Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 234 (2004) 185–202
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Effect of ionization mode in the analysis of proteolytic protein digests

Paul M. Bigwarfe, Jr.[a,1], Troy D. Wood[a,b,*]

[a] *Department of Chemistry, Natural Sciences Complex, University at Buffalo, The State University of New York, Buffalo, NY 14260-3000, USA*
[b] *Department of Structural Biology, Natural Sciences Complex, University at Buffalo, The State University of New York, Buffalo, NY 14260-3000, USA*

Received 7 November 2003; accepted 23 February 2004

Dedicated to Professor Alan G. Marshall on the occasion of his 60th birthday.

Abstract

The bottom-up sequencing approach relies on an accurate mass fingerprint of the peptide fragments produced from enzymatic digestion. However, this technique is complicated by "dirty" mass spectra due to matrix interferences and, especially with regard to electrospray ionization data, the presence of multiply-charged species, both of which produce false and incorrect hits in database searches. Here, it is shown that the quality of mass spectra of peptide digests can be improved significantly by changing the polarity of the mass spectrometer's ionization source. Using the negative ion mode in both matrix-assisted laser desorption ionization and electrospray ionization of peptide digests results in the detection of a number of peaks in the mass spectra that are not observed in positive ion mode. In addition, there is considerable reduction in interference from contaminant species in the negative ion mode versus the positive ion mode. By combining the positive ion and negative ion mode data, greater sequence coverage and improved database search scores were obtained. When digestion is incomplete, a dramatic enrichment in the observed abundances for digest product peptides versus the intact proteins could be obtained by changing the pH and the ionization mode. The present results highlight the importance of using positive ion mode and negative ion mode combined to improve overall amino acid sequence coverage and improve confidence in protein identification.

Keywords: Bottom-up; Ionization; Mass fingerprint

1. Introduction

The advent of soft ionization mass spectrometric techniques such as electrospray ionization (ESI) [1] and matrix-assisted laser desorption ionization (MALDI) [2–4] has made the analysis of large biomolecules such as proteins feasible. Henzel et al. showed that by using mass spectrometry alone, one could identify proteins isolated using two-dimensional electrophoresis, ushering in the era of proteomics [5]. In particular, with the ability to detect minute quantities of proteins as well as having low sample requirements, nanoelectrospray ionization (nanospray or nanoESI) [6] and MALDI [7] have enabled routine sequencing of proteins. Presently, proteome analysis is typically performed using the bottom-up sequencing approach [5,8–11]. This method entails separating the proteins in the sample, enzymatically digesting them, followed by mass spectrometric analysis. The protein sequence is then reconstructed from the resulting peptides. This method is dependent on obtaining an accurate mass fingerprint in order to generate a match for protein identity on a sequence database search, such as MASCOT [12]. The quality of the mass spectra (and the ease in their interpretation) is therefore critical. Many database search algorithms do not take into account salt adductions or multiple charging, the latter of which is commonly observed in ESI, even for peptide digests. Also, samples from cell lysates and gels often contain salts and detergents, which are not particularly amenable to MS and give poor mass spectral quality, thus increasing the number of incorrect hits in the database search. Since MADLI mass spectra are generally more tolerant to salts and produce mostly singly charged species, it is the ionization method most commonly employed for mass fingerprinting [13], but some sample cleanup steps are still required [14].

* Corresponding author. Tel.: +1-716-645-6800x2176; fax: +1-716-645-6963.
E-mail address: twood@acsu.buffalo.edu (T.D. Wood).
[1] Present address: Discovery Sciences, aaiPharma Inc., 2320 Scientific Park Dr., Wilmington, NC 28405, USA.

1387-3806/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2004.02.015

A comparison of different MALDI matrices for their utility in detecting protein digest products was investigated [15], but the negative ion mode was not reported.

However, if a protein not in a sequence database is present, sequencing with ESI-tandem mass spectrometry (MS/MS) becomes necessary [10]. Additionally, "shot-gun" proteomics places an increased emphasis on nanospray, as MS/MS is needed to identify every protein component in the mixture [16–18]. Using nanospray also allows the ability to routinely couple separations and sample cleanup procedures with ionization [19,20], thereby improving mass spectral quality.

Both MALDI and ESI analysis are usually performed in the positive ion mode. However, the negative ion mode is a surprisingly underutilized tool in the analysis of proteins and protein digests. In particular, nanoESI has been shown to analyze oligonucleotides and peptides of low isoelectric point using both metal [21] and polyaniline [22] coated emitters as well as fused silica capillaries [23] and a wire in a pulled capillary [24]. Fewer molecules produce negative ions [25,26] and less salt adduction is observed in the negative ion mode [27], so cleaner mass spectra result. Data has already been shown that the negative ion mode can produce less interference for peptides than the positive mode [22]. It should be expected then, that by using the negative ion mode, more information and higher quality spectra from protein digests could be obtained. To investigate this possibility, a number of different proteins were digested, placed in both an acidic and a basic solvent, and finally analyzed in the positive ion and negative ion modes, respectively.

2. Experimental

2.1. Apparatus

Most nanospray spectra were acquired on a commercially available PE-SCIEX Biosystems API 3000 (Concord, Ontario) triple quadrupole mass spectrometer equipped with a homebuilt nanospray source [28]. This source contains within it an XYZ translatable emitter platform and mount through which voltage is applied. Data was acquired either with Q_1 scans or in the product ion mode. The quadrupoles were set to unit resolution in all cases. The nebulization gas was not used, as it is incompatible with our homebuilt nanospray source. All instrumental parameters were controlled by the manufacturer's Analyst 1.1 data system on a Windows NT platform. In most cases the mass spectra shown here are the sum of 15 scans.

MALDI spectra were acquired on a commercially available Bruker Daltonics (Billerica, MA) Biflex IV MALDI-Tof mass spectrometer operating in the reflectron mode. External calibration was performed before each use using a mixture of angiotensin II and gramicidin D. Laser fluence was typically 35–40% and 200 shots averaged for each spectrum. Data acquisition was controlled by the manufacturer's FlexControl program and the data processing was performed by Bruker's XTOF software.

2.2. Reagents

All peptide samples, trypsin for enzymatic digestion, and MALDI matrices were obtained from Sigma (St. Louis, MO) and used without further purification. Solvents were purchased from Aldrich (Milwaukee, WI) and were of HPLC grade. NiagaraFlow™ nanospray emitters were provided as a gift from Nanogenesys, Inc. (Amherst, NY) and used to obtain all nanospray spectra. MALDI matrices, α-cyano-4-hydroxy-cinnamic acid (CHCA) and sinapinic acid (SA), and dihydroxybenzoic acid (DHB) were dissolved in 1:1 water:methanol solutions with 2% tri-fluoroacetic acid (CHCA and DHB) or 1:1 water:acetonitrile with 2% ammonium hydroxide (sinapinc acid) until the solution was saturated. Zip Tips™, C_{18} and μC_{18} desalting pipette-tip columns were purchased from Millipore (Bedford, MA) and were used according to the manufacturer's specified procedure.

2.3. Procedure

Approximately 0.5 mg of protein sample was dissolved in 30 μl of 50 mM ammonium bicarbonate pH 8.0 (AmBic)/2 mM dithiothreitol and incubated for 30 min at 45 °C. Then 30 μl of cold 1 mg/ml tosylphenylalanyl chloromethyl ketone (TPCK) treated (to inactivate any chymotrypsin which may be present) trypsin in AmBic was added to the solution and allowed to digest for 18 h at 37 °C. After completely drying under vacuum the sample was then reconstituted in 8–10 μl of 0.5% TFA solution. For nano-ESI analysis, samples were prepared using C_{18} Zip Tips™ and ∼3 μl of sample was eluted into the NiagaraFlow™ emitter with 49:49 water:acetonitrile and either 2% concentrated formic acid (88%) or 0.1 M ammonium hydroxide. Samples for MALDI were desalted with μC_{18} Zip Tips™ designed for smaller elution volumes. Approximately 1 μl was eluted from the Zip Tip™ directly onto the sample plate with the matrix solutions. Mass spectra were acquired in the positive and negative mode for each matrix. Isoelectric points (pI) of identified peptides shown in the tables were obtained theoretically by the SWISS-PROT database [29].

3. Results and discussion

3.1. MALDI analysis

MALDI analysis in the positive ion mode has been the method of choice in analyzing protein digests because of the high throughput and ease of interpretation. However, Fig. 1 shows that different results can be obtained by using the negative ion mode. The mass spectra in Fig. 1 represent extremes between the appearance of the mass spectra in the positive

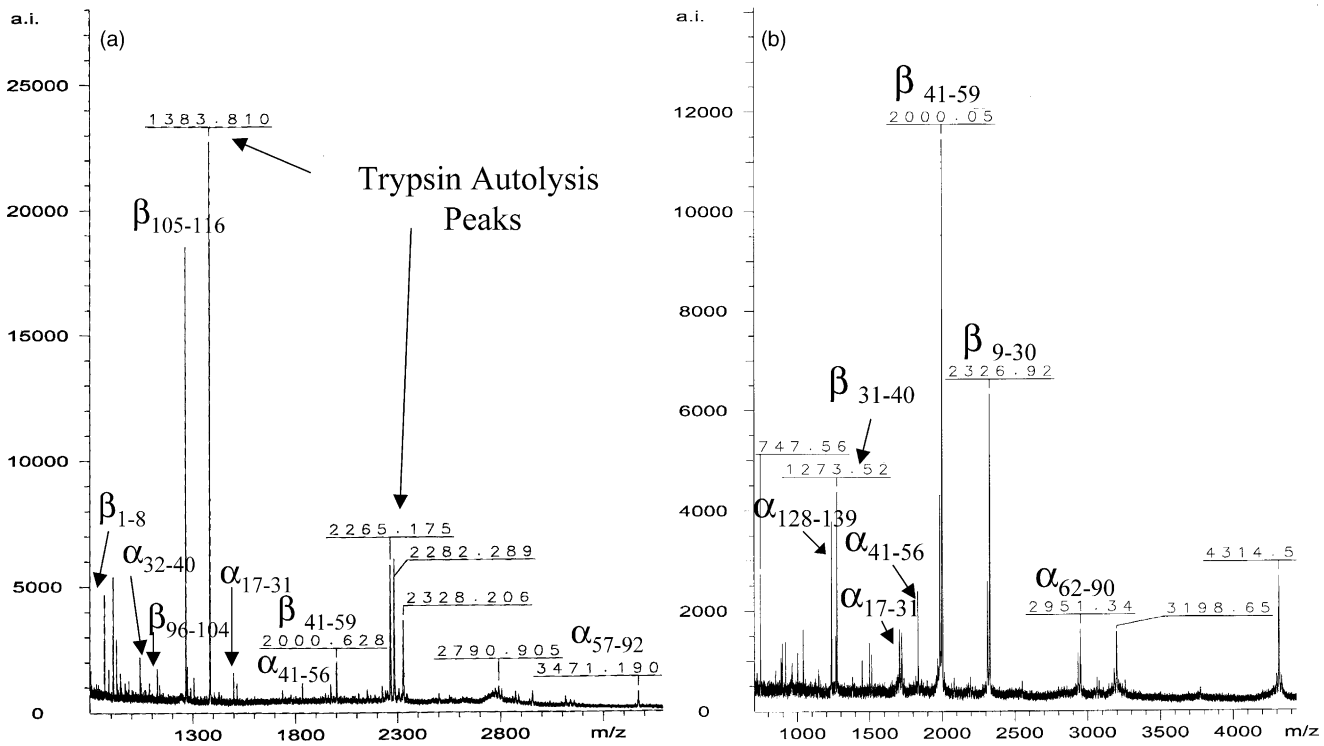

Fig. 1. MALDI spectra of a hemoglobin tryptic digest. (a) In a CHCA matrix in the positive ion mode. (b) In a sinapinic acid matrix using the negative ion mode. See Table 1 for peak identities.

