

# *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***

Mascot Search Results - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.matrixscience.com/cgi/master\_results.pl?file=.../data/20030604/FTnmfaHw.dat

**Mascot Search Results**

User : Tanuja Chaudhary  
 Email : tanuja@matrixscience.com  
 Search title :  
 Database : MSDB 20030428 (1165316 sequences; 370264913 residues)  
 Taxonomy : Other mammalia (26750 sequences)  
 Timestamp : 4 Jun 2003 at 15:47:37 GMT  
 Top Score : 83 for **A59068**, beta-casein variant CnH - bovine

**Probability Based Mowse Score**

Score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event.  
 Protein scores greater than 57 are significant ( $p < 0.05$ ).

**Protein Summary Report**

[Switch to Concise Protein Summary Report](#)

To create a bookmark for this report, right click this link: [Protein Summary Report \(./data/20030604/FTnmfaHw.dat\)](#)

Done Internet

**ASMS 2003**

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

Mascot Search Results - Microsoft Internet Explorer

Address: http://www.matrixscience.com/cgi/master\_results.pl?file=../data/20030604/FTnmfaHw.dat

[Switch to Concise Protein Summary Report](#)

To create a bookmark for this report, right click this link: [Protein Summary Report \(/data/20030604/FTnmfaHw.dat\)](#)

Re-Search All    Search Unmatched

### Index

Accession	Mass	Score	Description
1. <a href="#">A59068</a>	23515	83	beta-casein variant CnH - bovine
2. <a href="#">KBB0A2</a>	25091	66	beta-casein precursor - bovine
3. <a href="#">AAA30431</a>	25131	66	BOVCASB NID: - Bos taurus
4. <a href="#">Q9GM99</a>	344583	52	Huntingtin.- Sus scrofa (Pig).
5. <a href="#">Q9BDG5</a>	16443	50	Beta casein B (Fragment).- Bos taurus (Bovine).
6. <a href="#">Q9TSD5</a>	13915	48	Beta-casein A2 variant (Fragments).- Bos taurus (Bovine).
7. <a href="#">CAA06408</a>	25090	45	BBAJ5165 NID: - Bubalus bubalis
8. <a href="#">AAB29137</a>	25082	44	S67277 NID: - Bos taurus
9. <a href="#">CAA34450</a>	7647	42	OCCMYA2R NID: - Oryctolagus cuniculus
10. <a href="#">Q9WMK3</a>	116803	42	Vinculin.- Sus scrofa (Pig).
11. <a href="#">P79704</a>	6751	42	Sex determining region Y (Fragment).- Choloepus didactylus (southern two-toed sloth)
12. <a href="#">A34154</a>	48245	42	calreticulin precursor, skeletal muscle - rabbit
13. <a href="#">AAA30430</a>	25072	41	BOVCASB NID: - Bos taurus
14. <a href="#">Q9WN43</a>	39619	37	Recombination activating protein 1 (Fragment).- Equus caballus (Horse).
15. <a href="#">MY01</a>	16967	36	myoglobin - southern American pika
16. <a href="#">Q9WH44</a>	40526	36	Recombination activating protein 1 (Fragment).- Talpa europaea (European mole).
17. <a href="#">Q9BEW5</a>	29028	36	Recombination activating protein 1 (Fragment).- Ceratotherium simum (White rhinoceros)
18. <a href="#">MY0H</a>	16973	36	myoglobin - western European hedgehog
19. <a href="#">Q9WN53</a>	41149	36	Recombination activating protein 1 (Fragment).- Nycteris grandis.
20. <a href="#">B34078</a>	24999	35	prolactin-related protein III - bovine

### Results List

1. [A59068](#)    Mass: 23515    Score: 83  
beta-casein variant CnH - bovine

Observed	Mr (expt)	Mr (calc)	Delta	Start	End	Miss	Peptide
830.41	829.40	829.44	-0.04	177	183	0	AVPYQPQR
873.43	873.43	873.48	0.05	98	106	1	VYKIMADY

ASMS 2003

MATRIX SCIENCE

The other protein is beta-casein precursor.

