Chapter 8

Manual Validation of Peptide Sequence and Sites of Tyrosine Phosphorylation from MS/MS Spectra

Amy M. Nichols and Forest M. White

Summary

Mass spectrometry-based analysis of protein phosphorylation has become increasingly powerful over the past decade and has been applied to many different biological systems. One of the most significant concerns facing the phosphoproteomics community and the proteomics field as a whole is the quality and accuracy of the data generated in these large scale efforts. For protein identification in a given sample, the solution has been to require multiple peptides per protein, eliminating "one-hit-wonders" (proteins identified on the basis of a single peptide assignment) which may increase false positives in the data set. Unfortunately, most of the phosphoproteomics data fall into the latter category, as each phosphorylation site will most likely be represented by a single tryptic peptide. Here we provide a detailed protocol describing our manual validation efforts to assure accurate peptide and phosphorylation site assignment for individual MS/MS spectra. In this procedure we use a combination of tools to assign b-, y-, neutral loss, and internal fragment ions, with the goal of assigning all significant ions in the MS/MS spectrum. Confident peptide and phosphorylation site assignment requires good coverage of the peptide with minimal unassigned fragment ions. Using this approach it is possible to maximize the quality of the phosphoproteomics data while minimizing database contamination associated with false positive identifications.

Key words: MS/MS, Mass Spectrometry, Spectra validation, Phosphorylation, Post-translational modification.

1. Introduction

Mass spectrometry has emerged as one of the primary techniques for large scale identification of proteins and post translational modifications such as phosphorylation. Analysis of proteome-wide protein phosphorylation by mass spectrometry can have a substantial impact in advancing our understanding of cell signaling networks if phosphorylation sites are correctly identified while minimizing false positive results. Phosphorylation site identification

M. Lipton and L. Paša-Tolic (eds.), *Mass Spectrometry of Proteins and Peptides, Vol. 492*

DOI: 10.1007/978-1-59745-493-3_8

is complicated because phosphorylation occurs only on specific residues in a protein sequence and typically, only one tryptic peptide will contain the phosphorylation site of interest. Therefore, the challenges in using mass spectrometry for phosphoproteomics are to extract and analyze the few peptides of interest and to definitively validate the sequence of the peptide and the location of the phosphorylation site from a single MS/MS spectrum. To accomplish these goals, phosphopeptide enrichment can be achieved with selective sample purification *(1)* and stringent manual data validation can be used to confidently identify peptide sequence and phosphorylation sites. The goal of manual sequence validation is to assign all peaks in the MS/MS spectra above m/z = 300 as fragments of the peptide sequence. Even when all the MS/MS peaks are assigned, if the fragments are not distributed over the full sequence, then the peptide and/or phosphorylation assignments can remain ambiguous.

2. Materials

2.1. Phosphopeptide Sample Preparation

1. Protocols for sample preparation have been documented in peer reviewed papers from our group *(1–6)* and in addition, a detailed version of the protocol is documented in *Methods in Molecular Biology (7)*.

2.2. Spectra Acquisition

1. Protocols for spectral acquisition have been documented in peer reviewed papers from our group *(1–6)* and in addition, a detailed version of the protocol is documented in *Methods in Molecular Biology (7)*.

2.3. Spectra Validation

1. Searching access on a MASCOT database from Matrix Science Ltd (*see* **Note 1**).
2. Analyst QS Software 1.0 from Applied Biosystems | MDS SCIEX.
3. A list of the monoisotopic masses of each amino acid, **Table 1**.
4. A list of common immonium ions, **Table 2**.

3. Methods

3.1. Phosphopeptide Sample Preparation

As mentioned above, general protocols for sample preparation have been published *(1–6)* and a detailed protocol is documented in *Methods in Molecular Biology (7)*. A brief synopsis of the protocol follows to help the reader understand the type of data being validated.

Table 1
Monoisotopic masses of the amino acids can be combined along with the charge state to calculate the theoretical *m/z* peak for a specific peptide or fragment. Note that leucine and isoleucine are isobaric isomers (same mass, different structures) and cannot be distinguished by mass spectrometry. In addition, the mass differential between amino acids such as glutamine (128.05858) and lysine (128.09496) may be below the resolution of some mass spectrometry instruments. Chemical modification during sample preparation may alter the monoisotopic masses, such as reduction and alkylation of cysteine residues

	Monoisotopic masses	
Amino acid	**Single letter code**	**Monoisotopic mass**
Alanine	A	71.03711
Arginine	R	156.10111
Asparagine	N	114.04293
Aspartate	D	115.02694
Cysteine	C	103.00919
Glutamate	E	129.04259
Glutamine	Q	128.05858
Glycine	G	57.02146
Histidine	H	137.05891
Isoleucine	I	113.08406
Leucine	L	113.08406
Lysine	K	128.09496
Methionine	M	131.04049
Phenylalanine	F	147.06841
Proline	P	97.05276
Serine	S	87.03203
Threonine	T	101.04768
Tryptophan	W	186.07931
Tyrosine	Y	163.06333
Valine	V	99.06841

A cell line of interest is grown to confluence on 10-cm culture plates. The cells are cultured in serum-free culture media 12 h prior to cytokine stimulation. Following the chosen stimulation

Table 2
Immonium ions can be generated during fragmentation of a peptide and result from the loss of CO from the C terminus and the addition of hydrogen to the N terminus. Although immonium ions will indicate the presence of selected amino acids in a peptide, the absence of these ions is insufficient to exclude an amino acid

Common immonium ions

Amino acid	Single letter code	Immonium ion mass
Alanine	A	44.0495
Arginine	R	129.1135
Asparagine	N	87.0553
Aspartate	D	88.0393
Cysteine	C	76.0215
Glutamate	E	102.055
Glutamine	Q	101.0709
Glycine	G	30.0338
Histidine	H	110.0713
Isoleucine	I	86.0964
Leucine	L	86.0964
Lysine	K	101.1073
Methionine	M	104.0528
Phenylalanine	F	120.0808
Proline	P	70.0651
Serine	S	60.0444
Threonine	T	74.06
Tryptophan	W	159.0917
Tyrosine	Y	136.0757
Valine	V	72.0808

conditions, cells are lysed on ice with a solution comprised of 8 M urea and 1 mM sodium orthovanadate to inhibit phosphatase activity. Following reduction (10 mM dithiothreitol at 56°C for 1 h) and alkylation (55 mM iodoacetamide in the dark at room temperature for 1 h), proteins are enzymatically digested for at least 14 h in 100 mM ammonium acetate with 40 μg of trypsin (Promega). Tryptic peptides are desalted in a C18 Sep-Pak Plus

cartridge (Waters) and eluted with 25% acetonitrile, 0.1% acetic acid; the eluted fraction is then lyophilized and stored at −80°C. For immunoprecipitation, lyophilized samples were dissolved in 400 µL of 100 mM Tris, 100 mM NaCl, 1% NP-40 at pH 7.4 (note: following solubilization of the sample, it may be necessary to re-adjust the pH to 7.4 for optimal immunoprecipitation). Samples are incubated overnight with a monoclonal anti-phosphotyrosine antibody (e.g., PT66, Sigma, Product #P5872) to immunoprecipitate the tyrosine phosphorylated peptides. These peptides are eluted with 50 µL of 100 mM glycine pH 2.5 for 30 min at room temperature (*see* **Note 2**).