ion mode (with α-hydroxy-cinnamic acid (CHCA) matrix, Fig. 1a) and negative ion mode (with SA matrix, Fig. 1b) of the protein hemoglobin, which is a tetramer composed of two alpha and two beta subunits. A summary comparison of MALDI data for the hemoglobin digest as a function of ionization mode and matrix used is shown in Table 1. Both mass spectra from Fig. 1 exhibit many of the same peaks, but there are some striking differences. Most notable are the different base peaks observed between ionization modes. In the negative mode, the base peak results from the peptide β41–59, which has two acidic and one basic amino acid residues and a theoretical pI of 4.21. The positive mode's most intense peak corresponds to the peptide β105–116, which contains mostly neutral residues and a single Arg (theoretical pI = 9.75). It should be expected, then, that those peptides with high pI values tend to give more prominent signals in the negative mode, while those with low pI values tend to generate intense peaks in the positive mode. Other peaks, for example, that for peptide β9–30 (pI = 4.25, strong negative ion mode peak), follow the same pattern. This is not to say that peptides with low pI values always produce intense negative ion signals while those of higher pI always produce intense positive ion signals. Furthermore, there are some differences noted between the matrixes, in particular that DHB can produce base peaks from acidic peptides in the negative ion modes as well as the positive ion mode, as is observed for β-hemoglobin L96-R104 (pI = 5.32). An interesting cautionary note is also noted from the Table 1 data. The presence of two distinct protein subunits in hemoglobin prevents a good database search score for either subunit without prior separation. The database used, MASCOT [12], uses the Molecular Weight Search (MOWSE) [30] algorithm to score the accuracy of a peptide mass fingerprint hit. The score is expressed as minus 10 times the log of the probability of a match. By only searching the peaks from each individual hemoglobin subunit, scores of 65 and 62 were obtained for α and β, respectively.

Bovine serum albumin (BSA), cytochrome *c*, lysozyme, myoglobin, and β-casein were all examined according to the same procedure. Similar results to hemoglobin were obtained and are shown in Tables 2–6. BSA, the heaviest protein studied at 66 kDa, generated many sodiated peaks, which the database search algorithm does not take into account. A positive match was not returned without compensating for this manually. Cytochrome *c* and myoglobin, proteins with an iron atom metal center in the heme group, generated fewer peaks and the database search matches were not high enough to be considered significant with all the unmatched peaks taken into account. Lysozyme, a poorly soluble protein in ESI solvents, gave more extensive MALDI data in both the positive and negative modes and gave the highest MOWSE scores of any of the proteins in MALDI. Here, there are examples of one peptide generating an intense peak in one ionization mode, while being weak or absent in the other. For example, using SA as the matrix, the peptide corresponding to residues 24–31 (theoretical pI = 5.99) generates an intense peak in the negative ion mode, and is completely absent in the positive mode. In

Table 1
Positive and negative MALDI comparison for digests of α + β hemoglobin

		CHCA Matrix		Sinapinic Acid Matrix		DHB Matrix		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	Positive	Negative	(from SWISS-PROT)
AHGKKVGDALTLAVGHLDDLPG								
ALSNLSDLHAHKLR	57-90 *a*	W						6.43
VGDALTLAVGHLDDLPGALSNLS								
DLHAHK	62-90 *a*			W	I			5.10
AAVLALWDKVNEEEVGGEALGR	9-30 *b*		W	I	S	W		4.25
FFDSFGDLSNPGAVMGNPK	41-59 *b*	W	BP	I	BP	I	S	4.21
KVLHSFGEGVHHLDNLK	66-82 *b*					W	W	7.02
VQLSGEEKAAVLALWDK	1-17 *b*					W		4.68
TYFPHFDLSHGSAQVK	41-56 *a*	W	S	W	I	S	S	6.61
VLHSFGEGVHHLDNLK	67-82 *b*					I	I	6.26
VGGHAGEFGAEALER	17-31 *a*	W	S		I	S	BP	4.75
VVAGVANALAHKYH	133-146 *b*				W			8.58
DFTPELQASYQK	121-132 *b*		I					4.37
VNEEEVGGEALGR	18-30 *b*					S	I	4.09
LLVVYPWTQR	31-40 *b*			BP	S			8.75
LLGNVLVVVLAR	105-116 *b*	BP		W	I	I		9.75
VVAGVANALAHK	133-144 *b*					I		8.73
LHVDPENFR	96-104 *b*	W				BP	I	5.32
LRVDPVNFK	91-99 *a*					S	W	8.75
MFLGFPTTK	32-40 *a*	W		I	I	S	W	8.50
TNVKAAWSK	8-16 *a*				W			10.00
VQLSGEEK	1-8 *b*	S	S			S	I	4.53
VDPVNFK	93-99 *a*					I		5.81
	a = Alpha							
	b = Beta							
	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak						
		All Peaks + or - 1 Charge State Unless Noted Otherwise						
Sequence Coverage	(Alpha)	52%	20%	31%	34%	34%	33%	
	(Beta)	32%	48%	52%	52%	67%	45%	
For Both Positive and Negative		52% Alpha, 63% Beta		34% Alpha, 52% Beta		34% Alpha, 67% Beta		
For All 4 Combined		58% Alpha, 72 % Beta						
MOWSE Score (Matched Peaks Only)	(Alpha)	70	No Match	33	55	70	57	
	(Beta)	61	80	81	81	144	86	
For Both Positive and Negative		109 Beta		55 Alpha, 81 Beta		70 Alpha, 144 Beta		
For All 4 Combined (Alpha + Beta Treated Separately)		106 Alpha, 125 Beta						
For All Combined		80 For Beta				110 For Beta		
Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy								

contrast, the peptide corresponding to residues 115–130 (theoretical pI = 8.75) generates a strong peak in the positive ion mode but it is not observed in the negative ion mode. Thus, by using both positive ion and negative ion modes, a more complete peptide map is obtained, and the MOWSE score increases dramatically. In the case of myoglobin, where only 30% of the sequence is mapped using CHCA in negative ion mode (MOWSE score 53, below the level of statistical significance), nearly complete sequence coverage (94%) is obtained when positive and negative modes are combined using the CHCA and SA matrixes (MOWSE score 226). A special case is noted with β-casein, a multi-phosphorylated peptide that gave the poorest scores in the MALDI analysis. Three peaks were detected in which one or more phosphorylation sites were preserved, corresponding to peptides for residues 16–40 (tetraphosphorylated at residues 30, 32, 33, 34), 41–63 (single phosphorylation site at residue 50) and the truncated 44–63. It should be noted that the β-casein MALDI mass spectra exhibited successive losses of phosphate from the tetra-phosphorylated species, a phenomenon not detected in the nanospray mass spectra of β-casein. This illustrates that MALDI is a slightly harsher ionization method for β-casein than nanospray.

In some cases, the presence of multiple acidic residues in a peptide sequence produces a strong negative ion signal that may not be present in the positive mode. For example, in BSA, a 10 amino acid peptide E300-K309 has three acidic residues and generates a negative ion peak (independent of matrix) but only produces a peak in positive ion mode when DHB is used. Similarly, a tryptic peptide from myoglobin (H119-K133) has two acidic residues and will produce negative ions of this peptide without producing positive ions. Notably, β-casein peptide R16-R40 has seven acidic residues and four phosphates, and produces strong negative ion signal using DHB but nothing in positive ion mode. Though apparently less common, basic peptides may produce a strong positive ion signal and little or no negative ion signal. For example, a peptide from lysozyme (K115-R130) with two basic amino acids produces strong positive ion signal but weak or non-detectable negative ion signal. The C-terminal peptide from β-casein (V185-V224), which contains three basic residues, produces a strong positive ion signal using SA but nothing detectable in negative ion mode. These trends are not a significant difference, but it does show that using CHCA, the most widely accepted matrix for peptide analysis [31–33], may not always be the best choice for MALDI of protein digests. A summary of the MALDI results

Table 2
Positive and negative MALDI comparison for digests of bovine serum albumin

		CHCA Matrix		Sinapinic Acid Matrix		DHB Matrix		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	Positive	Negative	(from SWISS-PROT)
RHPYFYAPELLYYANK	168-183	W		W	W			8.39
DAIPENLPPLTADFAEDK	319-336					W	W	3.77
LGEYGFQNALIVRYTR	421-436	W		W	W			8.59
SLHTLFGDELCKVASLR	89-105	I		I	BP			6.47
LFTFHADICTLPDTEK	529-544					W		4.54
RPCFSALTPDETYVPK	508-523	W				W		6.06
MPCTEDYLSLILNR	469-482	W				W		4.37
KVPQVSTPTLVEVSR	437-451					I	W	8.75
YNGVFQECCQAEDK	184-197						I	4.14
DAFLGSFLYEYSR	347-359	I	S	I	I	W		4.37
VPQVSTPTLVEVSR	438-451					I	W	5.97
LGEYGFQNALIVR	421-433	BP	BP	BP	S	BP	S	6.00
RHPEYAVSVLLR	360-371	I		I	S	I	I	8.75
YICDNQDTISSK	286-297					W	I	4.21
SLHTLFGDELCK	89-100						W	5.30
TCVADESHAGCEK	76-88					W		4.65
HLVDEPQNLIK	402-412					S	I	5.32
HPEYAVSVLLR	361-371	S		S	S	I	W	6.75
ECCDKPLLEK	300-309		S		I	S	S	4.68
LVNELTEFAK	66-75						BP	4.53
KQTALVELLK	548-557					W		8.59
QTALVELLK	549-557	W						6.00
LVVSTQTALA	598-607				S			5.52
NECFLSHK	123-130		S		W			6.74
YLYEIAR	161-167							6.00
AEFVEVTK	249-256						S	4.53
CCAADDK	581-587				I			4.21
AWSVAR	236-241					I		9.79

