	807.47	404.24	806.49	403.75	S	1275.71	638.36	1258.68	629.84	1257.70	629.35	12
7	**895.54**	448.27	878.51	439.76	877.53	439.27	A	**1188.67**	594.84	1171.65	586.33	1170.66	585.84	11
8	992.59	496.80	975.56	488.29	974.58	487.79	P	**1117.64**	559.32	1100.61	550.81	1099.63	550.32	10
9	**1063.63**	532.32	1046.60	523.80	1045.62	523.31	A	1020.58	510.80	1003.56	502.28	1002.57	501.79	9
10	**1164.67**	582.84	1147.65	574.33	1146.66	573.84	T	949.55	475.28	932.52	466.76	931.54	466.27	8
11	**1221.70**	611.35	1204.67	602.84	1203.68	602.35	G	**848.50**	424.75	831.47	416.24			7
12	1278.72	639.86	1261.69	631.35	1260.71	630.86	G	791.48	396.24	774.45	387.73			6
13	**1377.79**	689.40	1360.76	680.88	1359.77	680.39	V	734.46	367.73	717.43	359.22			5
14	**1575.92**	788.47	1558.90	779.95	1557.91	779.46	K	**635.39**	318.20	618.36	309.68			4
15	**1760.04**	880.53	1743.02	872.01	1742.03	871.52	K	437.25	219.13	420.22	210.62			3
16	1857.10	929.05	1840.07	920.54	1839.09	920.05	P	253.13	127.07					2
17							H	156.08	78.54					1

Done Local intranet

ASMS 2003

MATRIX SCIENCE

Poster

Tuesday (#941) entitled "Analysis of Human Histone H3 Post-Translational Modification Site Patterns from Cells Arrested During Mitosis by Tandem Mass Spectrometry"

Benjamin A. Garcia[1], Scott A. Busby[1], A. Celeste Dunsmoor[1], Jeffrey Shabanowitz[1],

Cynthia M. Barber[2], C. David Allis[2], Donald F. Hunt[1,3]

Departments of Chemistry[1] and Pathology[3], University of Virginia, Charlottesville, VA; Department of Biochemistry and Molecular Genetics[2], University of Virginia Health Science Center, Charlottesville, VA

ASMS 2003

MATRIX SCIENCE