Mascot Search Results - Microsoft Internet Explorer

Address: http://www.matrixscience.com/cgi/master\_results.pl?file=.../data/20030604/FTnmfahw.dat

1. [A59068](#) Mass: 23515 Score: 83  
beta-casein variant CnH - bovine

Observed	Mr (expt)	Mr (calc)	Delta	Start	End	Miss	Peptide
830.41	829.40	829.44	-0.04	177	183	0	AVPYYPQR
873.43	872.43	872.48	-0.05	98	105	1	VKEAMAPK
911.42	910.42	910.47	-0.05	100	107	1	EAMAPKHK
1013.48	1012.47	1012.52	-0.04	106	113	1	HKEMPFK
1591.87	1590.86	1590.92	-0.06	170	183	1	VLPVQKAVPYYPQR
2061.73	2060.72	2060.82	-0.10	33	48	0	FQSEEQQTEDELQDK Phospho (ST)
2186.09	2185.09	2185.16	-0.08	184	202	0	DMPIQAFLLYQEPVLPVPR
2432.02	2431.02	2431.04	-0.03	30	48	1	IEKFSQSEEQQTEDELQDK Phospho (ST)
2909.56	2908.56	2908.59	-0.04	184	209	1	DMPIQAFLLYQEPVLPVPRGPFPIIV

No match to: 802.45, 815.40, 818.42, 868.35, 894.47, 938.44, 939.39, 941.45, 967.41, 1047.54, 1085.49, 1137.52, 1157.58, 125.

2. [KBBOA2](#) Mass: 25091 Score: 66  
beta-casein precursor - bovine

Observed	Mr (expt)	Mr (calc)	Delta	Start	End	Miss	Peptide
830.41	829.40	829.44	-0.04	192	198	0	AVPYYPQR
873.43	872.43	872.48	-0.05	113	120	1	VKEAMAPK
911.42	910.42	910.47	-0.05	115	122	1	EAMAPKHK
1013.48	1012.47	1012.52	-0.04	121	128	1	HKEMPFK
1591.87	1590.86	1590.92	-0.06	185	198	1	VLPVQKAVPYYPQR
2061.73	2060.72	2060.82	-0.10	48	63	0	FQSEEQQTEDELQDK Phospho (ST)
2186.09	2185.09	2185.16	-0.08	199	217	0	DMPIQAFLLYQEPVLPVPR
2432.02	2431.02	2431.04	-0.03	45	63	1	IEKFSQSEEQQTEDELQDK Phospho (ST)
2909.56	2908.56	2908.59	-0.04	199	224	1	DMPIQAFLLYQEPVLPVPRGPFPIIV

No match to: 802.45, 815.40, 818.42, 868.35, 894.47, 938.44, 939.39, 941.45, 967.41, 1047.54, 1085.49, 1137.52, 1157.58, 125.

3. [AAA30431](#) Mass: 25131 Score: 66  
BOVCASBE NID: - Bos taurus

Observed	Mr (expt)	Mr (calc)	Delta	Start	End	Miss	Peptide
830.41	829.40	829.44	-0.04	192	198	0	AVPYYPQR
873.43	872.43	872.48	-0.05	113	120	1	VKEAMAPK
911.42	910.42	910.47	-0.05	115	122	1	EAMAPKHK
1013.48	1012.47	1012.52	-0.04	121	128	1	HKEMPFK
1591.87	1590.86	1590.92	-0.06	185	198	1	VLPVQKAVPYYPQR
2061.73	2060.72	2060.82	-0.10	48	63	0	FQSEEQQTEDELQDK Phospho (ST)
2186.09	2185.09	2185.16	-0.08	199	217	0	DMPIQAFLLYQEPVLPVPR
2432.02	2431.02	2431.04	-0.03	45	63	1	IEKFSQSEEQQTEDELQDK Phospho (ST)

ASMS 2003

MATRIX SCIENCE

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

Mascot Search Results: Protein View - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://www.matrixscience.com/cgi/protein\_view.pl?file=../data/20030604/FTnmfahw.dat&hit=

## Mascot Search Results

### Protein View

Match to: **A59068**; Score: **83**  
**beta-casein variant CnH - bovine**

Nominal mass ( $M_r$ ): **23515**; Calculated pI value: **5.01**  
 NCBI BLAST search of **A59068** against nr  
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Bos taurus](#)

Variable modifications: Phospho (ST)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Number of mass values searched: **45**  
 Number of mass values matched: **9**  
 Sequence Coverage: **35%**

Matched peptides shown in **Bold Red**

1 RELEELNVPG EIVESLSSE ESITCINKKI **EKFQSEEQQQ TEDELQDKIH**  
 51 PFAQTQSLVY PFGPIPNLS PQNIPPLTQT PVVVPPFIQP EVMGVSKVKE  
 101 **AMAPKHKEMP** FPKYPVEPFT ESQSLTLTDV ENLHLPPLL QSMHQPHQP  
 151 LPPTVMFPQ SVLSLSQSKV **LPVQKAVPY** PQRDMPIQAF LLYQEPVLP  
 201 **VRGPFPIIV**