3.2. Spectral Acquisition

Peptides are loaded into an immobilized metal affinity column and nonspecifically bound peptides are removed with 25% acetonitrile, 1% acetic acid, 100 mM NaCl. Retained peptides are eluted into a reverse-phase C18 pre-column with 50 µL of 250 mM Na_2HPO_4, pH 8.0. The C18 pre-column is attached to a reverse phase analytical column and a high pressure liquid chromatography system. The peptides are slowly eluted with a 140 min gradient as described previously *(1)* and electrosprayed into a quadrupole time-of-flight mass spectrometer (QSTAR XL Pro, Applied Biosystems) for MS and MS/MS spectra acquisition. All the spectra from the sample are saved to a *.wiff file.

3.3. Initial MS/MS Results

1. Convert the data *.wiff file to a MASCOT readable *.msm or *.mgf file using MASCOT.dll in Analyst.
2. Search the data on MASCOT with the MS/MS Ion search form. Select settings relevant to the experimental sample.
3. The MASCOT search algorithm matches the MS/MS data with peptides from specific proteins (*see* **Note 3**) and provides a confidence score for each assignment. Although a high confidence score is generally indicative of a good assignment, manual validation is necessary to verify the peptide and phosphorylation site assignment.
4. To manually validate a peptide, find the MASCOT assignment with the highest confidence score and click on the query link to open the peptide view window, *see* **Fig. 1**.
5. From the peptide view window, record the mass-to-charge ratio (*m/z*) of the precursor ion and its elution time frame displayed at the top of the window, *see* **Fig. 2**.

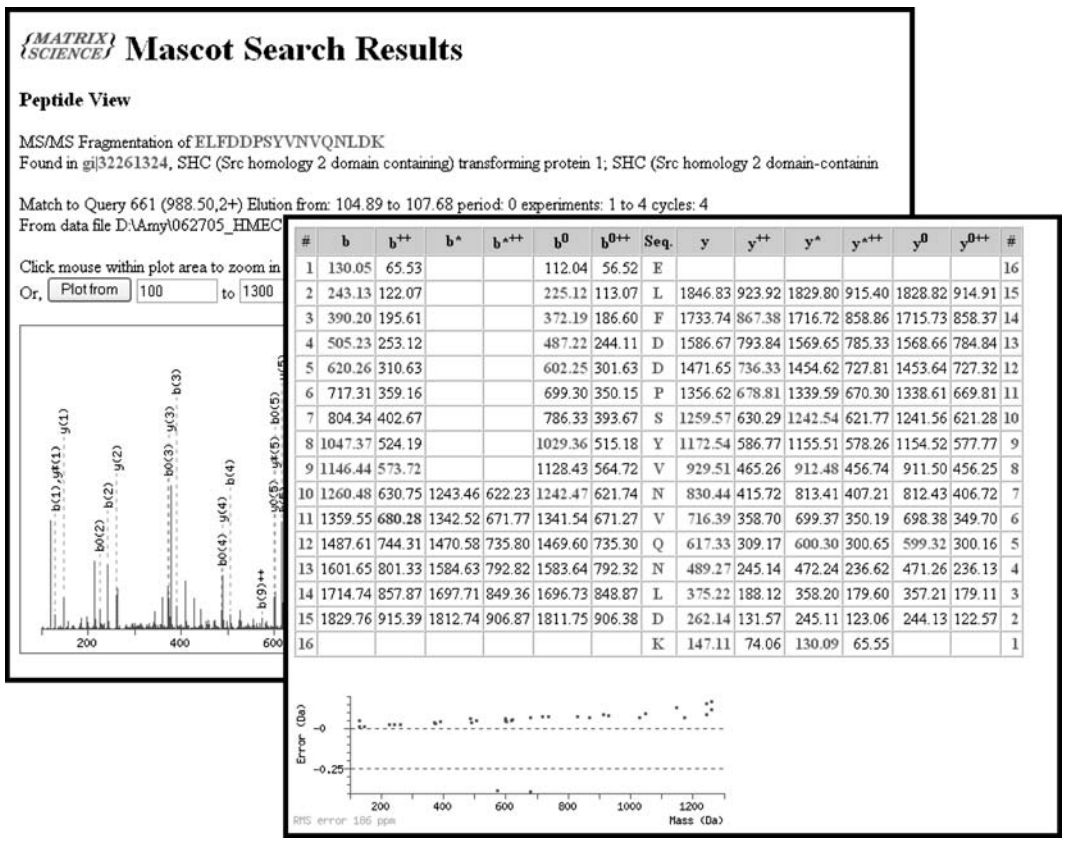
3.4. Identifying and Annotating MS/MS Spectra

1. Open the data *.wiff file in Analyst 1.0.
2. Double click anywhere on the total ion chromatogram (TIC) to display the MS and MS/MS spectra taken at that time.
3. Use the elution time frame and *m/z* for the precursor ion of interest to find the correct full scan mass spectrum (MS scan) and the corresponding MS/MS spectrum from which the assignment was made.