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	18%	7%	12%	17%	29%	22%
For Both Positive and Negative	22%		17%		35%	
For All 4 Combined	24%					NM = No Match
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	98, 46	36, NM	69, 54	103, 54	132, 84	106, 70
For Both Positive and Negative	115		103		164	
For All 4 Combined	132					

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

Table 3
Positive and negative MALDI comparison for digests of cytochrome *c*

		CHCA Matrix		Sinapinic Acid Matrix		DHB Matrix		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	Positive	Negative	(from SWISS-PROT)
EETLMEYLENPKKYIPGTK	61-79	W						4.95
KIFVQKCAQCHTVEK	8 To 22					I	W	8.86
IFVQKCAQCHTVEK	9 To 22	BP		BP	BP	I	I	8.06
KTGQAPGFTYTDANK	39-53				S			8.50
HKTGPNLHGLFGRK	26-39					W		11.17
GDVEKGKKIFVQK	1 To 13					I		9.53
HKTGPNLHGLFGR	26-38						I	11.00
TEREDLIAYLK	89-99					S	BP	4.68
TGPNLHGLFGRK	28-39						W	11.00
TGPNLHGLFGR	28-38	W	I	I	S			9.44
MIFAGIKK	80-87	W						10.00
MIFAGIK	80-86		S		W	BP	I	8.50
KIFVQK	8 To 13	W					I	10.00
GITWK	56-60		BP					8.75

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	52%	22%	24%	45%	51%	45%
For Both Positive and Negative	53%		45%		56%	
For All 4 Combined	70%					NM = No Match
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	77,44	46, NM	33, NM	66, 43	72, 62	86,62
For Both Positive and Negative	107		66		108	
For All 4 Combined	122					

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

Table 4
Positive and negative MALDI comparison for digests of lysozyme

		CHCA Matrix		Sinapinic Acid Matrix		DHB Matrix		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	Positive	Negative	(from SWISS-PROT)
GYSLGNWVCAAKFESNFNTQATNR	40-63			W				8.20
NLCNIPCSALLSSDITASVNCAK	92-114						W	5.82
KIVSDGNGMNAWVAWR	115-130	I	W	S				8.75
NTDGSTDYGILQINSR	64-79		W	I	I	I	I	4.21
IVSDGNGMNAWVAWR	116-130	S	I	S	W	I	I	5.84
WWCNDGRTPGSR	80-91			S	I			8.25
FESNFNTQATNR	52-63	S	S	S	BP	S	BP	6.00
CKGTDVQAWIR	133-143	S	S	I				8.22
GYSLGNWVCAAK	40-51		S		I	W	W	8.20
GTDVQAWIR	135-143	S	I	S	S	BP	S	5.84
RHGLDNYR	32-39					W		8.75
CELAAAMKR	24-32	BP	BP		S			8.22
WWCNDGR	80-86	S	I	BP	S	S	W	5.83
HGLDNYR	33-39	I	I	I		I	W	6.74
CELAAAMK	24-31		I		S			5.99
KVFGR	19-23	W	W					11.00

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	51%	64%	50%	57%	53%	68%	
For Both Positive and Negative	64%		64%		69%		
For All 4 Combined	66%						
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	109, 108	168, 114	126, 87	134, 91	119, 80	113, 48	
For Both Positive and Negative	168		198		126		
For All 4 Combined	209						

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

Table 5
Positive and negative MALDI comparison for digests of myoglobin

		CHCA Matrix		Sinapinic Acid Matrix		DHB Matrix		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	Positive	Negative	(from SWISS-PROT)
YLEFISDAIIHVLHSKHPGDFGADAQ								
GAMTK	103-133				W			5.74
LFTGHPETLEKFDKFK	32-47	I						6.76
YLEFISDAIIHVLHSK	103-118		S					5.99
GHHEAELKPLAQSHATK	80-96				W			7.03
GLSDGEWQQVLNVWGK	1 To 16			S	S			4.37
LFTGHPETLEKFDK	32-45				W			5.45
VEADIAGHGQEVLIR	17-31	S	S	S	I	I		4.65
HPGDFGADAQGAMTK	119-133		S		W	W		5.21
TEAEMKASEDLKK	51-63	S				W		4.87
HGTVVLTALGGILK	64-77			I				8.76
ALELFRNDIAAK	134-145	I						6.11
LFTGHPETELK	32-42			S	I	I	I	5.40
HLKTEAEMK	48-56	I						6.76
YKELGFQG	146-153				BP	W		6.00
ALELFR	134-139	BP				BP	S	6.05
ASEDLK	57-62	I					BP	4.37

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	38%	30%	36%	66%	44%	15%		
For Both Positive and Negative	58%		75%		45%			
For All 4 Combined	94%						NM = No Match	
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	107, 57	53, 31	62, 51	130, 93	82, 49	40, NM		
For Both Positive and Negative	124		146		95			
For All 4 Combined	226							

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

obtained is given in Fig. 2. This figure shows that increased sequence coverage information can be obtained by running the analysis in both the positive and negative modes.

3.2. Nanospray analysis

The same proteins used in the MALDI analysis were also studied using the positive ion and negative ion modes of nanospray ionization. The first protein studied was BSA, and the results are shown in Fig. 3. In Fig. 3, panels a and b show the peptide digest in an acidic solvent. The negative ion mode spectrum (Fig. 3b) is dramatically cleaner, with approximately an order of magnitude less noise and fewer unidentified peaks. The positive ion spectrum (Fig. 3a) exhibits higher absolute signal intensity as compared to negative-ion mode at the same pH, which should be

Table 6
Positive and negative MALDI comparison for digests of β-casein

Peptide Sequence	Residue Numbers	CHCA Matrix Positive	CHCA Matrix Negative	Sianpinic Acid Matrix Positive	Sianpinic Acid Matrix Negative	DHB Matrix Positive	DHB Matrix Negative	Theoretical pI (from SWISS-PROT)
EMPFPKYPVEPFTESQSLTLTDVENL HLPLPLLQSWMHQPHQPLPPTVMFP PQSVLSLQSK VLPVPQKAVPYPQRDMPIQAFLLYQE	123-184			W				5.32
PVLGPVRGPFPIIV AVPYPQRDMPIQAFLLYQEPVLGPVR GPFPIIV	185-224			S				8.47
RELEELNVPGEIVESLSSSEESITRINK	16-43			BP				6.11
RELEELNVPGEIVES*LS*S*S*EESITR	16-40				W			4.36
INKKIEKFQS*EEQQQTEDELQDK	41-63						BP	2.65
DMPIQAFLLYQEPVLGPVRGPFPIIV	199-224			W			S	3.80
KIEKFQS*EEQQQTEDELQDK	44-63			S				4.37
DMPIQAFLLYQEPVLGPVR	199-217	BP	S		I	S	I	3.75
HKEMPFPK	121-128		I				S	4.37
AVPYPQR	192-198	I				I		8.60
VLPVPQK	185-191	S	I			W	W	8.79
EMPFPK	123-128					W		8.72
GPFPIIV	218-224	I				BP		6.10
								5.52

* = Phosphorylation Site

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	17%	15%	55%	18%	20%	30%		
For Both Positive and Negative	21%		65%		35%			
For All 4 Combined	78%						NM = No Match	
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	48, NM	37, NM	52, NM	NM	52, NM	46, NM		
For Both Positive and Negative	60		71		60			
For All 4 Combined	116							

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

expected since the acidic solvent will promote formation of net positively-charged peptides. However, its *S*/*N* for the base peak is lower (∼25 versus ∼100) than the corresponding negative ion spectrum. Although because the pH of this solution is between 4 and 5, the higher chemical noise background may well be due to dissociated hydronium. The sequence coverage is improved as well; 33% of the residues in BSA are accounted for by the negative ion spectrum, as opposed to 29% for the positive mode. As can be seen from Table 7, there are a number of peaks that are unique to each ionization mode. By combining the data from both modes, the overall sequence coverage is improved to 57%. A direct consequence of this is an improved database search score as is discussed as follows. In the case of the nanospray mass spectra from BSA, a score of 92 was returned for the positive mode and 119 for the negative mode. By including only the peaks in the table (and no interference peaks), these numbers are improved to 180 and 167. The overall score was improved to 219 when data from both modes were combined. It should be noted that all multiply charged peptides, as identified by their isotope pattern, were entered as the molecular mass of the peptide since the database program does not take into account multiple charges.

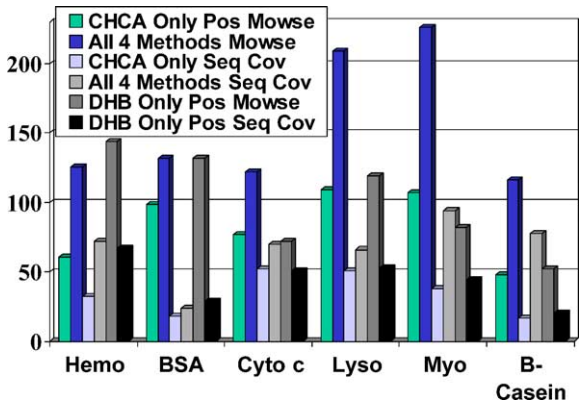

Fig. 2. Summary of the MOWSE score and percent sequence coverage obtained in the MALDI experiments. Each protein exhibits an increase when the negative ion mode is also taken into consideration.