Show predicted peptides also

Sort Peptides By  Residue Number  Increasing Mass  Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
30 - 48	2432.02	2431.02	2431.04	-0.03	1	TEKFQSEEQQQTEDELQDK Phospho (ST)
33 - 48	2061.73	2060.72	2060.82	-0.10	0	FQSEEQQQTEDELQDK Phospho (ST)
98 - 105	873.43	872.43	872.48	-0.05	1	VKEAMAPK

ASMS 2003

MATRIX SCIENCE

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

Mascot Search Results - Microsoft Internet Explorer

Address: [http://www.matrixscience.com/cgi/master\\_results.pl?fil](http://www.matrixscience.com/cgi/master_results.pl?fil)

Taxonomy : **Saccharomyces cerevisiae (baker's yeast)** (9722 sequences)

Timestamp : 5 Jun 2003 at 18:06:09 GMT

Significant hits:

<a href="#">ISBYSS</a>	protein disulfide-isomerase (EC 5.3.4.1) precursor - yeast (Saccharomyces cerevisiae)
<a href="#">CAA53935</a>	SCTRAAA MID: - Saccharomyces cerevisiae
<a href="#">RWBYS1</a>	glycophospholipid-anchored surface glycoprotein GAS1 precursor - yeast (Saccharomyces cerevisiae)
<a href="#">S53037</a>	PLB1 protein - yeast (Saccharomyces cerevisiae)
<a href="#">A54134</a>	aminopeptidase Y (EC 3.4.11.-) precursor, vacuolar - yeast (Saccharomyces cerevisiae)

**Probability Based Mowse Score**

Score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event.  
Individual ions scores  $> 34$  indicate identity or extensive homology ( $p < 0.05$ ).

**Peptide Summary Report**

[Switch to Protein Summary Report](#)

To create a bookmark for this report, right click this link: [Peptide Summary Report](#)

Select All    Select None    Search Selected     Error tolerant

1. [ISBYSS](#)    Mass: 58533    Total score: 123    Peptides matched: 2  
protein disulfide-isomerase (EC 5.3.4.1) precursor - yeast (Saccharomyces cerevisiae)

ASMS 2003

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

Mascot Search Results - Microsoft Internet Explorer

Address: http://www.matrixscience.com/cgi/master\_results.pl?fil

Select All    Select None    Search Selected     Error tolerant

1. [ISBY53](#)    Mass: 58533    Total score: 123    Peptides matched: 2  
 protein disulfide-isomerase (EC 5.3.4.1) precursor - yeast (Saccharomyces cerevisiae)  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> 37	824.02	2469.04	2469.22	-0.17	0	45	1	LAPTYQELADTYANATSDVLIAK + 2 Deamidation (NQ)
<input checked="" type="checkbox"/> 55	1368.23	4101.68	4102.91	-1.22	1	87	1	AAETLVEKHITLAQIDCTENQDLCEHNIPGPEPSLK + 4 Deamidation (NQ)

Proteins matching the same set of peptides:  
[AAK34848](#)    Mass: 58753    Total score: 123    Peptides matched: 2  
 YSCPDIAA NID: - Saccharomyces cerevisiae  
[CAA38402](#)    Mass: 59388    Total score: 123    Peptides matched: 2  
 SCPDI1 NID: - Saccharomyces cerevisiae

---

2. [CAA53935](#)    Mass: 48620    Total score: 99    Peptides matched: 2  
 SCTRAAA NID: - Saccharomyces cerevisiae  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> 34	1168.50	2334.99	2336.21	-1.23	0	101	1	VQTVGGAEIVTGNFSTLDLSSLK + Deamidation (NQ)
<input checked="" type="checkbox"/> 35	1170.04	2338.06	2336.21	1.85	0	(60)	1	VQTVGGAEIVTGNFSTLDLSSLK + Deamidation (NQ)

Proteins matching the same set of peptides:  
[S70297](#)    Mass: 48549    Total score: 99    Peptides matched: 2  
 SPS2 protein homolog YBR078w - yeast (Saccharomyces cerevisiae)  
[CAA85023](#)    Mass: 24859    Total score: 99    Peptides matched: 2  
 SCYBR079C NID: - Saccharomyces cerevisiae

---

3. [RNBYS1](#)    Mass: 60343    Total score: 93    Peptides matched: 3  
 glycopospholipid-anchored surface glycoprotein GAS1 precursor - yeast (Saccharomyces cerevisiae)  
 Check to include this hit in error tolerant search

ASMS 2003 

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.