17. gi|32261324 **Mass:** 52105 **Total score:** 85 **Peptides matched:** 5
SHC (Src homology 2 domain containing) transforming protein 1; SHC (Src homology 2 domain-containin
☐ Check to include this hit in error tolerant search or archive report

	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
☑	659	658.99	1973.96	1974.86	-0.90	0	(23)	1	ELFDDPSYVNVQNLDK + pS & pT
☑	660	659.26	1974.75	1974.86	-0.11	0	(52)	1	ELFDDPSYVNVQNLDK + Phospho (Y)
☑	661	988.50	1974.99	1974.86	0.13	0	55	1	ELFDDPSYVNVQNLDK + Phospho (Y)
☑	662	494.76	1975.01	1974.86	0.15	0	(28)	1	ELFDDPSYVNVQNLDK + pS & pT
☑	744	735.06	2202.16	2202.00	0.16	1	35	1	ELFDDPSYVNVQNLDKAR + Phospho (Y)

Fig. 1. MASCOT search result displayed as a function of proteins and the associated peptides found throughout the experimental sample. Each entry following a protein is a summation of the MS/MS spectra of a given precursor ion over a limited elution time; repeated peptides may be listed if the precursor's elution time is sufficiently long. Each query in the list will be given a score indicating the probability of correctly assigning the MS/MS spectra to a given peptide. In this example, five sets of MS/MS spectra were found to correlate to the same peptide (and one missed cleavage) in multiple charge states (+2, +3, +4). The peptides appear to vary in their phosphorylation, but this variation is actually due to incorrect assignment of the site of phosphorylation; upon manual validation these assignments will be clarified.



#	b	b^{++}	b^*	b^{*++}	b^0	b^{0++}	Seq.	y	y^{++}	y^*	y^{*++}	y^0	y^{0++}	#
1	130.05	65.53			112.04	56.52	E							16
2	243.13	122.07			225.12	113.07	L	1846.83	923.92	1829.80	915.40	1828.82	914.91	15
3	390.20	195.61			372.19	186.60	F	1733.74	867.38	1716.72	858.86	1715.73	858.37	14
4	505.23	253.12			487.22	244.11	D	1586.67	793.84	1569.65	785.33	1568.66	784.84	13
5	620.26	310.63			602.25	301.63	D	1471.65	736.33	1454.62	727.81	1453.64	727.32	12
6	717.31	359.16			699.30	350.15	P	1356.62	678.81	1339.59	670.30	1338.61	669.81	11
7	804.34	402.67			786.33	393.67	S	1259.57	630.29	1242.54	621.77	1241.56	621.28	10
8	1047.37	524.19			1029.36	515.18	Y	1172.54	586.77	1155.51	578.26	1154.52	577.77	9
9	1146.44	573.72			1128.43	564.72	V	929.51	465.26	912.48	456.74	911.50	456.25	8
10	1260.48	630.75	1243.46	622.23	1242.47	621.74	N	830.44	415.72	813.41	407.21	812.43	406.72	7
11	1359.55	680.28	1342.52	671.77	1341.54	671.27	V	716.39	358.70	699.37	350.19	698.38	349.70	6
12	1487.61	744.31	1470.58	735.80	1469.60	735.30	Q	617.33	309.17	600.30	300.65	599.32	300.16	5
13	1601.65	801.33	1584.63	792.82	1583.64	792.32	N	489.27	245.14	472.24	236.62	471.26	236.13	4
14	1714.74	857.87	1697.71	849.36	1696.73	848.87	L	375.22	188.12	358.20	179.60	357.21	179.11	3
15	1829.76	915.39	1812.74	906.87	1811.75	906.38	D	262.14	131.57	245.11	123.06	244.13	122.57	2
16							K	147.11	74.06	130.09	65.55			1

Fig. 2. Peptide view window displaying the totaled MS/MS spectra for a given peptide assignment in the MASCOT search results. MASCOT displays the b and y ions it is able to identify. In the example shown, the b and y ions overlap and appear to cover most of the sequence, indicating that this spectrum is appropriate for manual validation. This table of values can be used as an initial reference for manual validation and peak labeling, but the sequencing tool in Analyst 1.0 has an adaptable user interface. Note that the elution time is listed between 104.89 and 107.68; this time window can be used to find individual MS/MS spectra for the precursor ion in the Analyst 1.0 *.wiff file. The error plot shows a general trend away from zero as a function of *m/z*, an error trend commonly found in our data sets. The two points with substantial negative deviation are incorrectly assigned by MASCOT. (*See Color Plates*)

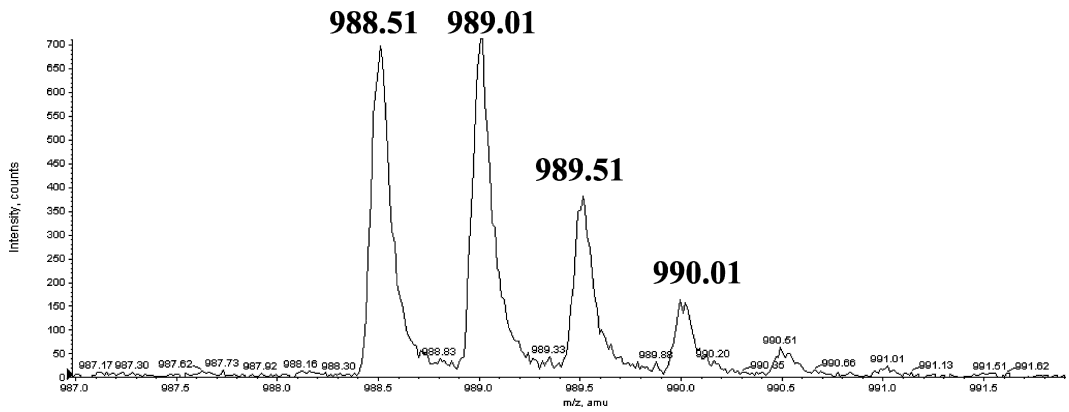


Fig. 3. Precursor ion isolation window in the full scan mass spectrum. For the purpose of this paper, a peptide from Src homology 2 domain-containing-transforming protein C1 (SHC) will be used to demonstrate the manual validation process (*see* **Table 3**). In this case, the precursor ion of interest is 988.51 (+2 charge state) and the area from −1 to 2.5 m/z units around the precursor is free of contamination and the noise around the peak is minimal (below 10% of the precursor peak).