Fig. 3c and d show the positive and negative ion mass spectra of BSA in a basic solvent. The results (also compiled in Table 7) are similar to the acidic solvent mass spectra. The negative ion mass spectrum is once again cleaner than the positive ion mass spectrum. It is also once again less intense, but with a higher *S*/*N* for the base peak. Overall, more peaks are present in these mass spectra, but that is most likely a result of the digestions being carried out on different days and not the elution solvent from the Zip Tip™. Table 7 shows that reduction in score is less for the negative analyses when the unmatched peaks are included, indicating the clarity of the spectra. The sequence coverage and MOWSE scores are improved when the positive ion and negative ion modes are combined as well. The overall sequence coverage increases from 33% in any one analysis to 57%, and the score dramatically improves to 381 when all four analyses are combined. It should also be noticed that peaks are present in the

Table 7
Positive and negative nanospray comparison for digests of bovine serum albumin

		Formic Acid Elution		NH4OH Elution		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	(from SWISS-PROT)
DAIPENLPPLTADFAEDKDVCK	319-340			I (+2)	W (-2)	3.96
ECCHGDLLECADDRADLAK	267-285		I (-2)			4.23
NECFLSHKDDSPDLPK	123-138	W				4.66
LAKEYEATLEECCAK	372-386	I	I			4.49
KVPQVSTPTLVEVSR	437-451	W, BP (+2)	I, I (-2)			8.75
DAFLGSFLYEYSR	347-359	I (+2)	I (-2)			4.37
LKPDPNTLCDEFK	139-151			S (+2)		4.56
VPQVSTPTLVEVSR	438-451			W, I (+2)	W, I (-2)	5.97
QTALVELLKHKPK	549-561			W		9.70
DDPHACYSTVFDK	387-399		W			4.41
LCVLHEKTPVSEK	483-495	S (+2)	S (-2)	I (+2)		6.75
LGEYGFQNALIVR	421-433	S	I	I	S	6.00
VGTRCCTKPESER	456-468	W				8.03
RHPEYAVSVLLR	360-371	S (+2)	I, I (-2)			8.75
LKECCDKPLLEK	298-309	W	W			6.17
TVMENFVAFVDK	569-580			I, I (+2)	I, S (-2)	4.37
EYEATLEECCAK	375-386			W	W	4.09
YICDNQDTISSK	286-297				S (-2)	4.21
SLHTLFGDELCK	89-100	W	W		W	5.30
TCVADESHAGCEK	76-88			W		4.65
HLVDEPQNLIK	402-412	I, S (+2)	I, I (-2)	W, S (+2)	W, I (-2)	5.32
HPEYAVSVLLR	361-371			I, S (+2)	I, I (-2)	6.75
DVCKNYQEAK	337-346		W			6.06
ECCDKPLLEK	300-309				I	4.68
LVNELTEFAK	66-75	I	S	I	I, S (-2)	4.53
AWSVARLSQK	236-245	I				11.00
KQTALVELLK	548-557		I			8.59
CCTESLVNR	499-507			BP		5.99
QTALVELLK	549-557	S	S	I	S	6.00
LVVSTQTALA	598-607	S	BP	I	BP	5.52
NECFLSHK	123-130				I	6.74
EKVLASSAR	210-218			W		8.85
YLYEIAR	161-167	S	S			6.00
AEFVEVTK	249-256	S	S	I	I	4.53
NYQEAK	341-346			S		6.00
SEIAHR	29-34			I		6.47
AWSVAR	236-241			W	I	9.79
KFWGK	156-160			I		10.00
ADEKK	152-156	S			I	6.11
QRLR	219-222	S				10.50
ADLAK	281-285				I	5.88
FGER	229-232	I			S	6.00

Intensities: BP = Base Peak S=Strong I=Intermediate W=Weak
All Peaks + or - 1 Charge State Unless Noted Otherwise

Sequence Coverage	33%	33%	34%	33%
For Both Positive and Negative	40%		38%	
For All 4 Combined	57%			
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)	180, 92	167, 119	190, 83	187, 116
For Both Positive and Negative	219		221	
For All 4 Combined	381			

Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy

ammonium hydroxide elution that are not present in the formic acid experiments, and vice versa. This suggests that it is probably necessary to collect nanospray mass spectra from both ionization modes and in both acidic and basic solutions in order to acquire the greatest extent of sequence coverage.

These experiments were repeated with cytochrome *c*, a mixture of α + β hemoglobin, lysozyme, myoglobin, and β-casein. The results of those experiments are summarized in Tables 8–12. The findings are consistent with the results from the BSA digestions. Sequence coverage for these smaller proteins, when both ionization modes are combined, is extraordinarily high. They gave high MOWSE scores and all scores were improved when positive and negative ion data were combined. Cytochrome *c* (Table 8) and myoglobin (Table 11), proteins that are quite soluble in ESI solvents, gave quality nanospray mass spectra and as a result, high MOWSE scores. Hemoglobin (Table 9) gave excellent sequence coverage but low MOWSE scores for the entire mass spectra since both subunits were present the sample as a mixture as noted above. More tryptic fragments of β-hemoglobin were detected, so that

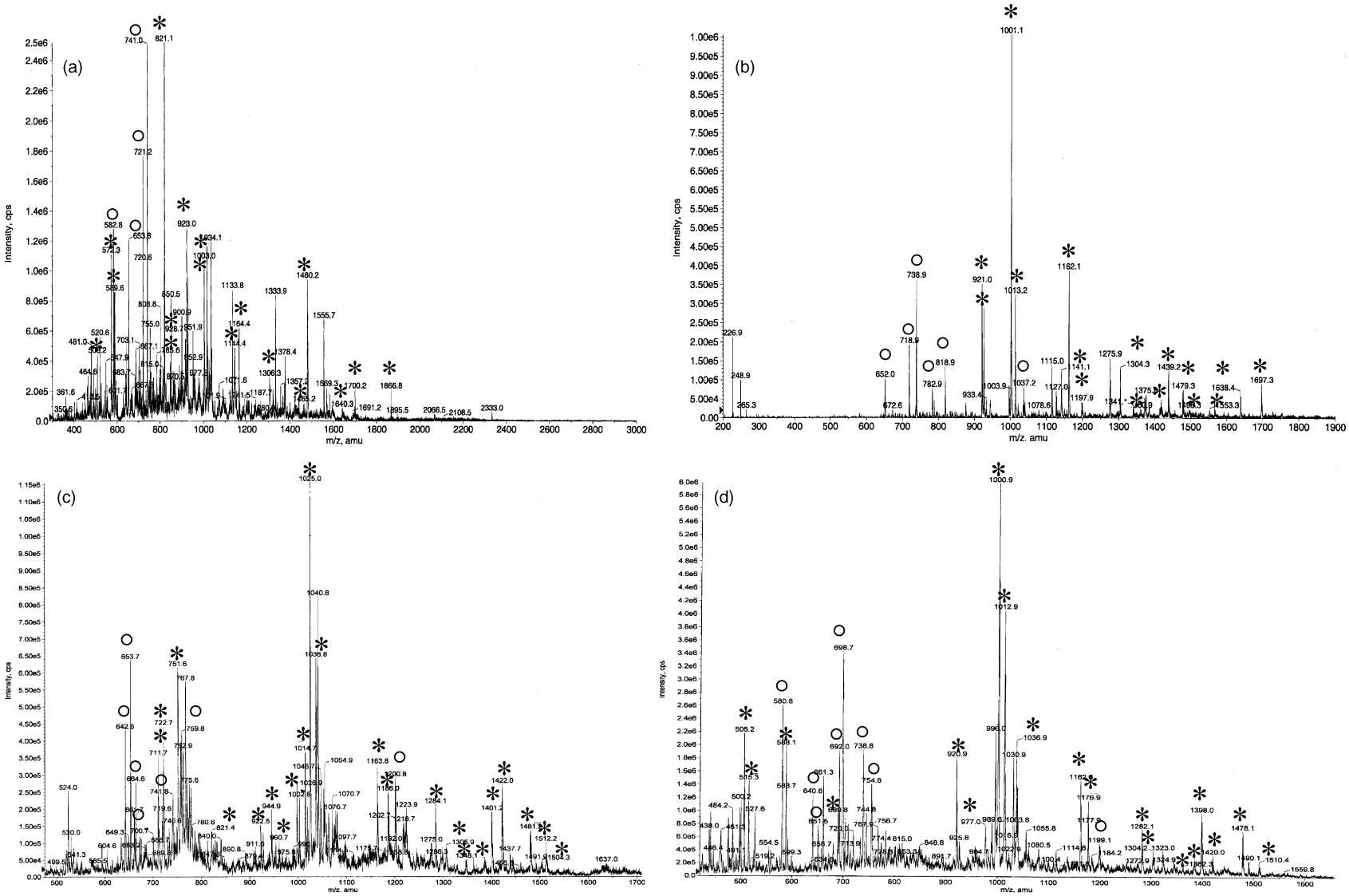

Fig. 3. Nanospray spectra of a bovine serum albumin tryptic digest. Peak legend: (∗) tryptic peptide in ±1 charge state, (○) tryptic peptide in ±2 charge state. (a) Positive ion spectrum eluted in an acidic solvent. (b) Negative ion spectrum in an acidic solvent. (c) Positive ion spectrum in a basic solvent. (d) Negative ion spectrum in a basic solvent. See Table 7 for peak identities.

Table 8
Positive and negative nanospray comparison for digests of cytochrome *c*

		Formic Acid Elution		NH4OH Elution		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive *b*	Negative*a*	(from SWISS-PROT)
GITWKEETLMEYLENPK	56-72	I	W			4.49
TGQAPGFTYTDANKNK	40-55	S (+2)	I (-2)	I (+2)	I (-2)	8.17
IFVQKCAQCHTVEK	9 To 22			I	I	8.06
EETLMEYLENPK	61-72	W	W	I	I	4.09
TGQAPGFTYTDANK	40-53			W	W	5.50
TGPNLHGLFGRK	28-39	W	I			11.00
TGPNLHGLFGR	28-38	S, I (+2)	W, S(-2)			9.44
EDLIAYLK	92-99	W				4.37
NKGITWK	54-60	S				10.00
MIFAGIK	80-86	S	BP	S	S	8.50
YIPGTK	74-79	I	I			8.59
IFVQK	9 To 13	I	I	W		8.75
GITWK	56-60			I		8.75
ATNE	101-104	I	I			4.99
KK	87-88	I				11.00
	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak				
		All Peaks +or - 1 Charge State Unless Noted Otherwise				
Sequence Coverage		69%	67%	50%	45%	
For Both Positive and Negative		69%		50%		
For All 4 Combined		82%				
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)		152, 90	139, 82	76, 48	50, 36	
For Both Positive and Negative		152		76		
For All 4 Combined		178				
		Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy				

a Incomplete digestion, many incorrect hits from intact proteins
b Incomplete digestion, intact protein mostly filtered out

was the protein identified by the database when searched. Lysozyme (Table 10) presented some solubility problems for nanospray, but increasing the acetonitrile content to 75% produced spectra that led to 82% overall sequence coverage and unambiguous identification.