Mascot Search Results - Microsoft Internet Explorer

Address: http://www.matrixscience.com/cgi/master\_results.pl?fi

3. [RWBYS1](#) Mass: 60343 Total score: 93 Peptides matched: 3  
 glycopospholipid-anchored surface glycoprotein GAS1 precursor - yeast (Saccharomyces cerevisiae)  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <a href="#">18</a>	822.33	1642.65	1642.74	-0.09	0	35	1	FFYSNNGSQFYIR + Deamidation (NQ)
<input checked="" type="checkbox"/> <a href="#">45</a>	935.36	2803.07	2804.35	-1.28	1	52	1	TAEFKNLSIPVFFSEYGCNEVTPR
<input checked="" type="checkbox"/> <a href="#">47</a>	1027.38	3079.13	3077.32	1.81	0	15	1	GVAYQADTANETSGSTVNDPLANYECSR + Deamidation (NQ)

Proteins matching the same set of peptides:  
[CAA37512](#) Mass: 60313 Total score: 93 Peptides matched: 3  
 SCGAS1 NID: - Saccharomyces cerevisiae

---

4. [S53037](#) Mass: 72136 Total score: 49 Peptides matched: 2  
 PLB1 protein - yeast (Saccharomyces cerevisiae)  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <a href="#">14</a>	761.32	1520.63	1520.73	-0.10	0	45	1	DAGFNISLADVWGR + Deamidation (NQ)
<input checked="" type="checkbox"/> <a href="#">41</a>	842.33	2523.96	2522.15	1.81	0	13	1	ATSNFSDTSLSTLFGSNSSNMPK + Deamidation (NQ); Oxidation

Proteins matching the same set of peptides:  
[AAA61611](#) Mass: 72138 Total score: 49 Peptides matched: 2  
 YSCPLB1A NID: - Saccharomyces cerevisiae

---

5. [A54134](#) Mass: 60328 Total score: 36 Peptides matched: 1  
 aminopeptidase Y (EC 3.4.11.-) precursor, vacuolar - yeast (Saccharomyces cerevisiae)  
 Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
<input checked="" type="checkbox"/> <a href="#">9</a>	698.31	1394.60	1394.69	-0.09	0	38	1	IISFNLSDAETGK + Deamidation (NQ)

ASMS 2003 

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

Mascot Search Results: Peptide View - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address [http://www.matrixscience.com/cgi/peptide\\_view.pl?file](http://www.matrixscience.com/cgi/peptide_view.pl?file)

**MASCOT** **SCIENCE** **Mascot Search Results**

**Peptide View**

MS/MS Fragmentation of **VQTVGGAIEVTGNFSTLDSLK**  
 Found in **CAA53935**, SCTRAAA NID: - Saccharomyces cerevisiae

Match to Query 34 (1168.50,2+)

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

Monoisotopic mass of neutral peptide (Mr): 2336.21  
 Fixed modifications: Carbamidomethyl (C)

ASMS 2003 **MATRIX**  
**SCIENCE**

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

Mascot Search Results: Peptide View - Microsoft Internet Explorer

Address: http://www.matrixscience.com/cgi/peptide\_view.pl?file

Variable modifications:  
 N13 : Deamidation (NQ)  
 Ions Score: 101 Matches (Bold Red): 39/260 fragment ions using 79 most intense peaks