4. The MS/MS spectrum may contain contaminating peptide fragments or species because any peptides or other species within the isolation window (typically −1 to +2.5 *m/z* of the precursor ion (*see* **Note 4**)) will be passed through the first quadrupole and fragmented in the collision cell. To assess the level of possible contamination in the MS/MS spectrum, zoom in on the precursor *m/z* within the MS scan and look for other ions near the precursor ion. A general threshold for contamination is 10% of the parent ion height excluding the peaks in the isotope envelope from the precursor ion, *see* **Fig. 3**.
5. If the full scan mass spectrum indicates potential contamination, return to the MASCOT results and find an alternate MS/MS spectrum for that peptide. Proceeding with a contaminated MS/MS spectrum may be frustrating and inconclusive because fragment peaks generated from the contaminating peak will not be assignable.
6. If the isolation window in the full scan mass spectrum is free of contaminating peaks, select the appropriate MS/MS spectrum and print a full page copy for annotation.
7. At the top of the MS/MS spectrum, write the peptide sequence in the single letter amino acid code assigned by MASCOT in the standard N- to C-terminal direction.
8. Next, label the *m/z* ratio and the charge state for all peaks above the background. The charge state can be found by zooming in on the peak in Analyst and evaluating the spacing between the isotope peaks. (Ex. A spacing of 1 *m/z* unit is a +1, 0.5 *m/z* unit is a +2, 0.33 *m/z* is a +3, etc.). Also, verify that the labeled *m/z* value printed by Analyst on the spectra corresponds to the monoisotopic peak (typically the first peak in the isotope envelope), **Fig. 4**.

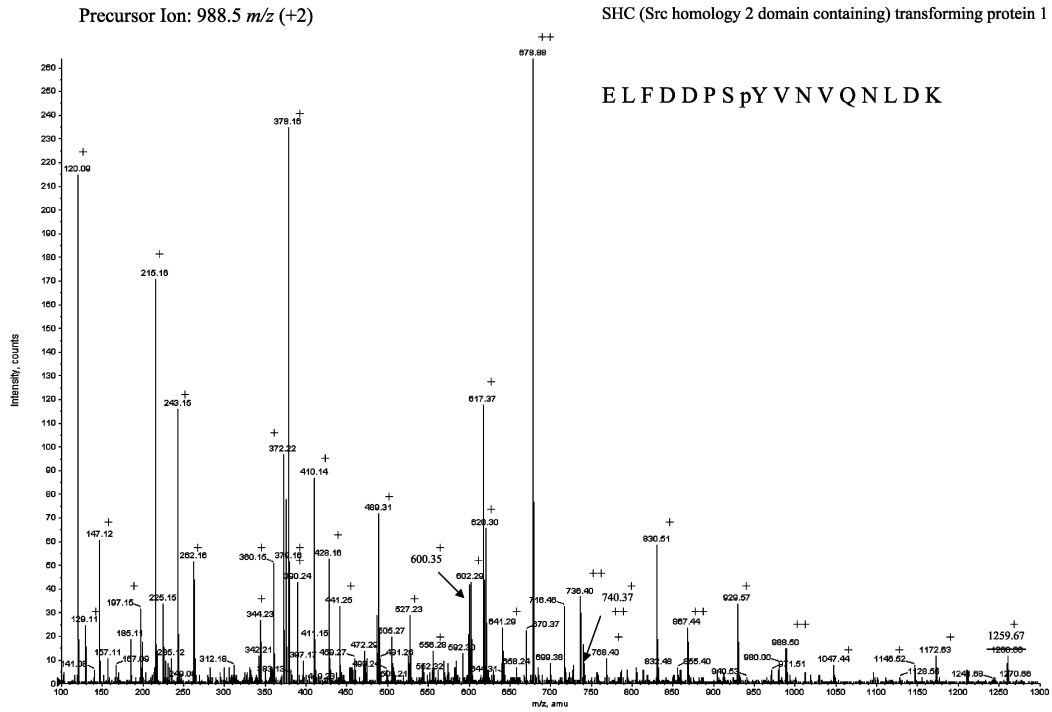


Fig. 4. Annotating charge state and *m/z* for fragment ions in the MS/MS spectrum. The basic annotation for an MS/MS scan includes the charge state of each peak, the protein it is assigned to by MASCOT, the peptide sequence matched by MASCOT and the parent ion *m/z* and charge state. The charge state for each peak in the MS/MS scan can be found by zooming in on the peak in Analyst and evaluating the spacing in the isotope envelope (+1 = 1 *m/z*; +2 = 0.5 *m/z*; +3 = 0.33 *m/z;* +4 = 0.25 *m/z*; etc.).

3.5. Peptide Sequencing

1. De novo peptide sequencing can be performed on each spectrum, but it can be time consuming even with experience. To facilitate fragment ion identification, select the MS/MS spectrum of interest in Analyst 1.0 and select sequence peptide from the BioExplore menu. A new window will appear with a variety of settings that can be specified for the experimental conditions, **Fig. 4**.
2. Once you have selected the settings, click on the sequence button. Analyst will propose a set of sequences with scores reflecting how well each sequence matches the spectrum (*see* **Note 5**).
3. If the sequence proposed by MASCOT is not listed in the Analyst window, type the sequence into the fragment window, **Fig. 5**. Make sure to follow the nomenclature in Analyst for modified residues, such as Z for phosphotyrosine (*see* **Note 6**).
4. Once the sequence is entered, the fragment ions identified by Analyst are highlighted in the table. First, use the table to identify the +1 charge state b and y ions, generated by