β-Casein (Table 12), as noted previously, is a special case because it contains phosphorylation sites. Each phosphate group shifts the peptide's pI values lower, which makes β-casein ideally suited for the negative ion mode. The positive and negative digest spectra of β-casein are shown in Fig. 4a and b. The negative ion spectrum exhibits peaks at 1561, 1041, and 624 *m*/*z* that correspond to the $[M-2H]^{-2}$, $[M-3H]^{-3}$, and $[M-4H]^{-4}$ states of a tetraphosphorylated peptide. The MS/MS spectrum of 1041 *m*/*z*, shown in Fig. 4c, clearly shows these phosphorylation events, evidenced by the successive losses of 96 or 80 Da (PO_4 and PO_3, respectively). These phosphate groups are lost both as neutrals and as negatively charged groups (the 1520 *m*/*z* peak). The phosphorylated peaks are of much lower intensity (or absent) in the positive ion mode. A low abundance post-translationally modified protein may need to be analyzed in the negative mode in order for it to be detected.

In general, nanospray and MALDI were equally effective at ionizing small peptides with the exception of BSA. Specifically, the number of peptides with 15 amino acids or less detected by either positive or negative ion nanospray is 39 for BSA (versus 22 by MALDI, the only case where there is a significant difference), 13 for cytochrome *c* (also 13 by MALDI), 14 for hemoglobin (13 for MALDI), 11 for lysozyme (12 by MALDI), 12 for myoglobin (also 12 by MALDI), and 6 for β-casein (versus 5 by MALDI). However, higher mass peptides (those 15 amino acids or larger) tended to be observed more readily by MALDI, but that *S*/*N* ratios for most peptides were higher in nanospray mass spectra because of the relatively high background noise in MALDI so that base peaks in MALDI are seldom more than 25:1 *S*/*N*. A summary of the nanospray results is shown in Fig. 5. Although MALDI is more commonly used in peptide mass fingerprinting, here nanospray revealed greater sequence coverage for five of the six proteins (except for β-casein) than MALDI. MOWSE scores also improved for five of the six proteins, with the exception of lysozyme, where the MOWSE score decreased from 209 to 161 although the percentage sequence coverage increased from 66 to 82%. It is interesting to note that β-casein's sequence coverage decreased from 78 to 66% while its MOWSE score increased slightly from 116 to 121. This highlights that unique tryptic peptides can sometimes help improve specificity of a match, even when overall sequence coverage decreases.

A correlation was sought between peptide characteristics and their detection in either the positive or negative ionization mode. At first, it was thought that the overall pI of the peptide would be the dominant factor in determining whether a given peptide was observed in the mass spectrum or not; however, it can be seen from the tables that is not true in all cases. It is worthy of note that what does appear to be an influencing factor is the presence of multiple acidic or basic residues in the peptide. The presence of multiple acidic residues (glutamic acid and aspartic acid), in particular, leads to a strong signal in the negative ion mode, and

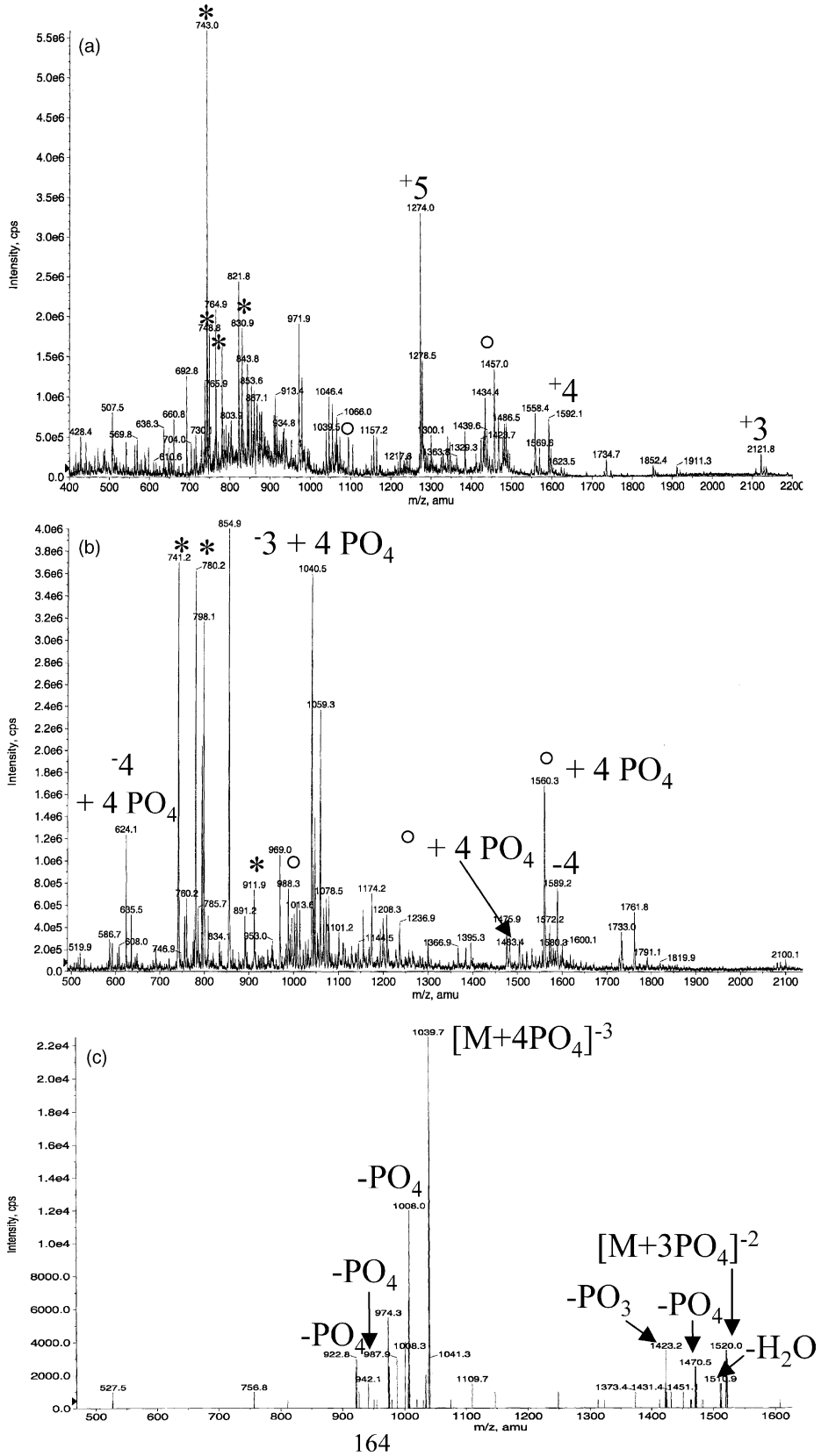

Fig. 4. Nanospray spectra of a β-casein tryptic digest. Peak legend: (∗) tryptic peptide in ±1 charge state, (○) tryptic peptide in ±2 charge state. (a) Positive ion spectra in an acidic solvent. (b) Negative ion spectrum in a basic solvent. (c) MS/MS spectrum of the 1040 *m*/*z* peak of a β-casein tryptic digest eluted in a basic solvent. This is the triply charged quadruply phosphorylated peak of the residues 16–40. Losses of 32 *m*/*z* for the $[M + 4PO_4]^{-3}$ form and 48 *m*/*z* for the $[M + 3PO_4]^{-2}$ form equate to 96 Da and are indicative of a phosphate group. See the text and Table 12 for further details.

Table 9
Positive and negative nanospray comparison for digests of α- and β-hemoglobin

Peptide Sequence	Residue Numbers	Formic Acid Elution		NH4OH Elution		Theoretical pI (from SWISS-PROT)
		Positive	Negative	Positive	Negative	
VGDALTLAVGHLDDLPGALS NLSDLHAHKLR	62-92 a				W	5.72
KVGDALTLAVGHLDDLPGAL SNLSDLHAHK	61-90 *a*			W	W	5.73
VGDALTLAVGHLDDLPGALSNLSDLHAHK	62-90 *a*	I (+2)	W (-2)	I (+2)	I (-2), S (-3)	5.10
VNEEEVGGEALGRLLVVYPWTQR	18-40 *b*				W (-2)	4.49
LHVDPENFRLLGNVLVVVLAR	96-116 *b*			I (+2), S (+3)		6.75
AAVLALWDKVNEEEVGGEALGR	9-30 *b*	S (+2)	BP (-2)	S (+2), I (+3)	S (-2), I (-3), S (-4)	4.25
AAWSKVGGHAGEFGAEALER	12-31 *a*			I (+2)	W (-2)	5.50
FFDSFGDLSNPGAVMGNPK	41-59 *b*	I, BP (+2)	I, S (-2)	W, S (+2), I (+3)	W, BP (-2), S (-3)	4.21
HFGKDFTPELQASYQK	117-132 *b*			S	I	6.75
VQLSGEEKAAVLALWDK	1-17 *b*			W, S (+2)		4.68
TYFPHFDLSHGSAQVK	41-56 *a*	W	W	W, S (+2), W (+3)	W, I (-2)	6.61
VLHSFGEGVHHLDNLK	67-82 *b*			W (+2)	W, W (-2)	6.26
VGGHAGEFGAEALER	17-31 *a*	W	W	I (+2)	W (-2)	4.75
VVAGVANALAHKYH	133-146 *b*	W	I	W	W	8.58
DFTPELQASYQK	121-132 *b*	I	I	W	W	4.37
GTFAALSELHCDK	83-95 *b*	W	I		W	5.32
LLVVYPWTQR	31-40 *b*	S	I	BP, S (+2)	I, W (-2)	8.75
LLGNVLVVVLAR	105-116 *b*	I	I	I	W	9.75
VLSAADKTNVK	1-11 *a*			I, I (+2)	I, I (-2)	8.56
LRVDPVNFK	91-99 *a*			W		8.75
MFLGFPTTK	32-40 *a*	S	S	S, S (+2)	S	8.50
AAVLALWDK	9-17 *b*	S	I	I		5.88
VQLSGEEK	1-8 *b*	S	I		I	4.53
VKAHGK	60-65 *b*	S	I			10.00
AAWSK	12-16 *a*	W				8.80
HFGK	117-120 *b*	W				8.50
	a = Alpha Subunit					
	b = Beta Subunit					
	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak				
		All Peaks +or - 1 Charge State Unless Noted Otherwise				
Sequence Coverage		52% Alpha	48% Alpha	66% Alpha	62% Alpha	
		82% Beta	79% Beta	86% Beta	95% Beta	
For Both Positive and Negative		52% Alpha, 82% Beta		67% Alpha, 95% Beta		
For All 4 Combined		67% Alpha, 99% Beta				
MOWSE Score (Matched Peaks Only)	(Alpha)	82 Alpha	68 Alpha	137 Alpha	145 Alpha	
	(Beta)	157 Beta	144 Beta	178 Beta	195 Beta	
For Both Positive and Negative		82 Alpha, 157 Beta		159 Alpha, 226 Beta		
For All 4 Combined (Alpha + Beta Treated Separately)		173 Alpha, 253 Beta				
All 4 Combined		187 Beta				
		Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy				

a weaker or nonexistent signal in the positive ion mode (by MALDI or nanospray), although there are exceptions to this as well. It was also observed that an acidic solvent favors the appearance of low pI peptides and the basic solvent favors high pI peptides, though peptides with pI 6–8 are often observed in either acidic or basic solvent. Other factors, such as whether the peptide is on the surface or the core protein, and peptide hydrophobicity may also play a role.