#	a	a <sup>++</sup>	a <sup>*</sup>	a <sup>+++</sup>	b	b <sup>++</sup>	b <sup>*</sup>	b <sup>+++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>+++</sup>	#
1	<b>72.08</b>	36.54			100.08	50.54			V					23
2	200.14	100.57	183.11	92.06	<b>228.13</b>	114.57	<b>211.11</b>	106.06	Q	2238.15	1119.58	2221.12	<b>1111.07</b>	22
3	301.19	151.10	284.16	142.58	<b>329.18</b>	165.10	312.16	156.58	T	2110.09	<b>1055.55</b>	2093.07	1047.04	21
4	400.26	200.63	383.23	192.12	<b>428.25</b>	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	<b>1909.98</b>	955.49	1892.95	946.98	19
6	514.30	257.65	<b>497.27</b>	249.14	542.29	271.65	525.27	263.14	G	<b>1852.95</b>	926.98	1835.93	918.47	18
7	585.34	293.17	568.31	<b>284.66</b>	<b>613.33</b>	307.17	596.30	<b>298.66</b>	A	1795.93	898.47	1778.91	889.96	17
8	698.42	349.71	681.39	341.20	<b>726.41</b>	363.71	<b>709.39</b>	355.20	I	<b>1724.90</b>	862.95	1707.87	854.44	16
9	827.46	414.24	810.44	405.72	<b>855.46</b>	<b>428.23</b>	838.43	419.72	E	<b>1611.81</b>	806.41	1594.79	797.90	15
10	926.53	463.77	909.50	455.26	<b>954.53</b>	477.77	937.50	469.25	V	<b>1482.77</b>	741.89	1465.74	733.38	14
11	1027.58	514.29	1010.55	505.78	<b>1055.57</b>	528.29	1038.55	519.78	T	<b>1383.70</b>	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	<b>1282.65</b>	641.83	1265.63	633.32	12
13	1199.63	600.32	1182.60	591.80	<b>1227.62</b>	614.31	1210.60	605.80	N	<b>1225.63</b>	<b>613.32</b>	1208.61	604.81	11
14	1346.70	673.85	1329.67	665.34	1374.69	687.85	1357.66	679.34	F	<b>1110.60</b>	555.81	1093.58	<b>547.29</b>	10
15	1433.73	717.37	1416.70	<b>708.85</b>	1461.72	731.37	1444.70	722.85	S	<b>963.54</b>	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	<b>876.50</b>	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	<b>775.46</b>	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	<b>662.37</b>	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	<b>547.35</b>	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	<b>434.26</b>	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	<b>347.23</b>	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	<b>260.20</b>	130.60	243.17	122.09	2

ASMS 2003



Most of the Y ions have been identified.

Mascot Search Results: Peptide View - Microsoft Internet Explorer

Address: [http://www.matrixscience.com/cgi/peptide\\_view.pl?file](http://www.matrixscience.com/cgi/peptide_view.pl?file)

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
23									K	147.11	74.06	130.09	65.55	1

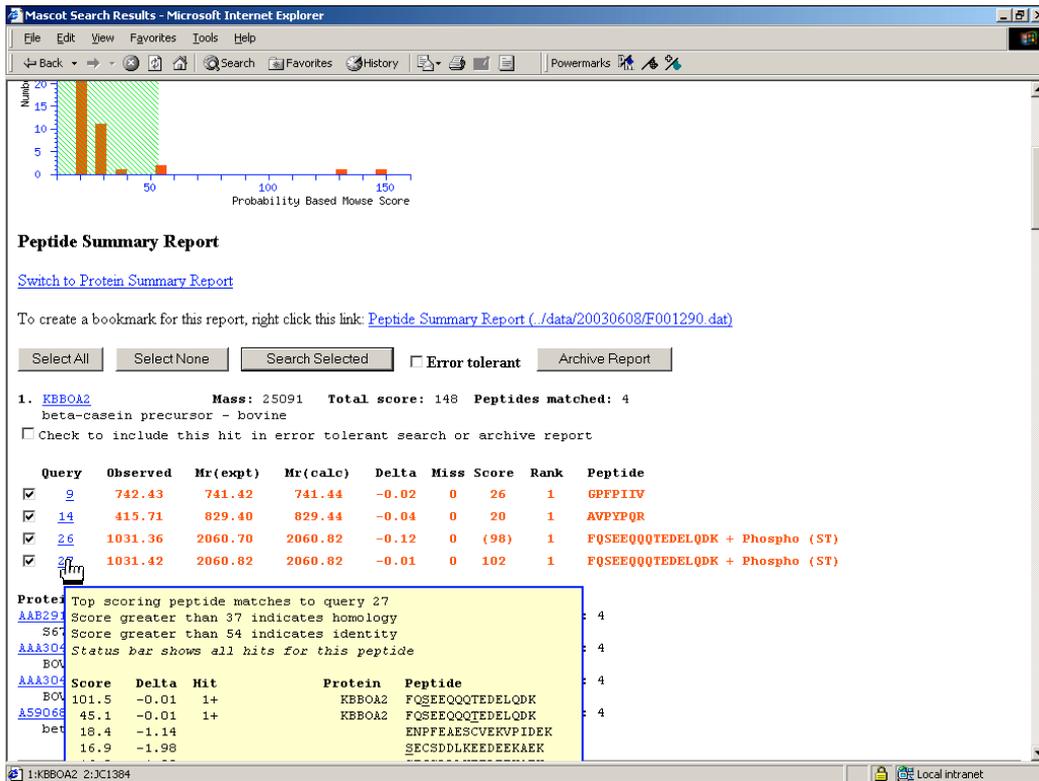
RMS error 413 ppm

NCBI BLAST search of [VQTVGGAIIEVTGNFSTLDLSSLK](#)  
 (Parameters: blastp, nr protein database, expect=20000, no filter, PAM30)  
 Other BLAST [web gateways](#)