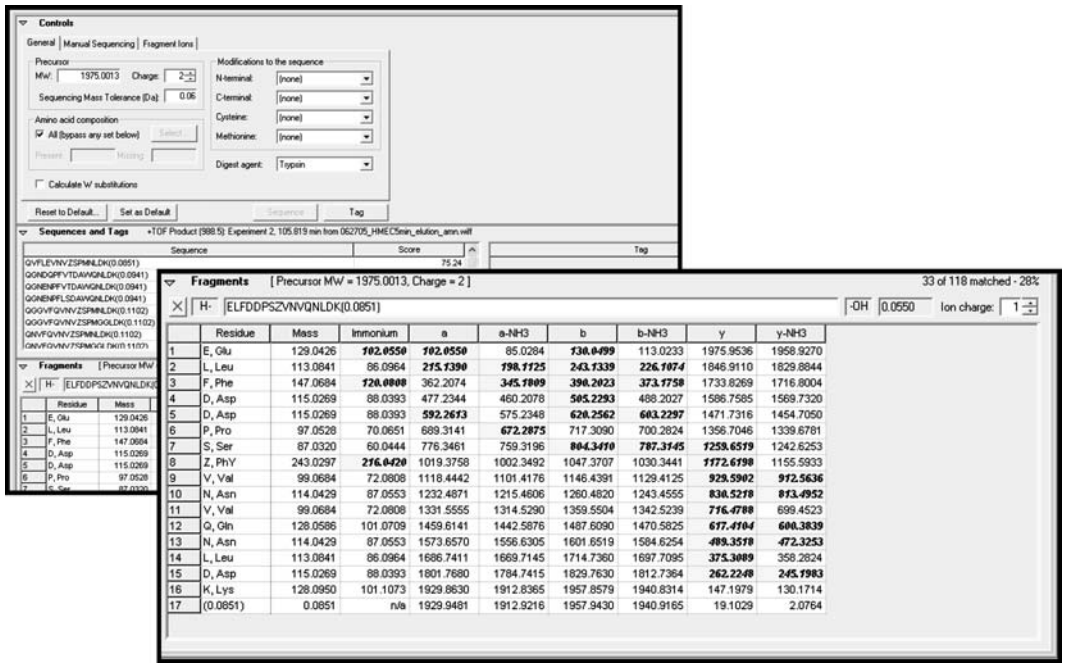


Fig. 5. Sequencing peptides with Analyst BioExplore. Under the BioExplore menu in Analyst 1.0, there is a function called sequence peptide which will bring up the window shown in the figure. The function will sequence the MS/MS spectrum selected from the *.wiff file and propose a set of possible peptide sequences. The sequences proposed by Analyst may not be the same as those proposed by MASCOT. The differential can be attributed to the databases and algorithms used by the two programs to match the MS/MS scans with the peptides. In the sequence peptide function, the fragment window is malleable, enabling other potential sequences to be entered, including the assignment provided by MASCOT. Fragment ions detected in the MS/MS scan will be highlighted and shown in *bold* in the fragment window. This function in Analyst is particularly helpful for calculating b ions, y ions and internal fragments. The charge state can also be changed in the upper right hand corner of the fragment window to calculate the *m/z* value for fragment ions with higher charge states.

fragmentation at the peptide bond and charge retention on the N or C terminus, respectively. If Analyst is not available, the theoretical *m/z* value for each b ion can be calculated by totaling the monoisotopic masses of each amino acid in a given N terminal fragment, adding the mass of the proper number of protons and dividing by the charge. The theoretical *m/z* value for each y ion can also be calculated by totaling the monoisotopic masses of the appropriate residues and adding the atomic mass of water, the proper number of protons and dividing by the number of charges. Water is added to the C terminus of the peptide during the peptide bond cleavage by trypsin or some other protease of choice.

5. For each b ion, draw a line over the fragment's C-terminal amino acid to denote the ion's presence. Also, label the fragment ion directly on the spectra with the notation $b_{\#}$, where the # is the number of residues in the fragment ion, **Fig. 6a**.

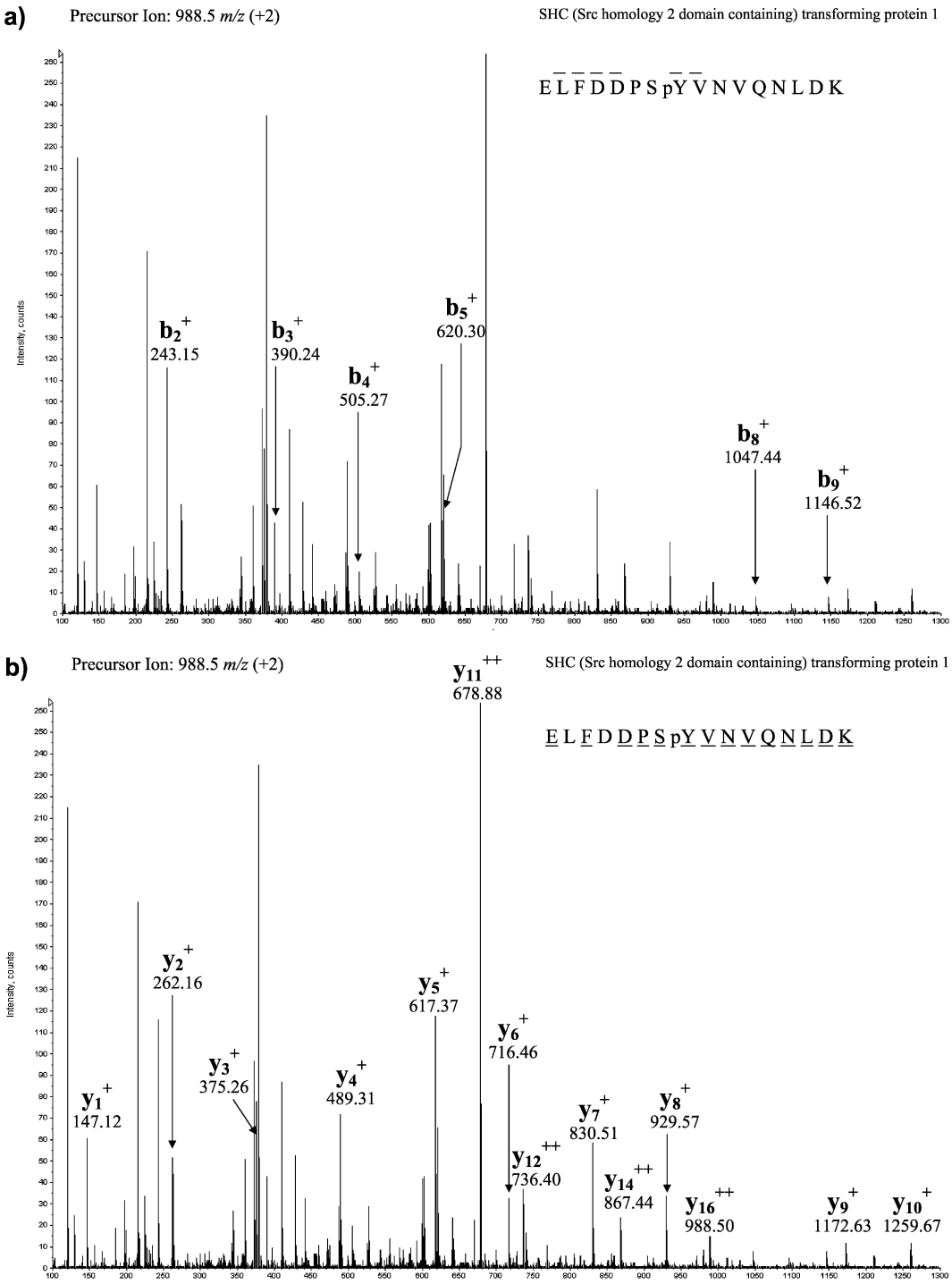


Fig. 6. Assignment of b- and y-type fragment ions in the MS/MS spectrum. Fragmentation of the amide bond results in b ions (**a**) and y ions (**b**), resulting from charge retention at the N or C terminus, respectively. Each b ion identified in the spectra is labeled and a *line* is drawn over the top of the amino acid C-terminal to the peptide fragmentation event. The same procedure is followed for the y ion annotation, but the corresponding amino acid N-terminal to the fragmentation event is underlined on the peptide sequence. The numbering system for b and y ions refers to the total number of residues in the fragment. Initial b and y ion identification is done using the table of values provided by MASCOT or Analyst.