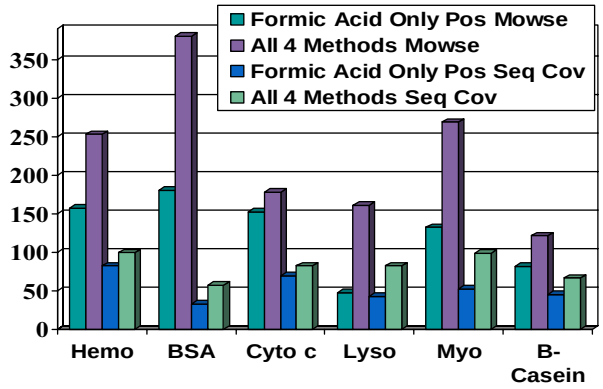

Fig. 5. Summary of MOWSE score and percent sequence coverage obtained for the nanospray experiments. Each protein exhibits an increase when the negative ion mode is also taken into consideration.

3.3. Tandem mass spectrometry of protein digests using polyaniline coated emitters

One of the main reasons that nanospray analysis for proteomics is performed in the positive mode is the commonly held belief that tandem mass spectrometry in the negative mode with nanospray ionization is unreliable. This stems from the poor emitter performance that was discussed in the introduction. The high risk of electrical discharge in the negative ion mode with traditional metal-coated emitters prevented MS/MS from being a viable option. However, with new technologies, such as polyaniline coated emitters, the negative ion mode has become much more stable [22]. Not only can negative ion MS/MS be used reliably for detection and localization of highly negative functional groups

Table 10
Positive and negative nanospray comparison for digests of lysozyme

		Formic Acid Elution		NH4OH Elution		Theoretical pI
Peptide Sequence	Residue Numbers	Positive *a*	Negative *b*	Positive	Negative	(from SWISS-PROT)
NLCNIPCSALLSSDITASVNCAKK	92-115			I (+2)	S (-2)	7.95
NLCNIPCSALLSSDITASVNCAK	92-114	I	I		S (-2)	5.82
KIVSDGNGMNAWVAWR	115-130		W	BP (+2)	W, BP (-2)	8.75
NTDGSTDYGILQINSR	64-79	S	BP		I (-2)	4.21
IVSDGNGMNAWVAWR	116-130		W, I (-2)	W	S (-2)	5.84
WWCNDGRTPGSR	80-91	S	I	W, S (+2)	W, S (-2)	8.25
FESNFNTQATNR	52-63	S (+2)	I, S (+2)			6.00
VFGRCELAAAMK	20-31				W	8.19
CKGTDVQAWIR	133-143			I (+2)	I (-2)	8.22
GYSLGNWVCAAK	40-51		I		I (-2)	8.20
GTDVQAWIR	135-143		I	I	I	5.84
RHGLDNYR	32-39		I			8.75
CELAAAMK	24-31		S			5.99
KVFGR	19-23				I	11.00
VFGR	20-23				I	9.20
	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak				
		All Peaks + or - 1 Charge State Unless Noted Otherwise				
Sequence Coverage		42%	78%	42%	69%	
For Both Positive and Negative		78%		69%		
For All 4 Combined		82%				
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)		47, No Match	114, 77	69, No Match	128, 74	
For Both Positive and Negative		114		128		
For All 4 Combined		161				
		Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy				

a Incomplete digestion, many incorrect hits from intact proteins
b Incomplete digestion, intact protein mostly filtered out

(Fig. 4c) but it also can be used to sequence peptides and proteins. There has been precedence for sequencing in the negative mode with ESI in the literature [25], but nanospray has been infrequently used. To demonstrate the capabilities of polyaniline coated nanospray emitters, the six proteins used in the study were subjected to digestion followed by product ion MS/MS analysis in the negative ion mode. Attaining a high level of signal for the MS/MS study was a

Table 11
Positive and negative nanospray comparison for digests of myoglobin

		Formic Acid Elution		NH4OH Elution		Theoretical pI
Peptide Sequence	Residue Numbers	Positive	Negative	Positive	Negative	(from SWISS-PROT)
GLSDGEWQQVLNVWGKVEADIAGHGQEVLIR	1 To 31	I (+2)	I (-2)			4.50
HPGDFGADAQGAMTKALELFR	119-139				W (-2)	5.38
KGHHEAELKPLAQSHATK	79-96				I (-3H+Na)	8.52
LFTGHPETLEKFDKFK	32-47	I (+2)	I (-2)	W (+2)	I (-2)	6.76
YLEFISDAIIHVLHSK	103-118			I (+2)	S (-2)	5.99
GHHEAELKPLAQSHATK	80-96			S (+2)	S (-2)	7.03
GLSDGEWQQVLNVWGK	1 To 16	I, S (+2)	I, S (-2)	W, W (+2)	W, BP (+2)	4.37
LFTGHPETLEKFDK	32-45			I (+2)	W, W (-2)	5.45
VEADIAGHGQEVLIR	17-31	W, S (+2)	W, S (-2)	S (+2)	I, S (-2), I (-3)	4.65
HPGDFGADAQGAMTK	119-133				I	5.21
HGTVVLTALGGILK	64-77	I, BP (+2)	I	W, BP (+2)	S	8.76
ALELFRNDIAAK	134-145	I, S (+2)	I, I (-2)			6.11
LFTGHPETLEK	32-42	W	I	W, I (+2)	I, I (-2)	5.40
HLKTEAEMK	48-56		W (-2)	W	W	6.76
YKELGFQG	146-153	S	BP		S	6.00
ASEDLKK	57-63				I	6.11
ALELFR	134-139	S	S	S	S	6.05
ELGFQG	148-153		I		W	4.00
IPIK	99-102				W	9.00
	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak				
		All Peaks + or - 1 Charge State Unless Noted Otherwise				
Sequence Coverage		52%	80%	71%	94%	
For Both Positive and Negative		80%		94%		
For All 4 Combined		98%				
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)		132, 79	198, 115	140, 74	237, 181	
For Both Positive and Negative		198		237		
For All 4 Combined		269				
		Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy				

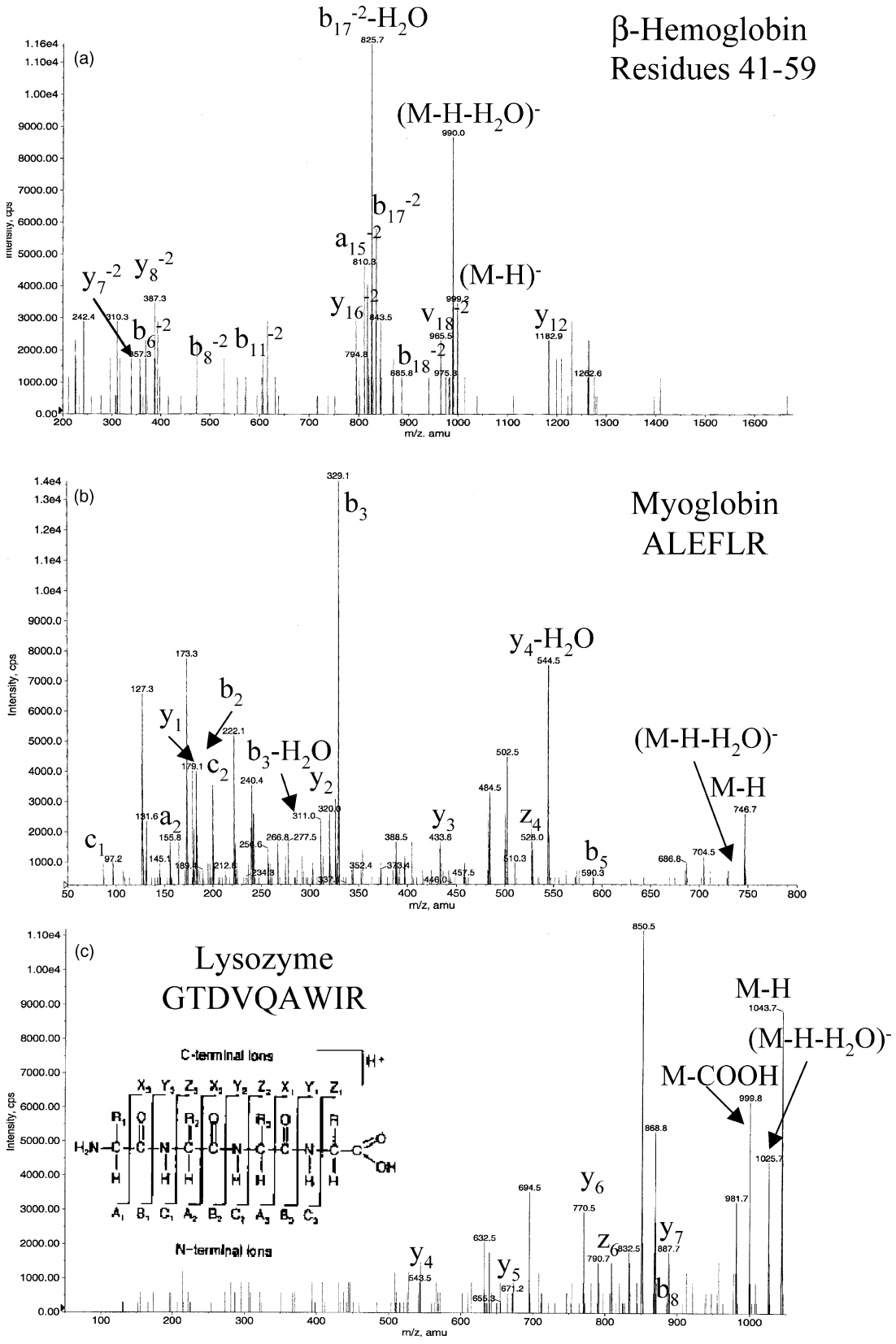

Fig. 6. Negative ion nanospray product ion MS/MS spectra from tryptic digest samples. (a) Hemoglobin, doubly charged ion at *m*/*z* 999.2, (b) myoglobin, singly charged ion at *m*/*z* 746.7, and (c) lysozyme, singly charged ion at *m*/*z* 1045.