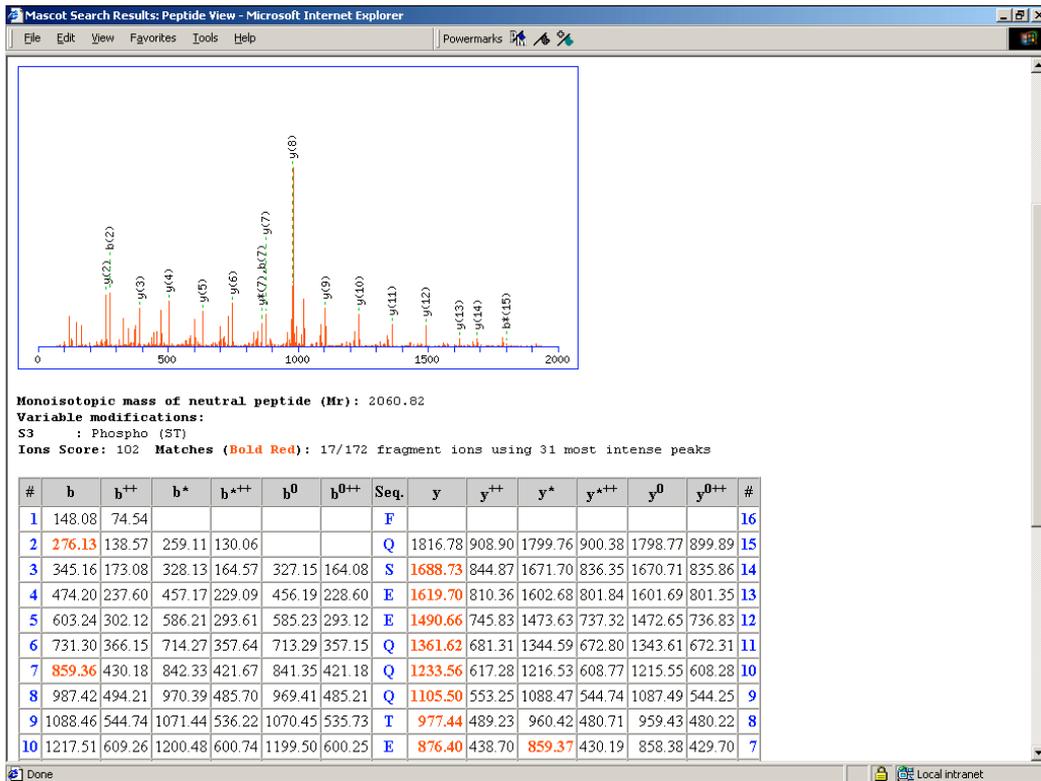
Mascot: <http://www.matrixscience.com/>

ASMS 2003

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. In such cases, there is no doubt about the location of the modification



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

## ***Background***

- **Eukaryotic chromosomes: DNA is associated with histone proteins to form nucleosomes, building blocks of chromatin**
- **PTMs to histones regulate the access to DNA in chromatin**
- **These mods can change the net charge and structure of histones**
- **Histone H3 is modified on the N-terminus (rich in Arg and Lys)**
- **Protease cleavage results in hydrophilic peptides (poorly retained on a RP-HPLC column)**

**ASMS 2003**

***MATRIX***  
***SCIENCE***

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.

## *Background continued...*

- CAD mass spectra of highly multiply charged peptides-interpretation is difficult.
- Peptides + Propionic anhydride converts N-termini and Lysines to propyl amides. This results in a decrease in net charge of the peptides and increased hydrophobicity.
- Now the peptides can be retained on RP-HPLC columns and their CAD spectra are simpler.

**ASMS 2003**

***MATRIX***  
***SCIENCE***

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.

## *Methods*

- Histone protein + Glu-C...isolate 1-50 residue N-terminus by off-line HPLC.
- Second digestion with Chymotrypsin of the 1-50 fragment generates appropriate length peptides for MS/MS.
- Derivatization with propionic anhydride.
- On-line RP-HPLC with direct elution into an electrospray ionization quadrupole ion trap mass spectrometer.
- MS/MS followed by peptide and PTM identification using Mascot.

**ASMS 2003**



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.