6. For each y ion, draw a line under the fragment's N-terminal amino acid to denote the ion's presence. Also, label the fragment ion directly on the spectra with the notation $y_{\#}$, where the # is the number of residues in the fragment ion, **Fig. 6b**.
7. If the peptide is long enough, all the b and y ions may not appear in the spectrum. To sequence the residues not covered by the +1 charge state b and y ions, identify any +2 b and y ions. Analyst will display the *m/z* for the +2 y and b ions if the ion charge in the upper right hand corner of the fragment ion window is adjusted to +2. If the peptide is of sufficient length, +3 b and y ions may also be found in the spectra.
8. Once all the b and y ions are identified and labeled, the remaining peaks must also be assigned, because unidentified peaks will significantly diminish confidence in the peptide assignment. With most MS/MS spectra, additional fragment ions will help to verify the identity of the peptide. These ions can result from neutral loss of water (H_2O ~ 18 Da), ammonia (NH_3 ~ 17 Da), carbonyl (CO ~ 28 Da (note that b ions displaying loss of CO are commonly referred to as 'a ions')), or phosphate (H_3PO_4 ~ 98 Da or HPO_3 ~ 80 Da) groups, **Fig. 7a, b**. These fragment ions will be separated from the appropriate b or y ion by the indicated number of atomic mass units divided by the charge state of the respective b or y ion.
9. Additional fragment ions may result from internal fragmentation, which is especially prevalent if there is a proline in the sequence. Proline residues normally fragment N-terminal to the residue, producing fragments in either direction, **Fig. 7c**.
10. Identify all the major peaks in the spectra above 300 *m/z* units. Peaks found below 300 *m/z* tend to be more difficult to identify and may contain a large number of immonium ions and internal fragments. However, if one of the top five most abundant peaks in the spectrum is below 300 *m/z*, it should be identified. Due to conservation of mass, all peaks in a given spectrum should be identifiable unless there is contamination in the MS scan. The more peaks that can be identified and assigned in a given spectra, the higher the confidence will be in the peptide identification.
11. If there still remain abundant unassigned peaks in the spectrum after assigning b and y ions, neutral losses and internal fragments, then it is likely that the peptide sequence assignment from Mascot is wrong or there is undefined contamination. To include the spectrum in the data set, attempt to find the actual sequence of the peptide which matches the spectrum, otherwise this peptide should be removed from the data set.

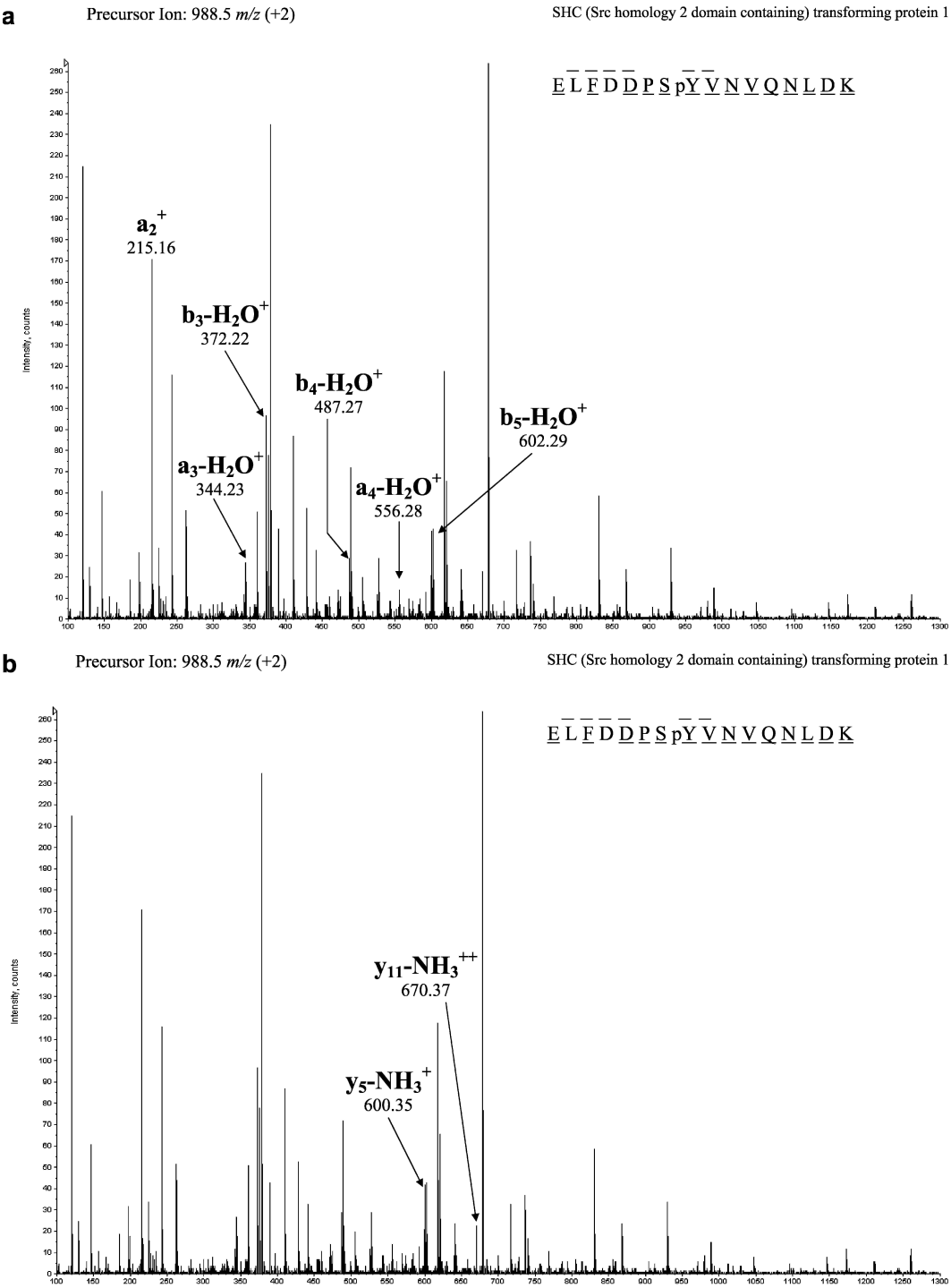


Fig. 7. Assignment of other fragment ions in the MS/MS spectrum. Additional fragment ions in the MS/MS spectrum are the result of neutral loss from b (**a**) or y (**b**) ions, or from internal fragments (**c**). Note that internal fragments can be much harder to identify, but proline cleavage is particularly facile. Internal fragments in either the N- or C-terminal direction often result from an initial cleavage N-terminal to proline and a second cleavage event elsewhere in the peptide sequence.