difficulty, but only because the triple quadrupole instrument has a spray voltage limitation of −4500 V for negative ion MS/MS whereas the limit is −5500 V for a Q_1 scan. That led to decreased ionization efficiency so only the most intense peaks yielded quality MS/MS spectra. However, no electrical discharge occurred at any time, and the signal remained steady throughout the course of the experiments. Examples of the data obtained from hemoglobin, myoglobin, and lysozyme are shown in Fig. 6a–c. All three spectra exhibit the characteristic b- and y-series fragmentation pattern typical of peptides, which led to their proper identification from human interpretation of their mass spectra. The

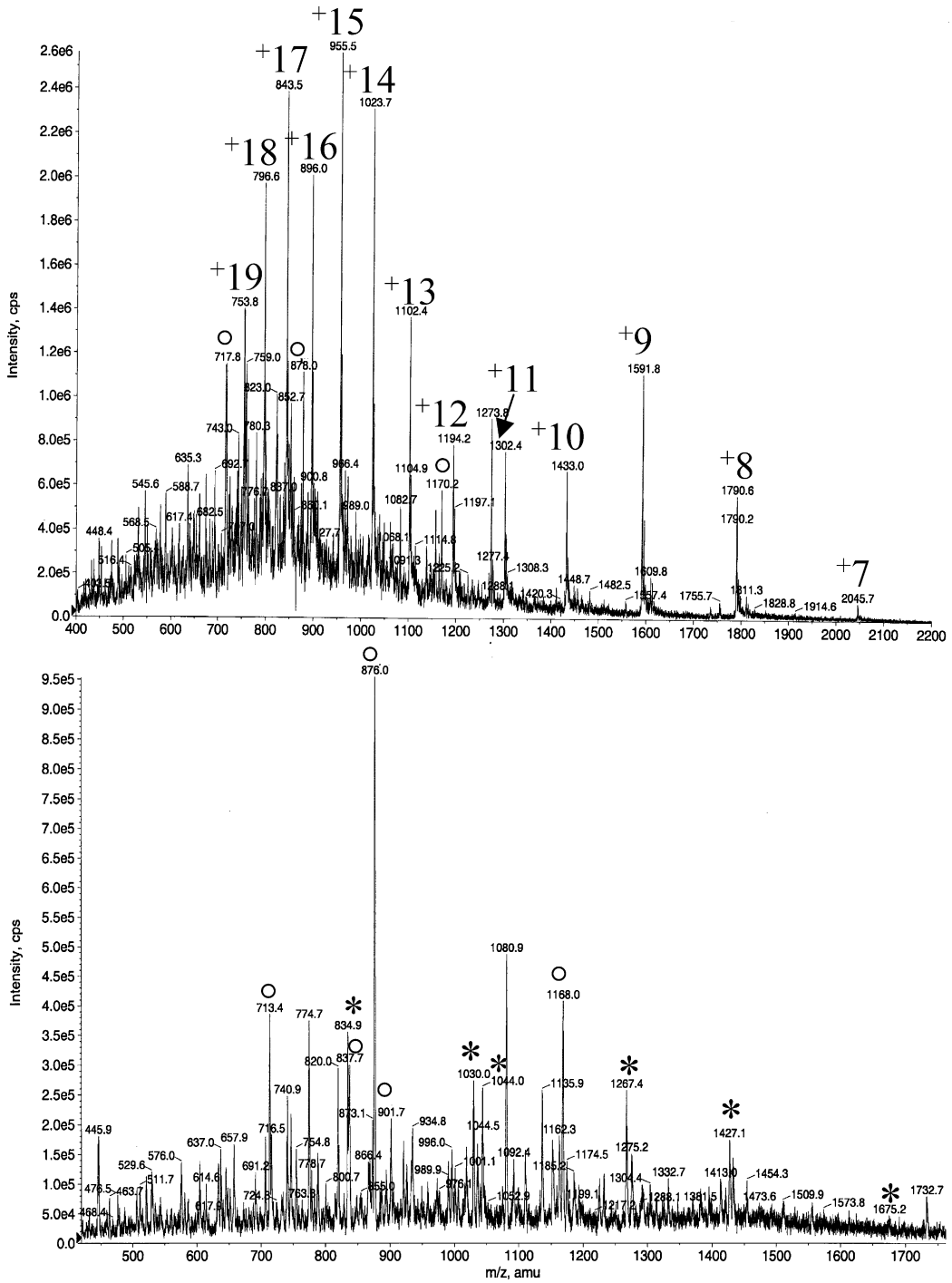

Fig. 7. Nanospray spectra obtained from an incomplete tryptic digest of lysozyme after elution in an acidic solvent as described in the text and in Table 10. Peak legend: (∗) tryptic peptide in ±1 charge state, (○) tryptic peptide in ±2 charge state, $^{+X} = [M + XH]^{+X}$. (a) Positive ion spectrum, intact protein is still present. (b) Negative ion spectrum, only tryptic peptides are detected.

MASCOT database could not generate a significant match because of its inability to search negative ion MS/MS spectra.

Negative ion nanospray MS/MS is a feasible technique, when the correct technology is used. However, because of its limited acceptance, the software required to fully exploit it has not been developed. If the usage of negative ion nanospray MS/MS goes up, this situation is likely to change.

3.4. Filtering incomplete digestions

Proteins may not always fully digest due to inaccessible sites, high levels of folding, insolubility, or even poor

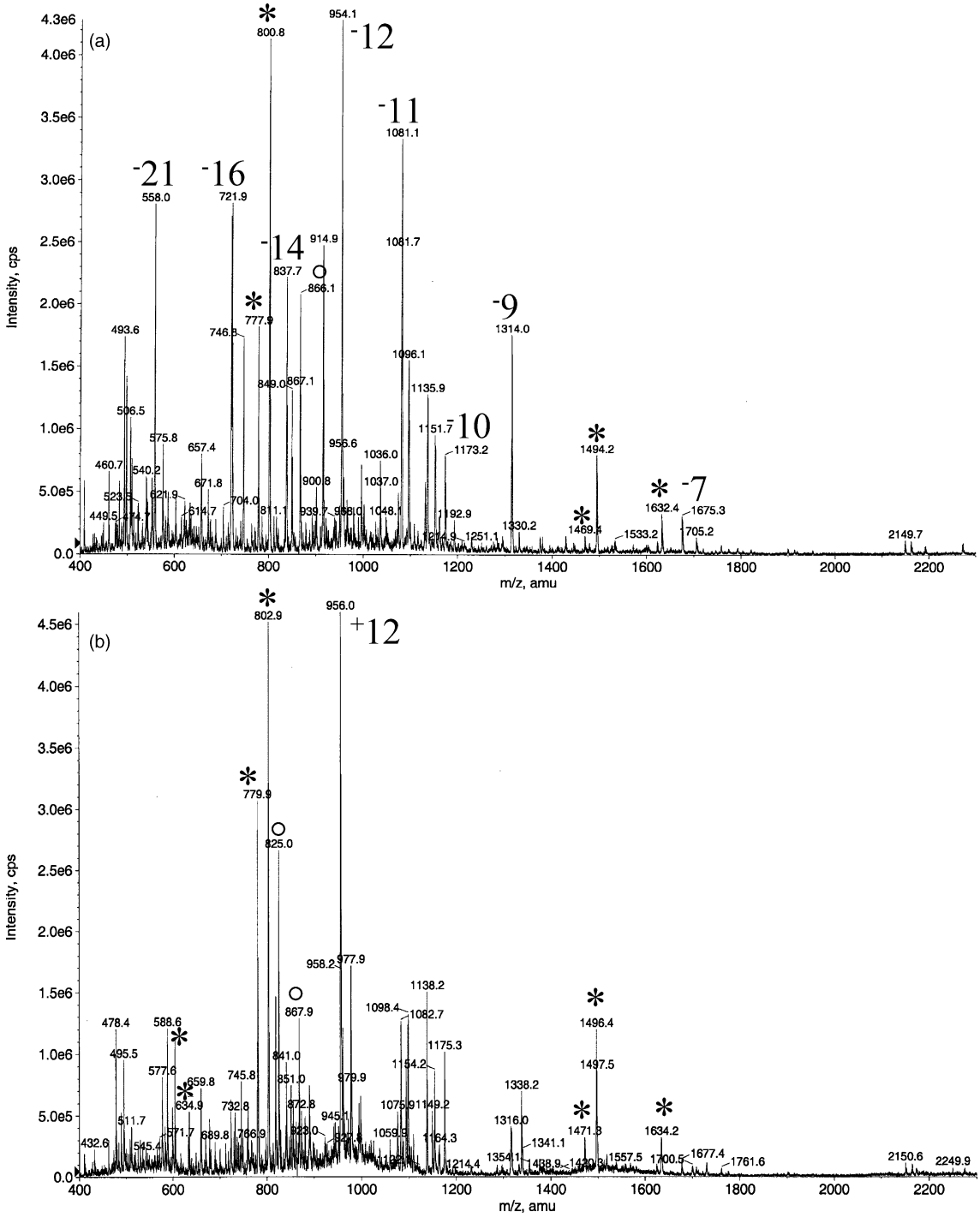

Fig. 8. Nanospray spectra obtained from an incomplete tryptic digest of cytochrome *c* after elution in a basic solvent as described in the text and Table 8. Peak legend: (∗) tryptic peptide in ±1 charge state, (○) tryptic peptide in ±2 charge state, $^{-}X = [M - XH]^{-X}$. (a) Negative ion spectrum, intact protein still present. (b) Positive ion spectrum, intact protein no longer detected but tryptic peptides are.