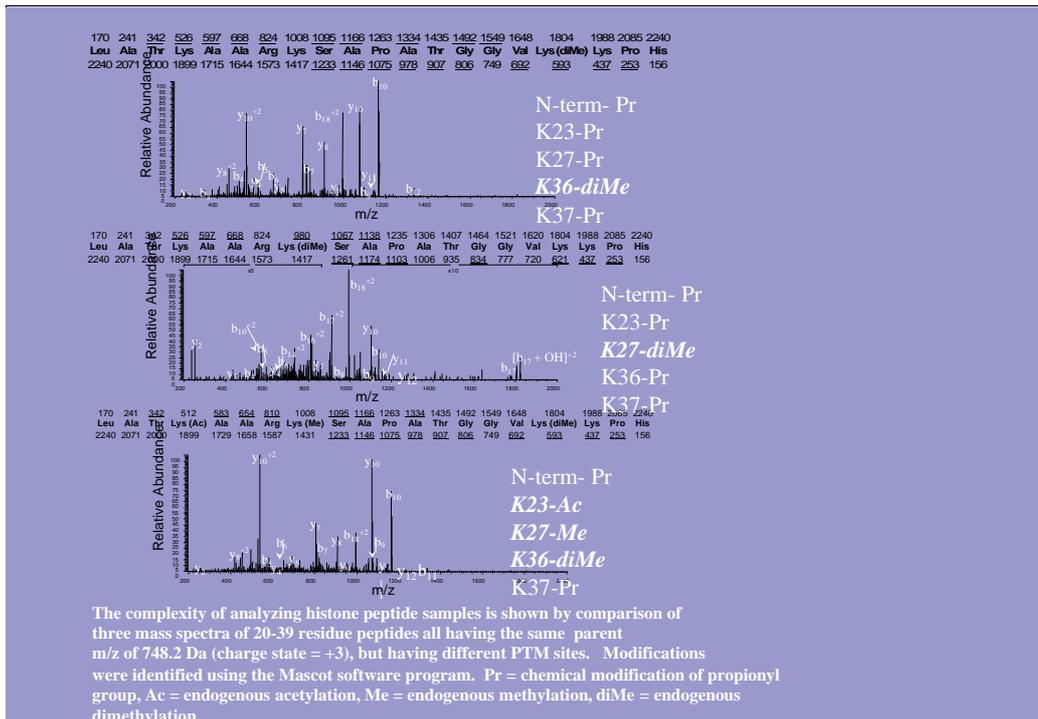
## Results

- Glu-C digestion gave the 1-50 piece
- Chymotrypsin digestion produced 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
- Primary modifications...found on Lys & Ser
- Mods could also be on Arg & Thr
- Ser & Thr can be phosphorylated
- Lys can have mono, di or tri-methyl groups or an acetyl group
- Arg can have mono or di methyl groups
- Propyl amide can be on the N-terminus or on unmodified Lys or Lys with mono-methyl group
- Mascot is one of the few software programs available for searching data with multiple modifications!

ASMS 2003



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.

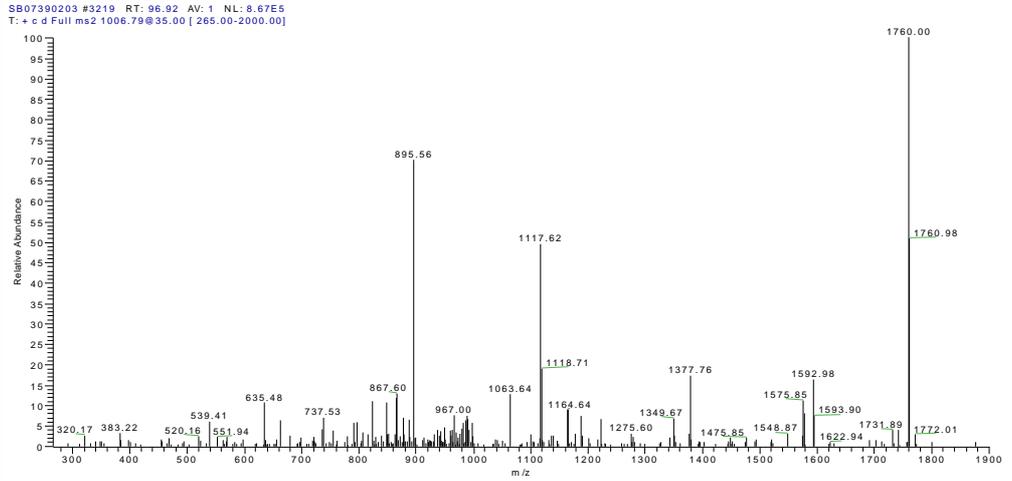


ASMS 2003



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**



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

Address: http://nash.mascot.csi.peptide\_view.phtml?...data/20021014/FO01531\_data?query=3963&h=1&index=g%2f10196611&px=1