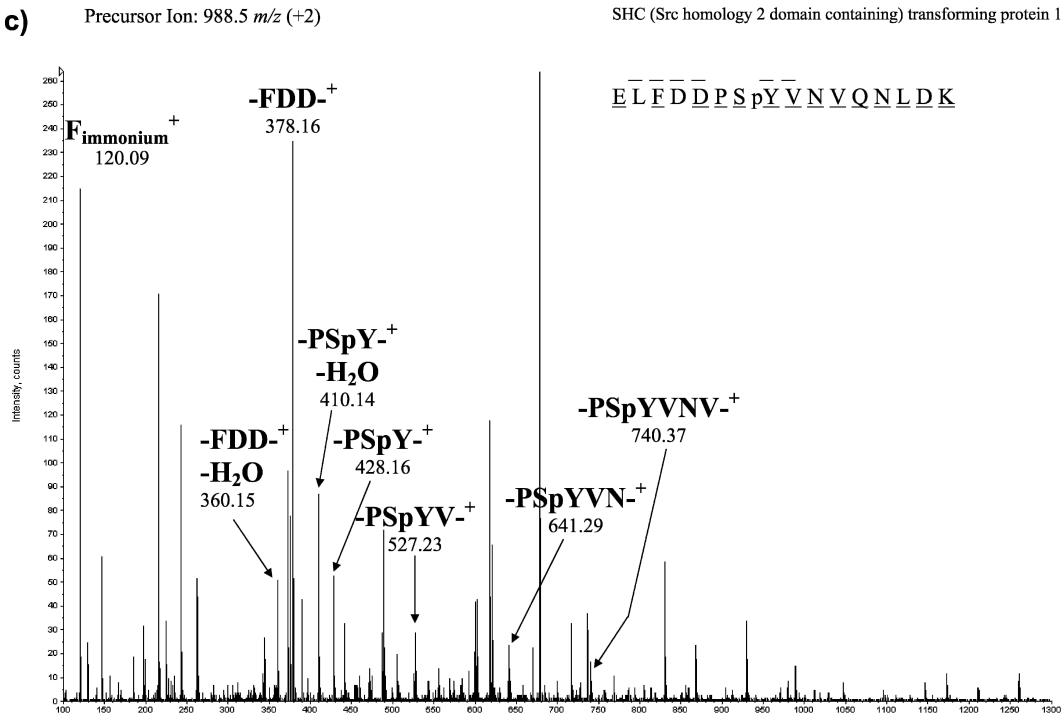


Fig. 7. (continued)

3.6. Phosphorylation Site Identification

1. Tyrosine phosphorylation (pTyr) can be identified by two pieces of evidence: the presence of the phosphotyrosine immonium ion at $m/z = 216.0426$ (+1 charge state) *(8)* and the increased mass to charge of the tyrosine residue in the sequencing from 163.0633 to 243.0297 m/z (+1 charge state) (*see* **Note 7**).
2. Phosphorylation of serine and threonine residues also increases the residue mass by 79.97 Da. However, serine and threonine phosphorylation is often associated with neutral loss of phosphoric acid from both the precursor ion and from the fragment ions (b, y and internal fragments) containing the phosphorylated residue (*see* **Note 8**).
3. To identify sites of serine and threonine phosphorylation it is often necessary to look for fragment ions corresponding to the addition of phosphate followed by neutral loss of phosphoric acid. In these instances, final nominal residue masses of 69 Da for phosphorylated serine and 83 Da for phosphorylated threonine are detected. Unfortunately, these masses can also result from neutral loss of water from non-phosphorylated serine and threonine, which can confound phosphorylation site assignment. Fortunately, neutral loss of phosphoric acid from phosphorylated serine and threonine

residues is typically much more abundant that seen for neutral loss of water from non-phosphorylated serine and threonine residues.

4. Pinpointing the exact phosphorylated residue can be difficult if the sequence of b and y ions does not cover the potential site of phosphorylation and multiple tyrosine, serine, or threonine residues are in the unsequenced portion of the peptide.
5. Phosphorylation sites should be reported on a specific residue only if phosphorylation of that residue is explicitly validated from the spectra. Otherwise, it should be explicitly stated that phosphorylation could be on any of several residues in the unsequenced portion of the peptide (**Table 3**).

Table 3
MS/MS peak list for fragmentation of parent ion 988.51 (+2). All peaks for which charge state can be ascertained are listed and annotated with amino acid sequence and appropriate nomenclature. Unassigned peaks are listed as such and their corresponding intensity is also listed. Other than *m/z* = 441.25, the unassigned peaks either fall below 300 *m/z* or are less abundant than 10% of the maximum spectra intensity of 260. Ideally it should be possible to also assign *m/z* = 441.25; however in practice it is often difficult to account for every ion in the MS/MS spectrum. In the case of a single unidentified ion (at ~15% of the most abundant ion in the spectrum), the confidence in the peptide and phosphorylation site assignment can be very high. The *final two columns* in the table show the theoretical mass of the assigned fragment and the error from the experimental value

Parental ion *m/z:* 988.50 (+2)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
120.09	1	Phenylalanine immonium ion		120.08	0.01
129.11	1	Unassigned			
147.12	1	K (y_1)		147.2	−0.08
185.11	1	Unassigned	20		
197.15	1	Unassigned	35		
215.16	1	EL (a_2)		215.14	0.02
225.15	1	Unassigned	34		
243.15	1	EL (b_2)		243.13	0.02
262.16	1	DK (y_2)		262.22	−0.06
342.21	1	Unassigned	12		

(continued)