Table 12
Positive and negative nanospray comparison for digests of β-casein

Peptide Sequence	Residue Numbers	Formic Acid Elution Positive	Formic Acid Elution Negative	NH4OH Elution Positive	NH4OH Elution Negative	Theoretical pI (from SWISS-PROT)
YPVEPFTESQSLTLTDVENLHLPLPLLQSWM						
HQPHQPLPPTVMFPPQSVLSLSQSK	129-184	W (+4), S (+5)	W (-3)		I (-4)	5.26
RELEELNVPGEIVES*LS*S*S*EESITR	16-40			I (+2), I (+3), W (+4)	S (-2), S (-3), S (-4)	2.65
ELEELNVPGEIVES*LS*S*S*EESITR	17-40				W (-2)	2.36
DMPIQAFLLYQEPVLGPVRGPFPIIV	199-224	I (+2)	I (-2)			4.37
DMPIQAFLLYQEPVLGPVR	199-217	W	I			4.37
FQS*EEQQQTEDELQDK	48-63		I (-2)			3.27
FQSEEQQQTEDELQDK	48-63				I (-2)	3.77
EAMAPKHK	115-122				I	8.69
AVPYPQR	192-198	I	W			8.79
VLPVPQK	185-191	I	I		S	8.72
EMPFPK	123-128	S	S		W	6.10
GPFPIIV	218-224	BP	BP	S	S	5.52
EAMAPK	115-120		W			6.10
* = Phosphorylation Site	Intensities:	BP = Base Peak S=Strong I=Intermediate W=Weak All Peaks + or - 1 Charge State Unless Noted Otherwise				
Sequence Coverage		45%	55%	14%	55%	
For Both Positive and Negative		55%		55%		
For All 4 Combined		66%				
MOWSE Score (Matched Peaks Only, All Peaks In Spectrum)		81, 40	101, 52	No Match	90, 41	
For Both Positive and Negative		101		90		
For All 4 Combined		121				
		Constrictions: +-1.5 Da Tolerance, Metazoa Taxonomy				

reactivity with trypsin. However, there is still valuable information that can be obtained in the mass spectra of such digestions. The nanospray spectrum of an incomplete digestion of lysozyme nanosprayed in an acidic solvent and detected in the positive ion mode is shown in Fig. 7a. The intact protein dominates the mass spectrum, but some tryptic peptides are evident. If the same sample is run in the negative ion mode (Fig. 7b), the intact protein is no longer detected and the tryptic peptides now dominate the mass spectrum, providing enough information for the protein to be identified. This is due to the fact that the protein is highly positively charged in the acidic solvent, so multiply-charged ions are selected against when changing the polarity negative. The tryptic peptides themselves on average are much less positively charged than the entire protein, so it is much more facile for these peptides to generate negative ions. This phenomenon also applies to highly negatively charged proteins. Fig. 8a is an incomplete digestion of cytochrome *c* in a basic solvent analyzed in the negative mode. The intact protein once again dominates the spectrum. When analyzed in the positive ion mode (Fig. 8b) the intact protein is selected against and not detected, but some tryptic peptides are detected. The sequence coverage is not as great as in a more complete digestion, but the mass spectral quality is improved enough to identify the protein. This methodology could be advantageous when sample is limited. There may not be enough sample to rerun the digest or another sample may not be available. Another limiting factor is time. Throughput issues may not allow for a delay of 6–18 h for a re-digestion and this procedure could circumvent that.

4. Conclusions

Using MALDI or nanospray in the negative ion mode, peptides from digests rich in aspartic acid and glutamic acid residues that are not observed in the positive ion mode were detected. By combining data from the positive ion and negative ion modes, better sequence coverage was obtained (particularly in nanospray) and higher MOSWE scores were attained for the proteins analyzed. The limitations of database search programs were also addressed. Future algorithms should take into account multiple charging and salt adductions. In the case of the nanospray experiments, the mass spectra were much cleaner in the negative ion mode. The presence of an acidic or basic solvent also has an influence on the appearance of certain peptides in the mass spectrum. Incompletely digested proteins were filtered from their tryptic digest peptides by changing the ionization polarity. A highly positively charged protein will not be detected in the negative ion mode, but a slightly net positively charged tryptic peptide can form ions with negative charge, and thus be detected in the negative ion mode. This could be important when sample quantities are limited which prevents the sample from being re-digested.

Analyzing protein mixtures in both the positive ion and negative ion modes with nanospray simplifies the interpretation and reduces unwanted peaks in protein digest mass spectra. This could make nanospray a viable alternative to MALDI for peptide mass fingerprinting, especially when sequencing unknown peptides, since post source decay is neither a trivial nor a sensitive technique. Using both the positive and negative ion mode in the sequencing

of bovine brain tubulin has thus far yielded 25% of the protein sequence, with approximately half the identified peptides being detected by each mode. Regardless of ionization technique used, these data demonstrate the value of combined positive ion and negative ion mode peptide mapping, and we urge the proteomics community to include negative ion mode mass spectrometry for proteomics analysis.

Acknowledgements

P.M.B. thanks Thomas White for assistance in fabricating the nanospray emitters as well as Douglas Smith and Gary Sagerman for the design and construction of the nanospray source. Additionally, the authors thank the UB Department of Pharmaceutics LCMS facility and UB Department of Biochemistry for usage of the ESI-Quadrupole and MALDI-Tof instruments, respectively. Funding from the NIH (subcontract of R43RR016399-01 to Marie A. Moy), the NSF (CHE-0094961), and UB is also greatly appreciated.

References

[1] M. Yamashita, J.B. Fenn, J. Phys. Chem. 88 (1984) 4451.
[2] M. Karas, D. Bachmann, F. Hillenkamp, Anal. Chem. 57 (1985) 2935.
[3] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, Rapid Commun. Mass Spectrom. 2 (1988) 151.
[4] M. Karas, F. Hillenkamp, Anal. Chem. 60 (1988) 2299.
[5] W.J. Henzel, T.M. Billeci, J.T. Stults, S.C. Wong, C. Grimley, C. Watanabe, Proc. Natl. Acad. Sci. U.S.A. 90 (1993) 5011.
[6] M.S. Wilm, M. Mann, Int. J. Mass Spectrom. Ion Process. 136 (1994) 167.
[7] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal. Chem. 63 (1991) 1193A.
[8] M. Senn, R. Venkataraghavan, F.W. McLafferty, J. Am. Chem. Soc. 88 (1966) 5593.
[9] W.R. Mathews, R.S. Johnson, K.L. Cornwell, T.C. Johnson, B.B. Buchanan, K. Biemann, J. Biol. Chem. 262 (1987) 7537.
[10] R. Aebersold, D.R. Goodlet, Chem. Rev. 101 (2001) 269.
[11] A. Pandey, M. Mann, Nature 405 (2000) 837.
[12] D.N. Perkins, D.J. Pappin, D.M. Creasy, J.S. Cottrell, Electrophoresis 20 (1999) 3551.
[13] P. Berndt, U. Hobohm, H. Langen, Electrophoresis 20 (1999) 3521.
[14] W.S. Hancock, S.Z. Wu, P. Shieh, Proteomics 2 (2002) 352.
[15] S.L. Cohen, B.T. Chait, Anal. Chem. 68 (1996) 31.
[16] G.J. Opiteck, K.C. Lewis, J.W. Jorgenson, R.J.A.C. Anderegg, Anal. Chem. 69 (1997) 1518.
[17] E.C. Koc, W. Burkhart, K. Blackburn, A. Moseley, L.L.J.B.C. Spremulli, J. Biol. Chem 276 (2001) 19363.
[18] A.L. McCormack, D.M. Schieltz, B. Goode, S. Yang, G. Barnes, D. Drubin, J.R. Yates, Anal. Chem. 69 (1997) 767.
[19] R.D. Smith, J.H. Wahl, D.R. Goodlett, S.A. Hofstadler, Anal. Chem. 65 (1993) 574A.
[20] K. Vanhoutte, W. Van Dongen, I. Hoes, F. Lemiere, E.L. Esmans, H. Van Onckelen, E. Van den Eeckhout, R.E.J. van Soest, A.J. Hudson, Anal. Chem. 69 (1997) 3161.
[21] J.J. Gaumet, G.A. Khitrov, G.F. Strouse, NanoLetters 2 (2002) 375.
[22] P.M. Bigwarfe Jr., T.P. White, T.D. Wood, Rapid Commun. Mass Spectrom. 16 (2002) 2266.
[23] F. Zappacosta, M.J. Huddleston, R.L. Karcher, V.I. Gelfand, S.A. Carr, R.S. Annan, Anal. Chem. 74 (2002) 3221.
[24] A. Lim, T. Prokaeva, M.E. McComb, P.B. O'Connor, R. Theberge, L.H. Connors, M. Skinner, C.E. Costello, Anal. Chem. 74 (2002) 741.
[25] J.A. Loo, R.R. Loo, K.J. Light, C.G. Edmonds, R.D. Smith, Anal. Chem. 64 (1992) 81.
[26] A.M. Lawson, W. Chai, G.C. Cashmarc, E.F. Hounsell, T. Feizi, Carbohydr. Res. 200 (1990) 47.
[27] W. Chai, J. Luo, C.K. Lim, A.M. Lawson, Anal. Chem. 70 (1998) 2060.
[28] D.R. Smith, G. Sagerman, T.D. Wood, Rev. Sci. Instrum. 74 (2003) 4474.
[29] R.D. Appel, A. Bairoch, D.F. Hochstrasser, Trends Biochem. Sci. 19 (1994) 258.
[30] D.J.C. Pappin, H.F. Hansen, M. Bartlet-Jones, W. Jeffery, A.J. Bleasby, Mass Spectrom. Biol. Sci. (1996) 135–150.
[31] G. Li, M. Waltham, N.L. Anderson, E. Unsworth, A. Treston, J.N. Weinstein, Electrophoresis 18 (1997) 391.
[32] J. Gobom, M. Schuerenberg, M. Mueller, D. Theiss, H. Lehrach, E. Nordhoff, Anal. Chem. 73 (2001) 434.
[33] P. Chaurand, R.M. Caprioli, Electrophoresis 23 (2002) 3125.