### Mascot Search Results

#### Peptide View

MS/MS Fragmentation of **KAARKSAPATGGVCKPHI**  
 Found in **gi|10198611**, dJ34E20.17 (H3 histone family, member B) [Homo sapiens]

Match to Query 3963 (1006.79.24) sb07390203.3219.3219.2.da  
 From data file VDBuser\Users\Ben\Histones\H3 Mitotic Arrested\Fraction 22-23\sb-07390203.RAW

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

Done

Monoisotopic mass of neutral peptide (H): 2011.16  
 Fixed modifications: P (N-term)  
 Variable modifications: K, R, S, T, G, V, H

Ion Score: 60 Matches (Bold Red): 17/164 fragment ions using 33 most intense peaks

#	b	b <sup>++</sup>	b <sup>*</sup>	b <sup>+++</sup>	b <sup>0</sup>	b <sup>0++</sup>	Seq.	y	y <sup>++</sup>	y <sup>*</sup>	y <sup>+++</sup>	y <sup>0</sup>	y <sup>0++</sup>	#
1	241.16	121.08	224.13	112.57			<b>K</b>							17
2	312.19	156.60	295.17	148.09			<b>A</b>	<b>1772.03</b>	886.51	1754.99	878.00	1754.01	877.51	16
3	<b>383.23</b>	192.12	366.20	183.61			<b>A</b>	1700.98	850.99	1683.95	842.48	1682.97	841.99	15
4	<b>539.33</b>	270.17	522.30	261.66			<b>R</b>	1629.94	815.48	1612.92	806.96	1611.93	806.47	14
5	<b>737.47</b>	369.24	720.44	360.72			<b>K</b>	1473.84	<b>737.43</b>	1456.82	728.91	1455.83	728.42	13
6	<b>824.50</b>	412.75	807.47	404.24	806.49	403.75	<b>S</b>	1275.71	638.36	1258.68	629.84	1257.70	629.35	12
7	<b>895.54</b>	448.27	878.51	439.76	877.53	439.27	<b>A</b>	<b>1188.67</b>	594.84	1171.65	586.32	1170.66	585.84	11
8	992.59	496.80	975.56	488.29	974.58	487.79	<b>P</b>	<b>1117.64</b>	559.32	1100.61	550.81	1099.63	550.32	10
9	<b>1063.63</b>	532.32	1046.60	523.80	1045.62	523.31	<b>A</b>	1020.58	510.80	1003.56	502.28	1002.57	501.79	9
10	<b>1164.67</b>	582.84	1147.65	574.33	1146.66	573.84	<b>T</b>	949.55	475.28	932.52	466.76	931.54	466.27	8
11	<b>1231.70</b>	611.35	1204.67	602.84	1203.68	602.35	<b>G</b>	<b>848.50</b>	424.75	831.47	416.24			7
12	1278.72	639.86	1261.69	631.35	1260.71	630.86	<b>G</b>	791.48	396.24	774.45	387.73			6
13	<b>1377.79</b>	689.40	1360.76	680.88	1359.77	680.39	<b>V</b>	734.46	367.73	717.43	359.22			5
14	<b>1575.92</b>	788.47	1558.90	779.95	1557.91	779.46	<b>K</b>	<b>635.39</b>	318.20	618.36	309.68			4
15	<b>1760.04</b>	880.53	1743.02	872.01	1742.03	871.52	<b>K</b>	437.25	219.13	420.22	210.62			3
16	1857.10	929.05	1840.07	920.54	1839.09	920.05	<b>P</b>	253.13	127.07					2
17							<b>H</b>	156.08	78.54					1

Done

ASMS 2003



## ***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<sup>1</sup>, Scott A. Busby<sup>1</sup>, A. Celeste Dunsmoor<sup>1</sup>, Jeffrey Shabanowitz<sup>1</sup>,**

**Cynthia M. Barber<sup>2</sup>, C. David Allis<sup>2</sup>, Donald F. Hunt<sup>1,3</sup>**

**Departments of Chemistry<sup>1</sup> and Pathology<sup>3</sup>,  
University of Virginia, Charlottesville, VA;**

**Department of Biochemistry and Molecular  
Genetics<sup>2</sup>, University of Virginia Health Science  
Center, Charlottesville, VA**

**ASMS 2003**

***MATRIX*  
*SCIENCE***