Table 3 (continued)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
344.23	1	ELF (a_3-H_2O)		344.2	0.03
360.15	1	–FDD–internal fragment, minus NH_3		360.2	–0.05
372.22	1	ELF (b_3-H_2O)		372.19	0.03
375.26	1	LDK (y_3)		375.31	–0.05
378.16	1	–FDD–internal fragment		378.21	–0.05
390.24	1	ELF (b_3)		390.2	0.04
410.14	1	–PSpY–internal fragment, minus H_2O		410.11	0.03
428.16	1	–PSpY–internal fragment		428.12	0.04
441.25	1	Unassigned	34		
487.27	1	ELFD (b_4-H_2O)		487.22	0.05
489.31	1	NLDK (y_4)		489.35	–0.04
505.27	1	ELFD (b_4)		505.23	0.04
527.23	1	–PSpYV–internal fragment		527.19	0.04
556.28	1	ELFDD (a5–2(H_2O))		556.25	0.03
600.35	1	QNLDK (y_5-NH_3)		600.39	–0.04
602.29	1	ELFDD (b_5-H_2O)		602.25	0.04
617.37	1	QNLDK (y_5)		617.41	–0.04
620.3	1	ELFDD (b_5)		620.26	0.04
641.29	1	–PSpYVN–internal fragment		641.23	0.06
670.37	2	PSpYVNVQNLDK (y_{11}-NH_3)		670.34	0.03
678.88	2	PSpYVNVQNLDK (y_{11})		678.86	0.02
716.46	1	VQNLDK (y_6)		716.48	–0.02
736.4	1	DPSpYVNVQNLDK (y_{12})		736.37	0.03
740.36	1	–PSpYVNV–internal fragment		740.3	0.06
768.4	1	Unassigned	11		
830.51	1	NVQNLDK (y_7)		830.52	–0.01
867.44	2	FDDPSpYVNVQNLDK (y_{14})		867.42	0.02
929.57	1	VNVQNLDK (y_8)		929.59	–0.02
988.5	2	ELFDDPSpYVNVQNLDK (y_{16}), parent ion		988.48	0.02

(continued)

Table 3 (continued)

Experimental *m/z*	Charge state	Fragment	Peak height	Theoretical *m/z*	Delta
1,047.44	1	ELFDDPSpY (b_8)		1,047.37	0.07
1,146.52	1	ELFDDPSpYV (b_9)		1,146.44	0.08
1,172.63	1	pYVNVQNLDK (y_9)		1,172.62	0.01
1,210.61	1	Unassigned	6		
1,259.67	1	SpYVNVQNLDK (y_{10})		1,259.65	0.02

4. Notes

1. Alternative searching programs exist and can be employed for matching MS/MS spectra and peptide sequences. Each program will have its own caveats and biases. Other programs include but are not limited to SEQUEST, PROQUANT, ProteinPilot, SpectrumMill, and others.
2. Our protocol *(7)* specifically isolates phosphotyrosine containing peptides. This method will rarely identify peptides containing phosphoserine and phosphothreonine. To analyze the latter, a general phosphorylation enrichment strategy should be employed, such as IMAC, strong cation exchange *(9)*, or titanium dioxide (TiO_2) *(10, 11)*.
3. The MASCOT results from a search of your data set will produce peptide sequence assignments and associated protein assignments. It is important to realize that multiple proteins may have the same peptide sequence. For a comprehensive list of proteins containing a given peptide sequence, perform a BLAST search through NCBI (http://www.ncbi.nlm.nih.gov/BLAST).
4. The isolation window taken around a precursor ion for fragmentation will depend on the resolution and settings of the mass spectrometer being employed.
5. The sequence assigned by Analyst 1.0 may not be the same as the sequence assigned by MASCOT. Differential sequence assignment between Analyst and MASCOT reflects the different biases in how the databases and algorithms are set up. Disagreement between the two programs exemplifies the need for manual spectral validation prior to presenting mass spectrometry data.

6. Nomenclature for altered residues can be found and modified in the Data Dictionary in the BioTools menu of Analyst 1.0.
7. The intensity and presence of phosphotyrosine immonium ion (m/z = 216.0462, +1 charge state) is dependent on the sequence of the peptide and the fragmentation energy of the instrument. The absence of the immonium ion peaks is not sufficient to exclude phosphotyrosine residues from a peptide sequence.
8. As with tyrosine phosphorylation, the peptide sequence plays a significant role in determining the extent of neutral loss of phosphoric acid from serine or threonine phosphorylated precursor or fragment ions. Multiply phosphorylated peptides and peptides with basic residues (R, K, H) tend to exhibit intense neutral loss peaks.

Acknowledgements

The authors thank Paul Huang, Alejandro Wolf-Yadlin, and other members of the White lab, for helpful discussions. This work was supported in part by a National Science Foundation Graduate Research fellowship, NIAID grant R21-AI065354 and by NIH grant P50-GM68762.

References

1. Zhang, Y, Wolf-Yadlin, A, Pappin, DJ, Rush, J, Lauffenburger, DA, White, FM. Time resolved mass spectrometry of tyrosine phosphorylation sites in the EGF receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* 4, 1240–1250 (2005).
2. Schmelzle, K, Kane, S, Gridley, S, Lienhard, GE, White, FM. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* 55(8), 2171–2179 (2006).
3. Schmelzle, K, White, FM. Phosphoproteomic approaches to elucidate cellular signaling networks. *Curr. Opin. Biotechnol.* 17(4), 406–414 (2006).
4. Kim, JE, White, FM. Quantitative analysis of phosphotyrosine signaling networks triggered by CD3 and CD28 costimulation in Jurkat cells. *J. Immunol.* 176(5), 2833–2843 (2006).
5. Moser, K, White, FM. Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. *J. Proteome Res.* 5(1), 98–104 (2006).
6. Kim, JE, Tannenbaum, SR, White, FM. Global phosphoproteome of HT-29 human colon adenocarcinoma cells. *J. Proteome Res.* 4, 1339–1326 (2005).
7. Zhang, Y, Wolf-Yadlin, A, White, FM. Quantitative proteomics analysis of phosphotyrosine-mediated cellular signaling networks. *Methods Mol. Biol.* 359, 203–212 (2007).
8. Steen, H, Küster, B, Fernandez, M, Pandey, A, Mann, M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal. Chem.* 73(7), 1440–1448 (2001).
9. Beausoleil, SA, Jedrychomski, M, Schwartz, D, Elias, JE, Villén, J, Li, J, Cohn, MA, Cantley, LC, Gygi, SP. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc. Natl Acad. Sci. U. S. A.* 101(33), 12130–12135 (2004).

10. Pinkse, MWH, Uitto, PM, Hilhorst, MJ, Ooms, B, Heck, AJR. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal. Chem*. 76(14), 3935–3943 (2004).

11. Larsen, MR, Thingholm, TE, Jensen, ON, Roepstorff, P, Jørgensen, TJD. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol. Cell. Proteomics* 4(7), 873–886 (2005).